

**Application of machine learning
classifiers to arterial disease
detection, utilising virtual patient
databases**

Gareth R. Jones

Submitted to Swansea University in fulfilment of
the requirements for the Degree of Doctor of
Philosophy.

Swansea University
2021

Summary


Two of the most common forms of arterial disease are stenosis and aneurysm, estimated to affect between 1% and 20% of the population. Ruptured abdominal aortic aneurysms alone are estimated to be the cause of between 6,000 and 8,000 deaths a year within the United Kingdom. Patients with stenosis have been shown to have a mortality hazard ratio of 1.42 compared to a control population [2], and an unadjusted death rate of 3.35 per 100 person-years compared to 1.23 per 100 person-years in a control population [97]. Current methods for the detection of arterial disease are generally impractical for large scale screening, expensive, or both. If an inexpensive method for the detection of both stenosis and aneurysm is created, that minimises the need for invasive measurements, the cost effectiveness of large scale screening could be improved making both continuous monitoring and screening feasible. One such method is to use easily acquirable haemodynamic measurements at accessible peripheral locations within the circulatory system for diagnosis. Within this thesis an initial exploratory study into the potential of using machine learning classification algorithms to detect arterial disease from such measurements is presented.

It is likely that the indicative biomarkers of arterial disease held within pressure and flow-rate profiles consist of micro inter- and intra- measurement details. To facilitate the use of a data driven approach to the discovery of any biomarkers a framework for the creation of virtual patients, through the employment of a mathematical model of blood flow, is presented. This framework is utilised to create a series of virtual patient databases, as the balance between simplicity and realism progresses through the thesis. The most realistic of these databases is made publicly available (<https://doi.org/10.5281/zenodo.4549764>). The aforementioned framework for the creation of virtual patients is a major contribution of this thesis, and can be applied to a wide range of biological systems given a mathematical description.

The synthetic data sets are used to train and subsequently test a series of machine learning classifiers, to predict the presence of both stenosis and aneurysm, using various combinations of pressure and flow-rate measurements. It is shown that the inclusion of a diseased vessel (either stenosis or aneurysm) produces consistent and significant biomarkers in haemodynamic profiles, irrespective of a patients unique underlying arterial network. These biomarkers are found to be differentiable from the natural variability present across a large cohort of patients, showing that arterial disease has a clear and unique effect on pressure and flow-rate profiles. This suggests strong potential in the use of haemodynamic measurements to detect arterial disease.

Declarations

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed:  **Date: 29 September 2021**

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed:  **Date: 29 September 2021**

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed:  **Date: 29 September 2021**

The University's ethical procedures have been followed and, where appropriate, that ethical approval has been granted.

Signed:  **Date: 29 September 2021**

Acknowledgements

I would like to thank my supervisors Dr Sanjay Pant, Dr Jim Parr, and Dr Perumal Nithiarasu, whose support, guidance, and expertise has been fundamental to the completion of my doctorate. Special mention, however, must be given to Dr Sanjay Pant who has continuously gone above and beyond his duty and responsibility.

I am very grateful to have completed my doctorate in such a supportive professional and social circle. I would like to thank all the members of the Biomedical engineering department and the occupants of the office A130 for their interest, input, and friendship.

Finally, I would thank all my friends and family for their unwavering support throughout the entirety of my education.

Contents

Summary	i
Declarations	ii
Acknowledgements	iii
Peer-reviewed papers published during PhD candidature	xi
Conference presentations during the PhD Candidature	xii
Common Symbols and Abbreviations	xxxviii
1 Introduction	1
1.1 Arterial disease	2
1.2 Use of haemodynamic measurements for disease detection	4
1.3 Machine learning classification algorithms and their application to arterial disease detection	6
1.4 Aims and objectives	9
1.5 Outline of Thesis	10
2 Important methodologies	13
2.1 Physics based model of pulse wave propagation	13
2.1.1 Inlet boundary conditions	15
2.1.2 Outlet boundary conditions	15
2.2 Bayes' theorem	16
2.3 Metropolis-Hastings Markov chain Monte-Carlo	17
2.3.1 Pre-fetching	19
2.4 Machine learning classifiers	21
2.4.1 Naive Bayes' classifier	22
2.4.2 Logistic regression	23
2.4.3 Multi-layer Perceptron	26
2.4.4 Support vector machine	28
2.4.5 Random forest	31
2.4.6 Gradient boosting classifier	32

2.4.7	Motivation for the choice of methods	34
3	Three vessel network proof-of-concept study	36
3.1	Introduction	36
3.2	Virtual patient database	37
3.2.1	Motivation and procedure	37
3.2.2	Topology and parameterisation	39
	Healthy subjects	39
	Disease subjects	42
3.2.3	Probability distributions	42
	Healthy subjects	43
	Disease subjects	43
3.2.4	Sampling to create the VPD	44
	Representation of haemodynamic profiles	44
3.2.5	Post processing	45
3.3	Available pressure and flow-rate measurements	46
3.4	Machine learning set up	47
3.4.1	Test/train split	47
3.4.2	Standardisation of input data	47
3.4.3	Machine learning algorithms	48
3.4.4	Required size of the VPD	49
3.4.5	Classifier configurations	50
	Binary configurations	50
	Multiclass ML configurations	51
3.4.6	Quantification of results	55
3.4.7	ML implementation	57
3.4.8	Hyper-parameter optimisation	58
	LR, SVM, and NB	58
	Random forest	59
3.5	Results and discussion	60
3.5.1	<i>A posteriori</i> analysis of the size of the VPD	60
3.5.2	ENBC results	63
	Like for like input measurement comparison.	64
	Effect of number of input measurements	65

	Importance of inlet pressure and flow-split	69
	Linear versus non-linear partitions	71
	Effect of disease severity	74
3.5.3	IVBC results	76
3.5.4	Multiclass analysis	77
	Comparison of the accuracy of different classifications	78
	Importance of number of input measurements	79
	Importance of pressure at the inlet of the system	81
3.5.5	Custom probabilistic configuration analysis.	83
3.6	Conclusions	87
3.7	Limitations	88
4	Creation of a physiologically realistic virtual patient database	90
4.1	Introduction	90
4.2	Key criteria of the new VPD	92
4.3	One-dimensional solver	93
4.4	Arterial network topology	101
4.4.1	Initial network	101
4.4.2	Important locations	102
	Locations of measurements	102
	Locations of disease	106
4.4.3	Network reduction	107
4.5	Parameterisation of the arterial network	118
4.5.1	Mechanical properties.	128
	Exponential property profiles	128
	Hierarchical assignment of vessel properties	129
4.5.2	Geometric properties	132
	Arterial vessel reference radius	132
	Arterial vessel length	133
4.5.3	Boundary conditions	133
	Inlet flow-rate	133
	Terminal lumped models	134
4.5.4	Final network summary	134
4.5.5	Scaling with respect to the reference network	135

4.6	Statistical modelling	136
	Classical inversion problem	137
	Bayesian formulation	138
4.6.1	Prior distributions	139
4.6.2	Likelihood	142
	Scalar pressure and flow-rate measurements	143
	Time varying inlet flow-rate and cardiac output measurement.	144
	Vessel length scaling term measurement	151
4.7	Sampling from the posterior distribution	152
4.8	Results and discussion	153
4.8.1	Burn-in analysis and chain diagnostics	153
4.8.2	Posterior vs. literature reported measurements	158
4.8.3	Evaluation of individual VPs	165
4.8.4	Analysis of resistance and compliance ratios of opposite extremities	168
4.9	Conclusions	170
5	Pre-processing and creation of literature advised unhealthy VPDs	172
5.1	Introduction	172
5.2	Application of physics advised filters to the healthy VPD	173
5.3	Creation of unhealthy VPDs	175
5.3.1	Disease forms	175
5.3.2	Severities of diseases	177
	Stenosis severities	177
	Aneurysm severities	177
5.3.3	Parameterisation of diseased vessels	179
5.4	Results and discussion	180
5.4.1	Evaluation of pressure and flow-rate profiles measured within individual VPs.	180
5.4.2	Analysis of scalar pressure and flow-rate distributions	182
5.5	Conclusions	187
6	Application of machine learning classifiers to the physiologically realistic VPDs	190

6.1	Introduction	190
6.2	Methodology	191
6.2.1	Available measurements	191
	Provision of measurements to ML classifiers	193
6.2.2	Machine learning classifiers	193
	Training and test sets	194
	ML methods	194
	Quantification of results	195
6.2.3	Adequacy of the number of VPs within the VPD	195
6.2.4	Hyper-parameter optimisation	196
	LR, SVM, and NB	196
	Random Forest	197
	Gradient Boosting	198
	Multi-layer perceptron	199
6.3	Results and discussion	200
6.3.1	Input measurement grid search	200
	Measurement combinations	201
	CAS classification:	206
	SAS classification:	210
	PAD classification:	210
	AAA classification:	213
	Importance of carotid artery flow-rate	216
6.3.2	Feature importance	217
6.3.3	Lower severity aneurysms	223
6.3.4	Unilateral aneurysm measurement tests	224
6.3.5	MLP over-fitting	226
6.4	Conclusions	230
7	Use of physics advised haemodynamic features	234
7.1	Introduction	234
7.2	Methodology	235
7.2.1	Physics advised features	235
	Ankle-brachial pressure index	235
	Pulse wave velocity	236

	Ballistocardiogram waveforms	238
7.2.2	Available measurements	241
7.2.3	Machine learning classifiers	242
	Training, validation, and test sets	242
	ML methods	244
	Quantification of results	244
7.2.4	Hyper-parameter optimisation	244
	ABPI	245
	PWV	246
	BCG	247
7.3	Results and discussion	247
7.3.1	ABPI	247
7.3.2	PWV	249
7.3.3	BCG	251
7.4	Conclusions	253
8	Conclusions and contributions	265
8.1	Haemodynamic affects of arterial disease	265
8.2	Potential for arterial disease detection through easily acquirable peripheral measurements	267
8.3	Tools and frameworks	268
8.3.1	Creation of a physiologically realistic virtual patient database, for the study of arterial haemodynamic	268
8.3.2	Methodology for the creation of physiologically realistic virtual patients	269
8.3.3	Primer for the application of machine learning classifiers	269
9	Future work	271
9.1	Improving classification accuracies	271
9.2	Expanding the current analysis	273
9.3	Validating results and conclusions	275
A	Evaluation of the distribution of pressure and flow-rate profiles across the VPD	277

B	Fourier representation	280
C	ENBC combination search results	284
D	IVBC combination search results	290
E	OVA and OVO multiclass combination search results	296
F	Mesh independence study	302
G	Arterial vessel length	317
H	MCMC trace plots	337
I	Pressure and flow-rate profiles from random VPs	356
J	Hyperparameter contour plots	360
	J.1 CAS	360
	J.2 SAS	360
	J.3 PAD	360
	J.4 AAA	360
K	CAS combination search results	368
L	SAS combination search results	374
M	PAD combination search results	380
N	AAA combination search results	386
O	AAA-L combination search results	392
P	GB results for all disease forms	394
	References	396

Peer-reviewed papers published during PhD candidature

Jones G., Parr J., Nithiarasu P., Pant S. A proof of concept study for machine learning application to stenosis detection. *Medical and Biological Engineering and Computing*. 2021 August <http://dx.doi.org/10.1007/s11517-021-02424-9>.

Jones G., Parr J., Nithiarasu P., Pant S. A physiologically realistic virtual patient database for the study of arterial haemodynamics. *International Journal for Numerical Methods in Biomedical Engineering*. 2021 May 10:e3497.

Jones, G., Parr, J., Nithiarasu, P., Pant S. Machine learning for detection of stenoses and aneurysms: application in a physiologically realistic virtual patient database. *Biomechanics and Modeling in Mechanobiology*. 2021 July <https://doi.org/10.1007/s10237-021-01497-7>.

Conference presentations during the PhD Candidature

G. Jones, J. Parr, P. Nithiarasu, S. Pant. A proof of concept for machine learning application to stenosis detection. *BioMedEng18*, 6-7 September 2018, Imperial College London, London, UK.

G. Jones, J. Parr, P. Nithiarasu, S. Pant. Proof of concept for machine learning applications to arterial disease detection. *Liverpool Early Career Researcher Conference 2019*, 3-5 April 2019, The University of Liverpool, Liverpool, UK.

G. Jones, J. Parr, P. Nithiarasu, S. Pant. Proof of concept for machine learning applications to arterial disease detection. *6th International Conference on Computational and Mathematical Biomedical Engineering*, 10-12 June 2019, Tohoku University, Sendai, Japan.

G. Jones, J. Parr, P. Nithiarasu, S. Pant. A proof of concept for machine learning application to stenosis detection. *BioMedEng19*, 5-6 September 2019, Imperial College London, London, UK.

List of Figures

- 2.1 The configuration of the three elements within the Windkessel model are shown. Q_{1D} and P_{1D} represent the volumetric flow-rate and pressure at the terminal boundary of the 1D system respectively. . . . 16
- 2.2 An example of an MCMC decision tree across two iterations of a chain. $i_{(l)}^{(k)}$ represents the l^{th} possible set (candidates) of the arterial network parameter scaling terms at the k^{th} step of the chain. Green branches denote the path of the chain taken if a candidate is accepted, while red branches show the path of the chain if a candidate is rejected. 20
- 2.3 A logistic regression algorithm returns a prediction of the probability of an observed set of input data resulting in a positive binary classification using a single activation function. $\{y_i^{(1)}, \dots, y_i^{(k)}\}$ represents the 1st to the N^{th} measurement taken from patient i , and $y^{(0)}$ represents a bias term—and so is patient independent. 25
- 2.4 A simple MLP classifier, containing one hidden layer consisting of three neurons. $\{y_i^{(1)}, \dots, y_i^{(N)}\}$ represents the 1st to the N^{th} measurement taken from patient i , $y^{(0)}$ represents the bias term, and $\theta^{(i,j)}$ represents the measurement specific weightings associated with the i^{th} parallel activation function within the j^{th} sequential layer. In the above figure \mathbf{u}_i represents the vector of inputs into the output layer, and within the case shown above is equal to $\mathbf{u}_i = \{h(\mathbf{y}_i, \theta^{(1,1)}), h(\mathbf{y}_i, \theta^{(2,1)}), h(\mathbf{y}_i, \theta^{(3,1)})\}$. Θ represents the matrix of input specific weightings for all activation function within the MLP. 27
- 2.5 Multiple instances of two different classifications are shown. The support vector corresponding to each of the two classifications are highlighted, and the resulting margin is represented by M . The partition between positive and negative events is shown by the solid black line. 30
- 3.1 Coupling of the inlet and outlet boundary conditions to the numerical model are shown. How this model represents the aortic bifurcation is also shown through comparison to an angiogram [26]. 40

3.2	A comparison of the flow-rate profiles produced when expanding the same FS, however truncated at different orders.	41
3.3	An example of a 60% stenosis is shown. This disease is created with a start location $o = 0.2$ and an end location $e = 0.8$	43
3.4	The differences between sensitivity, specificity, recall, and precision. Each of these relationships to the TP, FN, FP, and TN classification accuracies is also included.	56
3.5	The results of the analysis of the adequacy of the size of the VPD, when using pressure and flow-rate measurements at all the three locations, are shown above. The first, second, and third rows show the results of aortic, first iliac, and second iliac classification respectively.	62
3.6	The F_1 scores achieved by the ENBCs employing the NB, LR, SVM, and RF methods when providing each combination of input measurements are shown within the upper subplot. The lower subplot indicated which discrete measurements are present within a combination. Measurements included within each combination are highlighted with a black square.	66
3.7	Two examples of like for like input measurements that should result in identical classification accuracies.	67
3.8	Histograms of the discrepancy between the F_1 score of like-for-like classifiers, when employing the LR, SVM, and RF classification methods.	67
3.9	The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits. The combination of input measurements that produces the highest F_1 score is identified for the SVM method when using each number of input measurements, up to five input measurements.	68
3.10	The histograms of the F_1 scores achieved by each of the three different classification methods are shown for all input measurement combinations that include P_1 in the upper plot, and exclude P_1 in the lower plot.	72

3.11	The histograms of the F_1 scores achieved by each of the three different classification methods are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.	73
3.12	Logistic Regression: predicted probability of disease against stenosis severity for diseased patients (left) and histogram of predicted probability of disease for healthy subjects (right). TP: True Positive; FN: False Negative; FP: False positive; and TN: True Negative. . . .	75
3.13	SVM with an RBF kernel: histograms of True Positives (TP) and False Negatives (FN) against stenosis severity.	75
3.14	The average, maximum, and minimum F_1 score achieved by all IVBCs when providing different numbers of input measurements to detect disease in each of the three vessels are shown. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.	77
3.15	The average, maximum, and minimum F_1 score achieved by all classifiers trained within the multiclass combination search using different numbers of input measurements. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.	80
3.16	The histograms of the F_1 scores achieved for aortic, first iliac, and second iliac classification are shown for all input measurement combinations that include P_1 in the upper plot, and exclude P_1 in the lower plot. This analysis is shown for the results produced by the LR method.	82
3.17	The ROC curve of healthy VP classification within the CPC ensemble trained and tested using pressure and flow-rate measurements at all three available measurement locations are shown. The accuracy of classification is compared to 193 current screening methods (general methods shown in red, and cardiovascular methods shown in blue), recorded in [132] and [7].	86

4.1	Comparison of the pressure (upper row) and flow-rate (lower row) profiles produced at the distal end of the abdominal aorta (left column) and first common iliac (right column), using each of the two numerical methods to solve the healthy VP.	96
4.2	Comparison of the pressure (upper row) and flow-rate (lower row) profiles produced at the distal end of the abdominal aorta (left column) and first common iliac (right column), using each of the two numerical methods to solve the stenosed VP.	97
4.3	Comparison of the pressure (upper row) and flow-rate (lower row) profiles produced at the distal end of the abdominal aorta (left column) and first common iliac (right column), using each of the two numerical methods to solve the aneurysm VP.	98
4.4	The connectivity of the initial network, taken from [20]. At the inlet (free end) of vessel 1, a volumetric flow-rate is specified and at all outlets (free ends of the terminal vessels), a Windkessel model is specified. Locations at which pressure and flow-rate can be measured; and disease is likely to occur are also highlighted, see Section 4.4.2. .	103
4.5	The process of incorporating peripheral 1D vessels into the 0D terminal boundary Windkessel model parameters is outlined above. In the above figure Q_1 and P_1 represent the flow-rate and pressure at the proximal end of the vessel that is being removed; Q_2 and P_2 represent the flow-rate and pressure at the distal end of the vessel that is being removed; Q_3 and P_3 represent the the flow-rate and pressure at the outlet of the system; R_1 , R_2 , and C represent the resistances and compliance of the original terminal Windkessel model; R_v , and C_v represent the viscous resistance and compliance of the vessel being removed; and R_{new} and C_{new} represent the resistance and compliance of the new 2 element Windkessel model after the incorporation of the 1D vessel.	108
4.6	Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the first and second splenic segments; the left gastric; and the common hepatic arteries.	111

4.7	Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the common interosseous, the posterior interosseous, and the second ulnar segment.	113
4.8	Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the the posterior interosseous.	114
4.9	Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the the first and second popliteal segments; the anterior tibial; the tibiofibular trunk; and the posterior tibial.	116
4.10	Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the posterior tibial.	117
4.11	Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the first and second splenic segments, left gastric, common hepatic, and posterior interosseous. Percentage errors have been computed based on error metric presented in [58].	119
4.12	The location of the vessels that have been found to be able to be removed from the initial network without inducing errors of greater than 2%. In the above figure ‘a’, ‘b’, ‘c’, ‘d’, and ‘e’ indicate the second ulnar segments, the common hepatic artery, the first splenic segment, the left gastric artery, and the second splenic segment respectively.	126

4.13	The location of the vessels within each of the five chains. The aortic chain consists of: the first to fourth aortic arch segments denoted by “a” through to “d”, the first to sixth thoracic aorta segments denoted by “e” through to “j”, and the first to fifth abdominal aorta segments denoted by “k” through to “o”. The right and left arm chains consist of: the first and second subclavian segments denoted by “p” and “q” respectively, the axillary artery denoted by “r”, and the brachial artery denoted by “s”. The right and left leg chains consist of: the external iliac denoted by “t”, the first and second femoral segments denoted by “u” and “v” respectively, and the first and second popliteal segments denoted by “w” and “x” respectively. .	130
4.14	Illustration of hierarchical procedure used to assign intra-vessel and intra-chain property profiles to segments within the arterial network. $\Omega_v^{(n,m)}$ and $\Omega_s^{(n,m)}$ represent the decay and scaling parameters of the m^{th} branch at the n^{th} generation respectively; and $\beta^{(n,m)}(x)$ represents the vessel wall mechanical property or radius of the m^{th} branch at the n^{th} bifurcation generation at location x	132
4.15	The three type of prior distributions used for all arterial network scaling parameters.	142
4.16	On the left the measured flow-rate profile taken from the initial network, being incorporated into the posterior distribution, is shown. On the right the corresponding standard deviation profile, computed using Equation (4.25), is shown in terms of the parameter \mathcal{F}	147
4.17	The affect of correlation length as a proportion of the cardiac period on the correlation between time points. The x-axis represents the relative distance between two time points as a proportion of the cardiac period, and the y-axis represents the correlation between these time points. Each line shows the correlation profile when a different correlation length to cardiac period ratio is used.	148
4.18	The percentage difference between the empirical mean of the average cardiac output and the associated standard deviation of 100,000 realisations drawn from GRFs with different standard deviation scaling terms, and the measured average cardiac output and associated standard deviation taken from literature [92][35].	149

4.19	The discrete time points at which the likelihood of the time domain inlet flow-rate profiles is evaluated is shown on the mean flow-rate profile when different numbers of evaluation points are used. The left, middle, and right plots show the locations of evaluation points when using 6, 9, and 12 points respectively.	151
4.20	MCMC trace plot of the scaling term applied to the length of the arterial vessels at every 100 th iteration. In this case, not net migration of the chain is observed.	156
4.21	Autocorrelation as a function of lag-length in the MCMC chain for the vessel length scaling parameter (left), and for the diastolic pressure in the ascending aorta (right). IACT denotes the integrated autocorrelation time [41, 155, 176].	158
4.22	Histogram of correlation coefficient (r) between the 19,701 pairs of parameters in the posterior.	159
4.23	Posterior marginal distribution for the vessel length scaling term, denoted as Φ_a (left); joint distribution between Φ_a and the right arm chain's scaling factor for the β daughter/parent ratio denoted as Φ_b (mid); and joint distribution between Φ_a and the brachiocephalic trunk's scaling factor for β decay denoted as Φ_c (right). While Φ_a and Φ_b show a largely uniform joint distribution, Φ_a and Φ_c show a potentially multi-modal distribution with local peaks at $\Phi_c \approx 1.0$ and $\Phi_c \approx 3.2$	159

4.24	Histograms of the MCMC distributions of the pressure and flow-rate measurements at all measurement locations. The literature based measurements and associated error distribution at each location are overlaid in black. Diastolic and systolic pressure in the right radial artery are shown in (a) and (b), respectively; the diastolic and systolic pressure in the left radial artery are shown in (c) and (d), respectively; the diastolic and systolic pressure in the ascending aorta are shown in (e) and (f), respectively; the diastolic and systolic pressure in the right common carotid artery are shown in (g) and (h), respectively; the diastolic and systolic pressure in the left common carotid artery are shown in (i) and (j), respectively; and the average flow-rate in the second segments of the right and left femoral artery are shown in (k) and (l), respectively.	161
4.25	Histograms of the MCMC distributions of the average flow-rate in the first segment of the right femoral artery (a), and the first segment of the left femoral artery (b). The literature based measurement of average flow-rate in the femoral arteries is overlaid in black.	163
4.26	Comparison of the empirical distribution of the time varying inlet flow-rate profiles between the MCMC samples and the pseudo measurement created through a GRF. All standard deviation curves in dashed lines depict mean ± 1 standard deviation. See Section 4.6.2 for details on the GRF and the literature reported measurements used to create it.	164
4.27	Each blue line shows the profile of correlation between a different starting position and all other evaluation points. The black line shows the correlation profile used within the pseudo measurement GRF, outlined in Section 4.6.2.	166

4.28	In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error (see Table 4.11) are shown by the solid and dashed grey lines, respectively.	167
4.29	The relationship between the ratio of femoral flow-rate pulse to the compliance and resistance ratios.	169
5.1	The limits imposed on the ratios of the post second bifurcation lower extremity compliances in opposite lower extremities are plotted over the ratios of compliances of all post burn-in VPs sampled from the posterior distribution with positive average flow-rate in all vessels. . .	174
5.2	The connectivity of the arterial network. The location of the four forms of disease (see Chapter 4) are highlighted.	176
5.3	An example of a stenosis of severity 0.6 and aneurysm of severity 8.0 are shown. These disease profiles are created with a start location of 0.2 and an end location of 0.8.	179
5.4	Examples of healthy and diseased SA_x , PA_x , and AA_x area profiles. The geometrical boundaries between vessel segments that form the chains are indicated by red dashed lines. See Section 5.3.1 for details of the vessels in each chain.	181

- 5.5 The subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively (see Chapter 4 for details).183
- 5.6 The subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively (see Chapter 4 for details).184
- 5.7 The subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively (see Chapter 4 for details).185

5.8	Histograms of the filtered distributions of the pressure and flow-rate measurements at all measurement locations. The literature based measurements and associated error distribution at each location are overlaid in black. Diastolic and systolic pressure in the right radial artery are shown in (a) and (b), respectively; the diastolic and systolic pressure in the left radial artery are shown in (c) and (d), respectively; the diastolic and systolic pressure in the ascending aorta are shown in (e) and (f), respectively; the diastolic and systolic pressure in the right common carotid artery are shown in (g) and (h), respectively; the diastolic and systolic pressure in the left common carotid artery are shown in (i) and (j), respectively; and the average flow-rate in the second segments of the right and left femoral artery are shown in (k) and (l), respectively.	186
5.9	The histograms of the scalar pressure and flow-rate measurements taken from each of the five VPDs are shown above.	188
6.1	The connectivity of the arterial network. The location of the four forms of disease (see Chapter 5); and six pressure and flow-rate measurements (see Section 6.2.1) are highlighted.	192
6.2	The F_1 scores achieved for CAS using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.	202
6.3	The F_1 scores achieved for SAS using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.	203
6.4	The F_1 scores achieved for PAD using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.	204
6.5	The F_1 scores achieved for AAA using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.	205

6.6	The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for carotid artery stenosis classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.	206
6.7	The average log loss cost across the training and test sets when using the combination of six to three input measurements that achieve the highest accuracies (as detailed in Tables 6.5 and 6.6).	211
6.8	The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for SAS classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.	212
6.9	The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for PAD classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.	214
6.10	The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for AAA classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.	216
6.11	The histograms of the F_1 scores achieved for CAS classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.	218
6.12	The histograms of the F_1 scores achieved for SAS classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.	219
6.13	The histograms of the F_1 scores achieved for PAD classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.	220

6.14	The histograms of the F_1 scores achieved for AAA classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.	221
6.15	The ratios of the F_1 scores for AAA-L classification relative to AAA classification, when providing each combination of input measurements are shown. Measurements included within each combination are highlighted with a black square.	225
6.16	The log loss cost profiles across the training and test sets when using the best performing combination containing three to six input measurements for CAS classification and employing early stopping.	232
6.17	The log loss cost profiles across the training and test sets when using the best performing combination containing three to six input measurements for AAA classification and employing early stopping.	233
7.1	The topological positions of the brachial arteries and the posterior tibials.	257
7.2	An example of brachial artery and ankle pressure profiles taken from a VP for which determination of the footwaves from geometric properties is suitable.	258
7.3	An example of brachial artery and ankle pressure profiles taken from a VP for which determination of the footwaves from geometric properties is problematic. Due to the oscillatory pressure profile in the brachial artery, the minimum pressure and maximum pressure gradient occur during different features in the waveform.	259
7.4	The topological positions of the boundaries of the short ascending and long descending aortic segments, used in the mathematical model of the BCG waveform.	260

7.5	Examples of BCG waveforms taken from five VPs when using constant cross sectional areas for the ascending and descending aortas; allowing spatial cross sectional area variation; and allowing spatial and temporal cross sectional area variation. When using constant cross sectional areas the BCG waveform is computed using Equation (7.5), and A_D and A_A are computed as the average of the mean cross sectional area through out the cardiac period at the inlet and outlet of each segment respectively. When spatial cross sectional area variation is allows the BCG waveform is computed using Equation (7.7) and A_0 , A_1 , A_2 computed as the mean cross sectional area through out the cardiac period at each location.	261
A.1	The mean and standard deviation envelope, <i>i.e.</i> the mean \pm one standard deviation, are shown at all measurement obtainment locations. The subplots show the flow-rate at the inlet of the system (a), pressure at the inlet of the system (b), flow-rate at the outlet of the first iliac (c), pressure at the outlet of the first iliac (d), flow-rate at the outlet of the second iliac (e), and pressure at the outlet of the second iliac (f).	278
B.1	Comparison of the use of the A/B and Z/ϕ forms of a FS to describe the pressure and flow-rate profiles used as inputs into ENBC classifiers.	281
B.2	Comparison of the use of the A/B and Z/ϕ forms of a FS to describe the pressure and flow-rate profiles used as inputs into IVBC classifiers. Comparison is made on the classifiers trained to predict the health of the aorta within the top subplot, the first iliac within the middle subplot, and second iliac within the bottom subplot.	283
F.1	The results of the second step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the spatial discretisation applied to the healthy VP is varied, are shown for the SDC implementation. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	305

F.2	The results of the second step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the spatial discretisation applied to the healthy VP is varied, are shown for the DCG implementation. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	306
F.3	The results of the third step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	307
F.4	The results of the third step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the aneurysms VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	308
F.5	The results of the third step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	309
F.6	The results of the third step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the aneurysm VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	310

F.7	The results of the forth step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the healthy VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	311
F.8	The results of the forth step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	312
F.9	The results of the forth step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the aneurysm VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	313
F.10	The results of the forth step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the healthy VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	314
F.11	The results of the forth step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	315

F.12	The results of the forth step of the mesh independence study, <i>i.e.</i> the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the aneurysm VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.	316
G.1	The mean and standard deviation of the average, maximum, and minimum pressure at all appropriate examination locations for the 30,000 VPs created using a constant length scaling term.	318
G.2	The mean and standard deviation of the average, maximum, and minimum pressure at all appropriate examination locations for the 30,000 VPs created using individual length scaling terms applied to each vessel within the network.	319
G.3	The mean and standard deviation of the average, maximum, and, minimum flow-rate at all appropriate examination locations for the 30,000 VPs created using a constant length scaling term.	320
G.4	The mean and standard deviation of the average, maximum, and minimum flow-rate at all appropriate examination locations for the 30,000 VPs created using individual length scaling terms applied to each vessel within the network.	321
G.5	Due to the high number of individual correlations, it is not possible to show all in a singular plot. Instead the pressure and flow-rate measurements have been broken down into a grid, and a series of subplots created. The location of each subplot within the larger correlation corner plot.	322
G.6	The Pearson correlation coefficients between the average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.	323

G.7	The Pearson correlation coefficients between the average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using an individual scaling factors applied to the length of each vessel within the network.	326
G.8	The Pearson correlation coefficients between the maximum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.	327
G.9	The Pearson correlation coefficients between the maximum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.	328
G.10	The Pearson correlation coefficients between the minimum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.	329
G.11	The Pearson correlation coefficients between the minimum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.	330
G.12	The Pearson correlation coefficients between the maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.	331

G.13	The Pearson correlation coefficients between the maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network	332
G.14	The Pearson correlation coefficients between the minimum and maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.	333
G.15	The Pearson correlation coefficients between the minimum and maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.	334
G.16	The Pearson correlation coefficients between the minimum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.	335
G.17	The Pearson correlation coefficients between the minimum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.	336

H.1	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the: aorta chain initialising value (a), aorta chain decay term (b), right arm chain daughter/parent ratio (c), right arm chain decay term (d), left arm chain daughter/parent ratio (e), left arm chain decay term (f), right leg chain daughter/parent ratio (g), right leg chain decay term (h), left leg chain daughter/parent ratio (i), and the left leg chain decay term (j) for the β properties of the network.	338
H.2	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the: brachiocephalic trunk daughter/parent ratio (a), brachiocephalic trunk decay term (b), right common carotid daughter/parent ratio (c), right common carotid decay term (d), left common carotid daughter/parent ratio (e), and the left common carotid decay term (f) for the β properties of the network.	339
H.3	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the: celiac trunk daughter/parent ratio (g), celiac trunk decay term (h), right common iliac daughter/parent ratio (i), right common iliac decay term (j), left common iliac daughter/parent ratio (k), and the left common iliac decay term (l) for the β properties of the network.	340
H.4	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the: aorta chain initialising value (a), aorta chain decay term (b), right arm chain daughter/parent ratio (c), right arm chain decay term (d), left arm chain daughter/parent ratio (e), left arm chain decay term (f), right leg chain daughter/parent ratio (g), right leg chain decay term (h), left leg chain daughter/parent ratio (i), and the left leg chain decay term (j) for the r_0 properties of the network.	341
H.5	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the: brachiocephalic trunk daughter/parent ratio (a), brachiocephalic trunk decay term (b), right common carotid daughter/parent ratio (c), right common carotid decay term (d), left common carotid daughter/parent ratio (e), the left common carotid decay term (f) for the r_0 properties of the network.	342

H.6	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the: celiac trunk daughter/parent ratio (g), celiac trunk decay term (h), right common iliac daughter/parent ratio (i), right common iliac decay term (j), left common iliac daughter/parent ratio (k), and the left common iliac decay term (l) for the r_0 properties of the network.	343
H.7	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the daughter/parent ratio of the: inferior mesenteric (a), right renal (b), left renal (c), superior mesenteric (d), second left posterior intercostal (e), second right posterior intercostal (f), first left posterior intercostal (g), first right posterior intercostal (h) for the β properties of the network.	344
H.8	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the daughter/parent ratio of the: first right ulnar (a), right common interosseous (b), right radial (c), right vertebral (d), right external carotid (e), right internal carotid (f) for the β properties of the network.	345
H.9	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the daughter/parent ratio of the: first left ulnar (a), left common interosseous (b), left radial (c), left vertebral (d), left external carotid (e), left internal carotid (f) for the β properties of the network.	346
H.10	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the daughter/parent ratio of the: right anterior tibial (a), left anterior tibial (b), right posterior tibial (c), left posterior tibial (d), right profunda femoris (e), left profunda femoris (f), right internal carotid (g), left internal carotid (h) for the β properties of the network.	347
H.11	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the daughter/parent ratio of the: inferior mesenteric (a), right renal (b), left renal (c), superior mesenteric (d), second left posterior intercostal (e), second right posterior intercostal (f), first left posterior intercostal (g), first right posterior intercostal (h) for the r_0 properties of the network.	348

H.12	The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: first right ulnar (a), right common interosseous (b), right radial (c), right vertebral (d), right external carotid (e), right internal carotid (f) for the r_0 properties of the network.	349
H.13	The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: first left ulnar (a), left common interosseous (b), left radial (c), left vertebral (d), left external carotid (e), left internal carotid (f) for the r_0 properties of the network.	350
H.14	The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: right anterior tibial (a), left anterior tibial (b), right posterior tibial (c), left posterior tibial (d), right profunda femoris (e), left profunda femoris (f), right internal carotid (g), left internal carotid (h) for the r_0 properties of the network.	351
H.15	The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: inferior mesenteric (a), right renal (b), left renal (c), superior mesenteric (d), second left posterior intercostal (e), second right posterior intercostal (f), first left posterior intercostal (g), and the first right posterior intercostal (h).	352
H.16	The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: second right ulnar (a), right common interosseous (b), right radial (c), right vertebral (d), right external carotid (e), and the right internal carotid (f).	353
H.17	The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: second left ulnar (a), left common interosseous (b), left radial (c), left vertebral (d), left external carotid (e), and the left internal carotid (f).	354

H.18	The parameter scaling terms at each 100 th iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: right anterior tibial (a), left anterior tibial (b), right posterior tibial (c), left posterior tibial (d), right profunda femoris (e), left profunda femoris (f), right internal iliac (g), left internal iliac (h).	355
I.1	In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively.	357
I.2	In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively.	358
I.3	In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right carotid artery (d), pressure profiles in the left carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively.	359

J.1	The contour plot of the F_1 scores achieved for CAS classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.	361
J.2	The contour plot of the F_1 scores achieved for CAS classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.	361
J.3	The contour plot of the F_1 scores achieved for CAS classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.	362
J.4	The contour plot of the F_1 scores achieved for SAS classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.	362
J.5	The contour plot of the F_1 scores achieved for SAS classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.	363
J.6	The contour plot of the F_1 scores achieved for SAS classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.	363
J.7	The contour plot of the F_1 scores achieved for PAD classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.	364
J.8	The contour plot of the F_1 scores achieved for PAD classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.	364
J.9	The contour plot of the F_1 scores achieved for PAD classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.	365
J.10	The contour plot of the F_1 scores achieved for AAA classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.	366
J.11	The contour plot of the F_1 scores achieved for AAA classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.	366

J.12	The contour plot of the F_1 scores achieved for AAA classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.	367
P.1	The F_1 scores achieved for all disease forms when employing the GB method. Measurements included within each combination are highlighted with a black square.	395

Common Symbols and Abbreviations

Name	Symbol/Abbreviation	Units
Virtual patient database	VPD	-
Virtual patient	VP	-
Carotid artery stenosis	CAS	-
Subclavian artery stenosis	SAS	-
Peripheral artery disease	PAD	-
Abdominal aortic aneurysm	AAA	-
Machine learning	ML	-
Naive Bayes'	NB	-
Logistic regression	LR	-
Multi-layer perceptron	MLP	-
Support vector machine	SVM	-
Random forest	RF	-
Gradient boosting	GB	-
Entire network binary classifier	ENBC	-
Individual vessel binary classifier	IVBC	-
Custom probabilistic configuration	CPC	-
One-versus-one	OVO	-
One-versus-all	OVA	-
Markov chain Monte-Carlo	MCMC	-
Mechanical properties of arterial vessel	β	dyn/cm
Reference radius	r_0	cm
Cross sectional area	A	cm ²
Pressure	P	mmHg

Name	Symbol/Abbreviation	Units
Flow-rate	Q	ml/s
Temporal coordinate	t	s
Spatial coordinate	x	cm
Time period of cardiac cycle	T	s
Fourier series sine coefficient	a	-
Fourier series cosine coefficient	b	-
Probability distribution of variable I	$p(I)$	-
Probability distribution of variable I , given J	$p(I J)$	-

Table 1: Glossary of frequently used terms, including units where applicable.

Chapter 1

Introduction

Arterial disease refers to any disease affecting the arterial system. Two of the most common forms of arterial disease are stenosis and aneurysm. These two forms of disease are estimated to affect between 1% and 20% of the population [66, 186, 129, 122]. Ruptured abdominal aortic aneurysms alone are estimated to be the cause of between 6,000 and 8,000 deaths a year within the United Kingdom [44]. Patients with stenosis have been shown to have a mortality hazard ratio of 1.42 compared to a control population [2], and an unadjusted death rate of 3.35 per 100 person-years compared to 1.23 per 100 person-years in a control population [97]. Current methods for the detection of arterial disease are primarily based on obtaining images of the area profiles of arterial vessels. The techniques for obtaining such images are generally impractical for large scale screening, expensive, or both. If an inexpensive method for the detection of both stenosis and aneurysm is created, that minimises the need for invasive measurements, the cost effectiveness of large scale screening could be improved making both continuous monitoring and screening feasible.

One such alternative is to use easily acquirable pressure and flow-rate measurements at accessible peripheral locations within the circulatory system for diagnosis. Within this thesis the potential of using machine learning (ML) classification algorithms to detect arterial disease from such measurements is investigated. A series of virtual patient databases (VPDs), similar to that presented in [217], are created using a physics based model of one dimensional pulse wave propagation [20]. These synthetic data sets are then used to train and test a series of ML algorithms, designed to distinguish between healthy and unhealthy patients.

This chapter provides an overview of the following: arterial disease, and why detection of arterial disease is important; the relationship between pressure and flow-rate profiles and the health of a patient's arterial system, and how this may be exploited to detect arterial disease; why the application of ML classification algorithms to arterial disease detection is logical; and how ML classification algorithms may be applied to the problem of arterial disease detection. This chapter culminates in an outline of the aims and objectives of the remaining thesis.

1.1 Arterial disease

The cardiovascular system comprises of the heart, blood vessels, and blood. The blood vessels within the cardiovascular system can be further broken down into arteries, the micro-circulation, and veins. The arterial system transports high pressure blood from the heart to the micro-circulation where it is permeated through organs and tissue. Cardiovascular disease is a general term given to conditions relating to the heart and blood vessels. When this disease affects the arterial system it is referred to as arterial disease. While there are many forms of arterial disease, two of the most common are stenosis and aneurysm.

A stenosis is a narrowing of an arterial vessel. This is normally caused by a build up of fatty deposits, known as atherosclerosis. The rate of atherosclerosis has been recorded to vary in different populations between 6.2% and 18.3% [134]. An exact value of the prevalence of resulting stenosis is hard to obtain due to the fact that stenosis can be be categorised into several sub-diseases, prevalence varies dependent on the demographics of the population studied, and the interpretation of the point atherosclerosis becomes stenosis is not fixed. One of the most common forms of stenosis is peripheral artery disease (PAD). PAD refers to the stenosis of any peripheral artery, however most commonly the legs [115, 1, 39]. The prevalence of PAD has been recorded to vary between 5.28% and 18.83% within different demographics [66]. Within the upper body, a common form of stenosis is subclavian artery stenosis (SAS), with the prevalence recorded to be 1.9% in a free-living population and 7.1% in a clinical population [186]. Another common form of stenosis is carotid artery stenosis (CAS), with 3.8% of men and 2.7% of women recorded to be affected [129]. While the occurrence of these different forms of stenosis are not mutually exclusive, 18.7% of patients with PAD have been recorded to also have SAS [78], these numbers do give an indication to the scale of the effect. Patients diagnosed with SAS have been shown to have a mortality hazard ratio of 1.42 compared to a control population [2], while patients with CAS have been shown to have a unadjusted death rate of 3.35 per 100 person-years compared to 1.23 per 100 person-years in a control population [97].

The second common form of arterial disease are aneurysms. An aneurysm is a localised weakening of an arterial vessel wall, causing the vessel to bulge. This bulge

will gradually grow over time until, if left untreated, it may eventually rupture [59, 179]. As with stenosis, aneurysm may be referred to by several different names, dependent on the location of the disease. The most common form of arterial aneurysm is abdominal aortic aneurysm (AAA), with a prevalence of 4.8% [122]. Ruptured AAAs are estimated to be the cause of between 6,000 and 8,000 deaths a year within the United Kingdom [44], and approximately 45,000 deaths a year in the United States of America [81]. Currently AAA is the only form of arterial disease for which the United Kingdom's National Health Service (NHS) offers screening. While this program of voluntary screening of 65 year old males is expected to reduce the rate of premature death from ruptured AAAs by 50% by 2023, the current expected cost of this is £7,600 per quality life-year gained [48].

Current methods for the detection of arterial disease are primarily based on obtaining images of the area profile of arterial vessels [205, 124, 163, 119]. Two frequently used techniques to obtain these images are:

- **Angiography.** Arterial vessels are not naturally visible within X-ray images. Angiography involves the insertion of a catheter into the arterial system. This catheter is guided to the portion of the arterial network of which an image is desired, and used to inject a contrast medium. This contrast medium highlights arterial vessels in X-ray images. X-ray images can then be taken of the arterial system, and examined to diagnose arterial disease. Angiography is invasive, and requires local anaesthesia.
- **Doppler ultrasound.** Doppler ultrasound involves a trained technician running a probe along the surface of the skin. This probe emits high frequency sound waves, and records the reflected waves. The shift in frequency between the emitted and received waves allows for the construction of an image of the arterial vessels beneath the probe. Assessing arterial health through the use of Doppler ultrasound requires the probe to be run over all arterial vessels for which an image is desired. A further limiting factor is that arterial vessels deep in the body are harder to see than superficial vessels.

It can be seen from the examples highlighted above that imaging techniques are generally impractical for large scale screening. While the NHS currently offers screening for AAAs through the use of Doppler ultrasound, this program is limited

in scope due to cost effectiveness. Currently all 65 year old males in the United Kingdom are invited to undergo scanning for the presence of an AAA. If the abdominal aorta of a patient is found to have a diameter less than 3cm, the patient is discharged with no future scanning carried out. If the abdominal aorta of a patient has a diameter between 3cm–4.4cm or 4.5cm–5.4cm the patient is invited back for a follow up scan in 1 year or 3 months time, respectively. If the diameter of a patient’s abdominal aorta is found to be greater than 5.4cm the patient is invited to undergo surgery [48]. Concerns have been raised about the current screening protocol due to the lack of follow up scans for patients with a diameter of less than 3cm [204], and the reduction in cost effectiveness as the prevalence of AAAs reduces [73]. Currently screening for AAAs is not offered to women, partially due to the fact that a similar protocol to that offered to men would cost £30,000 per quality life-year gained due to the lower prevalence [200]. The NHS currently offers no screening program for the presence of any form of stenosis.

If a new method for the detection of stenosis and aneurysm is created that minimises both the cost and the need for invasive measurements, the cost effectiveness and feasibility of large scale screening would be improved. This could, in turn, allow for an expansion to the current AAA screening program, and the implementation of a stenosis screening program.

1.2 Use of haemodynamic measurements for disease detection

A potential alternative is to use easily acquirable pressure and flow-rate measurements at peripheral locations. It is known from the principals of fluid mechanics that if the cross sectional area of a vessel is changed, the pressure and flow-rate profiles of fluid passing through that vessel will also change [133, 50, 208, 183]. Applying this to arterial disease, the inclusion of a stenosis or aneurysm within a patient’s arterial network should create biomarkers within the pressure and flow-rate profiles of blood. The implications of the presence of arterial disease on pressure and flow-rate profiles has already been investigated through the use of one-dimensional models (similar to those utilised later in this thesis) by a number of previous studies. It is found in [127] that even low severity AAAs have a global

impact on the pressure and flow-rate profiles. Further more [183] has shown that the presence of an aneurysm produces measurable biomarkers within flow-rate profiles, with these results validated by experimental data. Other similar studies that have shown the haemodynamic effects of arterial disease include [201], [107], and [90].

Further supporting the use of easily acquirable peripheral measurements is the significant work already carried out in creating inexpensive methods to measure both blood pressure and flow-rate. Continuous time varying flow-rate profiles can be obtained in the brachial [27], carotid [144], and femoral [164] arteries using Doppler ultrasound. It is important to note that the obtainment of flow-rate profiles through the use of Doppler ultrasound differs from obtaining an image of the vessel as measurements are only required at a single discrete location, rather than along the entirety of a vessels length. Non-invasive pressure profiles can be obtained in the radial and common carotid arteries using applanation tonometry [3, 146] and in the brachial arteries through reconstruction of finger arterial pressure [77]. A diagnostic method that utilised an abundance of already clinically obtainable pressure and flow-rate measurements would have the potential to significantly improve the cost effectiveness of large scale screening.

The relationship between blood pressure and cardiovascular health is already extensively documented and studied, supporting the use of haemodynamic measurements for arterial disease detection. Previous studies have highlighted correlations between blood pressure and the risk of cardiovascular disease [185]. It has been shown that male and female patients with high-normal blood pressure are 1.6 and 2.5 times more likely to experience an incidence of cardiovascular disease over a 10 year period than those with optimum blood pressure [210], respectively. The importance of blood pressure in assessing cardiovascular health is further supported by the known correlation between the ratio of ankle to brachial systolic blood pressure, referred to as ankle-brachial pressure index, and mortality rate [135]. It can be seen from the above examples, however, that generally the use of haemodynamic measurements for assessing cardiovascular health is currently advisory. The above examples do not offer fixed “rules” for the diagnosis of arterial disease, and instead only allows for the evaluation of the relative risk of a patient being affected. To allow for the creation of a diagnostic method specific evaluation criteria within pressure and flow-rate measurements are required.

1.3. MACHINE LEARNING CLASSIFICATION ALGORITHMS AND THEIR APPLICATION TO ARTERIAL DISEASE DETECTION

“Since the information which the pulse affords is of so great importance, and so often consulted, surely it must be to our advantage to appreciate fully all it tells us, and to draw from it every detail that is capable of imparting (Frederick Akbar Oratio Mahomed (1872)).”

One example of the current use of haemodynamic measurements as a diagnostic method is the use of ankle-brachial index to diagnose PAD. A previous meta-analysis study of this method has shown high levels of specificity (83.3 – 99.0%) however varying levels of sensitivity (15 – 79%) [218]. The arterial system, and thus the underlying pressure and flow-rate profiles, are patient specific. This poses a serious issue in using haemodynamic measurements to detect arterial disease. A single set of healthy baseline pressure and flow-rate profiles, which can be used for comparison, does not exist. It is, therefore, difficult to determine if characteristics of a patient’s pressure and flow-rate profiles are caused by natural variance or arterial disease. If a consistent and significant biomarker of arterial disease is found within pressure and flow-rate profiles, irrespective of a patients unique underlying arterial network, stenosis and aneurysm can be diagnosed from pressure and flow-rate measurements.

1.3 Machine learning classification algorithms and their application to arterial disease detection

To predict the presence of a stenosis or aneurysm through the use of haemodynamic profiles, a model is required that can map pressure and flow-rate measurements to a patients underlying arterial health. One method for the creation of such a mapping is to take a physics advised approach [183]. It is possible to use physics based models of blood flow, that compute the pressure and flow-rate profiles associated with particular realisations of arterial networks, to aid in the diagnosis of arterial disease. Direct inversion of these physics based models of blood flow is complex and impractical. Instead patient specific parameter estimation [29] can be carried out, and expected pressure and flow-rate measurements obtained. Comparison of these expected measurements to those recorded in patients allow for the assessment of arterial health. A possible method to detect CAS using an extended physics

1.3. MACHINE LEARNING CLASSIFICATION ALGORITHMS AND THEIR APPLICATION TO ARTERIAL DISEASE DETECTION

based model, containing a patient specific model of blood flow, is proposed in [34]. A problem with the inclusion of these computational models is that every time a patient is tested some form of parameter estimation is required, incurring a computational expense that is likely to be proportional to the accuracy of this estimation. Thus, accuracy of disease detection is likely to be proportional to the accuracy of the patient specific parameter estimation, and so consequently proportional to the computational expense.

An alternative method to create a model mapping pressure and flow-rate measurements to a patient's arterial health is to take a purely data driven approach. A model directly predicting the health of a patient from pressure and flow-rate measurements bypasses the need for physics based models. It is likely that the indicative biomarkers of arterial disease held within pressure and flow-rate profiles consist of micro inter- and intra- measurement details. Discovery of these biomarkers through a traditional hypothesis-driven scientific method [213] is likely infeasible—*i.e.* proposing a possible biomarker, testing to see if arterial disease can be detected using this proposed indicator, and repeating until suitable features are discovered. If a large database of pressure and flow-rate measurements taken from patients of known arterial health is available, it is possible for a ML classifier to be trained to not only discovery but also exploit any biomarkers held within pressure and flow-rate profiles. This methodology involves a significantly greater initial expense associated with the creation of the model, relative to a physics advised approach. Once the model predicting arterial health has been created, however, the diagnosis of patients is computationally inexpensive and near instantaneous. In the case of a data driven approach the limit to the achievable accuracy is likely to be imposed by the data set used to create the model. Creating the model mapping pressure and flow-rate measurements to arterial health from an unrealistic or incomplete data set will result in an imprecise and inaccurate diagnostic method.

A supervised ML classifier, essentially, uses pre-observed data to create an estimate of the model mapping a series of input variables to a dependent outcome [113]. This estimated model is then used to assign predicted outcomes to new unseen input variables. Due to the versatility, potential for high accuracy, and ever increasing ease of implementation [60] a large range of different classification algorithms—that exploit the information held within the pre-observed data through different mechanisms—exist [194, 103, 123, 172, 140]. In the case of this thesis,

1.3. MACHINE LEARNING CLASSIFICATION ALGORITHMS AND THEIR APPLICATION TO ARTERIAL DISEASE DETECTION

a model is desired that maps pressure and flow-rate measurements directly to the arterial health of the patient from which the measurements have been taken.

ML has a history of being used for medical applications [112]. Classification algorithms have been shown to be able to predict the presence of irregularities in heart valves [42], arrhythmia [193], and sleep apnea [106] from recorded time domain data. These studies suggest it is possible for a ML classifier to be trained to detect arterial disease using only time domain pressure and flow-rate measurements.

To train ML classifiers, and then consequently test their performance, a large database of measurements taken from patients of known arterial health is required. One method for the obtainment of such a database is to record measurements from a real cohort. While this method does ensure that the range and distribution of pressure and flow-rate measurements is physiologically accurate—given that the patient cohort is comprehensive—it also has several disadvantages. The difficulty in diagnosing arterial disease is the major motivating factor for this thesis, and thus obtainment of the health classification of patients is prohibitive. This issue is further amplified when the large number of patients required to train and test a ML classifier is factored in. A further limiting factor is the imbalance in the number of healthy and unhealthy patients within a real population. The prevalence of arterial disease is recorded to vary between approximately 1% and 20%. While medically this prevalence is high, for a ML application it is not. The imbalance in the number of healthy and unhealthy patients is likely to cause issues in the training and testing of ML classifiers.

An alternative method to obtain a large database of pressure and flow-rate profiles, and the corresponding health of patients, is to create a synthetic data set. Physics based models of one dimensional pulse wave propagation, that compute the pressure and flow-rate profiles produced by particular realisation of arterial networks, have been extensively developed and tested [20, 64, 6, 145, 168, 131]. An available pre-existing solver has been successfully validated against a 3D model of bloodflow through stenosed arterial vessels [21]. Virtual patients (VPs) can be created by sampling random realisations of arterial networks, and solving the physics based model of pulse wave propagation to obtain the corresponding pressure and flow-rate profiles. The generation of a virtual patient database (VPD) in this manner is supported by the employment of a very similar methodology in [217]. Using VPs to train and test ML classifiers provides several advantages, such as:

- **Expense:** creating VPs is relatively inexpensive. The primary cost associated with the creation of VPs is computational, and thus negligible in comparison to data acquisition in a real population.
- **Class imbalance:** creating VPs allows for the control of the distribution of different disease states. For example, in a real population the rate of arterial disease can vary between 1% and 20%. During the creation of VPs, however, 50% are created with disease to ensure a balanced data set.
- **Measurement availability:** using VPs allows for measurements of pressure and flow-rate to be taken at any location within the arterial system.

It is important to note that in a clinical application there will be limitations to the measurements that can be non-invasively and inexpensively obtained. The availability of pressure and flow-rate measurements throughout the arterial network are useful, however, as they allow for an *a priori* assessment of ML classifiers to be tested using all possible combinations of pressure and flow-rate measurements.

1.4 Aims and objectives

The primary aim of this thesis is to carry out an investigation into the potential of using ML classification algorithms to predict the presence of stenosis or aneurysm, directly from easily acquirable haemodynamic measurements. An exploratory stance is taken to these classifiers with focus on uncovering behaviours and patterns in the performance of classification, rather than optimisation and creation of increasingly complex ML models for maximum accuracy.

The secondary aims of this thesis are as follows:

- Development of a frame work, through the employment of a physics based model, for the creation of synthetic data sets. This frame work is to developed to allow for: i) the creation of several different VPDs as the required balance between simplicity and physiological realism progresses, and ii) the generation of a generic methodology, given a mathematical description of a biological system, that can be adopted to create virtual patients for any biological system while accounting for all available information.

- Exploitation of data driven methods, such as ML classifiers, to gain insight into the haemodynamic effects of arterial disease.

1.5 Outline of Thesis

The remainder of this thesis is organised as following:

Chapter 2: Important methodologies. The second chapter of this thesis provides an overview of the most important methodologies used. This chapter does not aim to provide a comprehensive, in depth explanation and examination of these methodologies. Instead a relatively simple understanding of the key elements and processes involved with each method are described. This chapter aims to build a sufficient understand of the key methodologies, so that the processes carried out within the subsequent chapters of the thesis are more intuitive.

Chapter 3: Three vessel network proof-of-concept study. The third chapter of this thesis carries out a proof-of-concept (PoC) study. The aim of this PoC is to assess the ability of ML classifiers to predict the presence of a stenosis within a simple three vessel arterial system. A series of ML classifiers—as previously outlined in Chapter 2—are presented, trained, and tested using different combinations of pressure and flow-rate measurements.

The majority of work in this chapter is published in: *Jones G., Parr J., Nithiarasu P., Pant S. A proof of concept study for machine learning application to stenosis detection. Medical and Biological Engineering and Computing. 2021 August <http://dx.doi.org/10.1007/s11517-021-02424-9>.*

Chapter 4: Creation of a physiologically realistic virtual patient database. To expand on the PoC study, Chapter 4 outlines the creation of a new physiologically realistic healthy VPD. Within this study a reduced version of the anatomically correct arterial network proposed by Blanco et al. [18, 17], referred to as the ADAN network, is used as the basis for the creation of new VPs. To ensure physiological realism across the new VPD, measurements of pressure and flow-rate are taken from literature. Bayes' theorem is used to incorporate the effect of these measurements into the distributions from which VPs arterial network parameters are sampled. Using an appropriate sampling method, in this case a Markov chain Monte-Carlo

(MCMC) method [69, 71, 25], realisations of VPs arterial networks are sampled from the distribution of arterial network parameters subject to the effect of literature based measurements. The VPD created in this chapter is made publicly available (<https://doi.org/10.5281/zenodo.4549764>).

The majority of work in this chapter is published in: *Jones G., Parr J., Nithiarasu P., Pant S. A physiologically realistic virtual patient database for the study of arterial haemodynamics. International Journal for Numerical Methods in Biomedical Engineering. 2021 May 10:e3497.*

Chapter 5: Pre-processing and creation of literature advised unhealthy VPDs. Using the physiologically realistic healthy VPD created within Chapter 4, Chapter 5 outlines the pre-processing carried out to create a usable data set. Physics based filters are applied to limit the occurrence of VPs displaying undesirable behaviours. The filtered healthy VPD is used to create both aneurysm and stenosis VPDs, paying consideration to the range of severities frequently observed in a real world clinical environment.

Chapter 6: Application of machine learning classifiers to the physiologically realistic VPDs. The literature advised unhealthy VPDs—created within Chapter 5—are used along side the filtered physiological realistic healthy VPD—created within Chapter 4—to train and test a series of ML classifiers. Where possible and when appropriate, clinical and physiological restrictions and considerations are applied to the training and testing of these classifiers.

The majority of work in this chapter is published in: *Jones, G., Parr, J., Nithiarasu, P., Pant S. Machine learning for detection of stenoses and aneurysms: application in a physiologically realistic virtual patient database. Biomechanics and Modeling in Mechanobiology. 2021 July <https://doi.org/10.1007/s10237-021-01497-7>.*

Chapter 7: Use of physics advised haemodynamics features. Hitherto, pressure and flow-rate profiles have been used in their raw form, *i.e.* a representation of the measurable pressure and flow-rate profiles, to train and test ML classifiers. It is possible, however, to use these raw pressure and flow-rate profiles to construct complex physics advised features. Examples of physics advised features, based on pressure and flow-rate profiles, include; ankle-brachial pressure index, pulse wave velocity [12], and ballistocardiogram waveforms [93]. While some ML classifiers are capable of combining input measurements to create high order features, the direct

provision of these physics advised features may aid in classification. Classifiers are trained and tested using physics advised features both in combination with raw pressure and flow-rate measurements, and in isolation. These results are compared with those achieved using solely raw pressure and flow-rate measurements (as outlined in Chapter 6).

Chapter 8: Conclusions and contributions The penultimate chapter of this thesis aims to put the results achieved within the previous chapters into the context of the overriding objectives of this research. The potential for using ML classifiers to detect arterial disease is evaluated both in an abstract sense and with consideration to clinical restrictions and requirements. Possible limitations in the methodologies taken within the thesis are identified and discussed, to build a rounded perspective of the current work. Finally this chapter aims to identify and evaluate the contributions of this thesis to both the computational engineering community, and specifically to the problem of arterial disease detection.

Chapter 9: Future work The final chapter of this thesis uses the evaluation of the progress achieved and limitations of this body of work—as outlined in Chapter 8—to assess and discuss future work that can either strengthen the conclusions reached within this thesis or progress the development of potential diagnostic methods.

Chapter 2

Important methodologies

In this chapter, an overview of the most important methodologies is presented. This presentation is relatively generic, with the implementation of each method being discussed in the subsequent chapters. This chapter seeks to build a relatively simple, intuitive, understanding of the methodologies. Sufficient references are provided to allow the reader to gain a deeper understanding if desired. The methods outlined in this chapter, and employed throughout the remainder of this thesis, are chosen due to their simplicity and “robustness”. One of the primary purposes of this thesis is to bridge physics based modelling, statistical modelling, and machine learning into a framework that can be used to create and then subsequently exploit VPs for a range of biological systems. While these methods may not be the “best” available—when evaluated on their efficiency or accuracy—they are easily applied to a range of different systems without the need for extensive problem specific optimisation or in-depth specialist knowledge. Thus, the resulting framework is accessible and implementable by future researchers with a wide range of different backgrounds and experiences.

The remainder of this chapter is structured as follows: first, the physics based model—used to compute the pressure and flow-rate profiles associated with VPs—is presented in Section 2.1; next, two statistical methods (Bayes’ theorem and Metropolis-Hastings Markov chain Monte-carlo) are outlined in Sections 2.2 and 2.3; and finally the ML classification methods used are described in Section 2.4.

2.1 Physics based model of pulse wave propagation

To compute the pressure and flow-rate waveforms associated with VPs, a physics based model of one-dimensional pulse wave propagation is adopted [20]. By considering each vessel within the network to be a deforming tube, a system of two governing equations can be derived. These equations represent conservation of mass and momentum balance with the assumption that blood is incompressible and

2.1. PHYSICS BASED MODEL OF PULSE WAVE PROPAGATION

that the tube walls are impermeable. The system of equations is (see [6] for details):

$$\frac{\partial A}{\partial t} + \frac{\partial(UA)}{\partial x} = 0, \quad (2.1)$$

$$\frac{\partial U}{\partial t} + U \frac{\partial U}{\partial x} + \frac{1}{\rho} \frac{\partial P}{\partial x} = \frac{f}{\rho A}, \quad (2.2)$$

where $P(x, t)$, $U(x, t)$, and $A(x, t)$ represent the pressure, flow velocity, and arterial cross sectional area, respectively, at spatial coordinate x and time t ; ρ represents the density of blood; and f represents the frictional force per unit length described as follows:

$$f(x, t) = -2(\zeta + 2)\mu\pi U, \quad (2.3)$$

where ζ is a constant that depends on the velocity profile across the arterial cross section, and μ represents the dynamic viscosity of blood. To close this system of equations, a mechanical model of the displacement of the vessel walls [20] is included:

$$P - P_{ext} = P_d + \beta \frac{\sqrt{A} - \sqrt{A_d}}{A_d}, \quad (2.4)$$

with

$$\beta = \frac{4}{3} E h \sqrt{\pi}, \quad (2.5)$$

where P_{ext} represents the external pressure, P_d represents the diastolic blood pressure, A_d represents the diastolic area of the vessel, β represents the mechanical property of the vessel, E represents the vessel wall's Young's modulus, and h represents the vessel wall's thickness. This system of equations has been previously used and tested extensively [20, 64, 6, 145, 168, 131].

A possible limitation associated with the use of one-dimensional models is the inability to capture complex three-dimensional behaviours, particularly when high severity diseases (*i.e.* large changes in vessel area) are introduced. Examples of such behaviour which may not be captured include artificial lumens; and turbulence and transitional flow (changing pressure drops). It is chosen to use one-dimensional models (rather than more complex three-dimensional models) for two reasons:

- Similar one-dimensional models have been used to compute the pressure and flow-rate profiles associated with patients with the presence of an aneurysm in [33], [96] and [183]; and stenosis in [21] and [96]—suggesting the suitability

2.1. PHYSICS BASED MODEL OF PULSE WAVE PROPAGATION

of this model for use on both healthy and unhealthy patients.

- The significantly lower computational expense associated with one-dimensional models makes them better suited to this initial exploratory study.

For the purposes of this thesis, it is computationally unpragmatic to model the entire arterial network. Instead a sub-network of interest is identified, and the arterial network truncated by incorporating all prior and subsequent vessels as appropriated boundary conditions. The inlet and outlet boundary conditions are explained next.

2.1.1 Inlet boundary conditions

The inlet of sub-networks are generally coupled to prescribed time domain volumetric flow-rates. The effect of vessels prior to a sub-network are incorporated into this inlet flow-rate. For this thesis all prescribed time domain inlet flow-rate profiles are described using a Fourier series (FS). A FS describes a periodic function through the synthesis of a series of harmonic components:

$$Q_{\text{inlet}}(t) = \sum_{n=0}^N a_n \sin(n\omega t) + b_n \cos(n\omega t), \quad (2.6)$$

where a_n and b_n represent the n^{th} sine and cosine FS coefficients, respectively; N represents the truncation order; and $\omega = 2\pi/T$, with T as the time period of the cardiac cycle.

2.1.2 Outlet boundary conditions

The terminal boundaries of sub-networks are generally coupled to three element Windkessel models [215]. This model replicates the effect of peripheral arteries using an electrical system. This system comprises of two resistors, R_1 and R_2 , which replicate the viscous resistance of large arteries and the micro-vascular system respectively, and a capacitor C which replicate the compliance of large arteries. It is common for the first resistor (R_1) to be equal to the characteristic impedance of the connecting one-dimensional vessel, to reduce reflections.

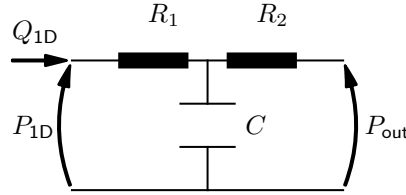


Figure 2.1: The configuration of the three elements within the Windkessel model are shown. Q_{1D} and P_{1D} represent the volumetric flow-rate and pressure at the terminal boundary of the 1D system respectively.

The configuration of these three elements is shown in Figure 2.1.

2.2 Bayes' theorem

Bayes' theorem is a method used to update the prior knowledge and beliefs held about the distribution of a parameter, given that new observations and data is obtained [117, 55]. In this thesis Bayes' theorem is primarily used to construct a distribution for the parameters describing a virtual patients arterial network that incorporates multiple pieces of dissimilar (geometric and haemodynamic) pieces of information. The prior knowledge held about a distribution, before the obtainment of new data, is described as $p(I)$. Given that the variable J is observed to be equal to j , *i.e.* $J = j$, the updated conditional distribution of variable I incorporating this new observation can be computed using the equation:

$$\underbrace{p(I | J = j)}_{\text{posterior}} = \frac{\overbrace{p(J = j | I)}^{\text{likelihood}} \overbrace{p(I)}^{\text{prior}}}{\underbrace{p(J = j)}_{\text{evidence}}}, \quad (2.7)$$

where $p(I | J = j)$ represents the distribution of the variable I incorporating the effect of observing $J = j$ —referred to as the posterior distribution. Within Equation (2.7) $p(J = j | I)$ represents the likelihood of the parameters, given the data—referred to as the likelihood. The final term in Equation (2.7), $p(J = j)$, represents

the probability of observing the measurement $J = j$, averaged (marginalised) over all values of I . This is referred to as the evidence. This marginal probability can be found by integrating the joint distribution of variables I and J , described by $p(J = j, I)$, across the support of variable I , *i.e.*:

$$p(J = j) = \int_I p(J = j, I) dI = \int_I p(J = j | I) p(I) dI. \quad (2.8)$$

Due to the required integration it is often unpragmatic to compute the evidence term. This term, however, is independent of the variable I and, thus, does not affect the shape of the posterior distribution. Instead this term purely acts as a normalising constant, ensuring the integral of the posterior is equal to 1. It is, therefore, possible to evaluate the posterior distribution, up to the limit of the normalising constant, without the need for computation of the evidence term, *i.e.*:

$$p(I | J = j) \propto p(J = j | I) p(I), \quad (2.9)$$

where $p(J = j | I) p(I)$ can be thought of as the unnormalised density of the posterior distribution.

2.3 Metropolis-Hastings Markov chain Monte-Carlo

Given that the probability density function (PDF), and ideally the cumulative distribution function (CDF), is known the ideal method to draw a series of samples from a distribution is to independently select points from the support based on the CDF. If, however, the CDF is not directly known and computation from the PDF is complex, or the CDF and PDF are both directly unknown this method can become computationally expensive. An alternative method is a Markov chain Monte-Carlo (MCMC) method [69, 71]. Starting from an initial sample, a chain of samples is built sequentially. An iteration within a Markov chain is required to be only dependent on the previous iteration of the chain, and thus given an adequate length chain an MCMC method is able to accurately sample from a distribution when independent sampling is unpragmatic. In this thesis an MCMC method is used to draw random realisations of virtual patients arterial networks from the distribution

2.3. METROPOLIS-HASTINGS MARKOV CHAIN MONTE-CARLO

created through the use of Bayes' theorem (see Section 2.2).

While there are many different MCMC methods, within this thesis a Metropolis-Hastings MCMC [40] is used. To draw samples from a distribution, an MCMC chain requires an initialising position, described as $I = i^{(k=0)}$. The only restriction imposed on the initialising position of a Metropolis-Hastings MCMC chain is that all parameters must be within their support, and so the density of the distribution at this position must be greater than zero [28]. Applying this to the posterior distribution described by Equation (2.7), the restriction on the initialising position is:

$$p(I = i^{(k=0)} | J) > 0. \quad (2.10)$$

From this initial position a candidate for the next sample is generated by a proposal distribution which depends only on $I = i^{(k=0)}$, ensuring the Markov chain has no memory (an important feature of the method). In this thesis this candidate sample $I = i^*$ is proposed by sampling from the following Normal distribution:

$$i^* \sim \mathcal{N}(i^{(k=0)}, \sigma_{\text{step}}^2), \quad (2.11)$$

where σ_{step}^2 represents the variance of the proposal distribution. This proposed candidate $I = i^*$ is either accepted or rejected based on the following ratio κ of posterior probabilities at $I = i^*$ and $I = i^{(k=0)}$:

$$\kappa = \frac{p(I = i^* | J)}{p(I = i^{(k=0)} | J)}, \quad (2.12)$$

which, through the Bayes' rule of Equation (4.17) applied to both the numerator and the denominator, can be written as:

$$\kappa = \frac{p(J|I = i^*) p(I = i^*)}{p(J|I = i^{(k=0)}) p(I = i^{(k=0)})}. \quad (2.13)$$

The first term and the second term in both the numerator and the denominator are the likelihoods and priors at the candidate and current point, respectively. To accept or reject the proposed candidate κ is compared to a random number v drawn from a uniform distribution between 0 and 1. If $\kappa \geq v$, the proposed candidate $I = i^*$ is accepted and the chain progresses with $i^{(k=1)} = i^*$. Otherwise, the proposed candidate is rejected and a new candidate is proposed through Equation (2.11).

2.3. METROPOLIS-HASTINGS MARKOV CHAIN MONTE-CARLO

Using a constant variance in the proposal distribution (see Equation (2.11)) ensures the acceptance criteria described above is reversible: the probability of accepting the sample $I = i^{(*)}$ given the current state of the chain is described by $I = i^{(k=0)}$ is equal to the probability of accepting $I = i^{(k=0)}$ given the current state of the chain is described by $I = i^{(*)}$. This feature is a requirement of the Metropolis-Hastings method [25]. The process of sampling candidate points through Equation (2.11), computing κ , and either accepting or rejecting the candidate is repeated until the chain is sufficiently long.

The acceptance rate, *i.e.* the proportion of candidate sample points that are accepted, is tuned through the variance used in Equation (2.11). A high variance results in the Markov chain exploring the support of posterior distribution quickly, however a high proportion of candidate points being rejected. A low variance results in the chain moving around the posterior distribution slowly, however a high proportion of candidate points being accepted. An optimum acceptance rate to balance these two opposing behaviours is stated in literature to be 0.234 [173].

The MCMC algorithm is sequential—each subsequent sample depends on the previous sample and hence the chain grows only one sample at a time. Generating a long chain this way, thus, leads to very high computational run times. To alleviate this issue and achieve some level of parallelisation, the concept of pre-fetching [24, 196] is employed in this thesis, explained next.

2.3.1 Pre-fetching

At each iteration of the Markov chain there are two possible outcomes that may happen; the candidate sample is accepted, or the candidate sample is rejected. If the candidate sample is accepted it becomes the new current state of the chain, progressing the Markov chain forward. If the candidate sample is rejected, the current state of the Markov chain remains unchanged. By extending this process of either accepting or rejecting at each step of the Markov chain to view multiple steps, a decision tree is formed. Figure 2.2 shows such a decision tree with a depth of $\eta = 2$. This results in 2^η leaf-nodes at the end of the tree. A careful consideration of the decision tree (see Figure 2.2) shows that the only unique parameters in the entire decision tree correspond to those at the leaf-nodes—due to the fact that when a candidate is rejected the Markov chain does not

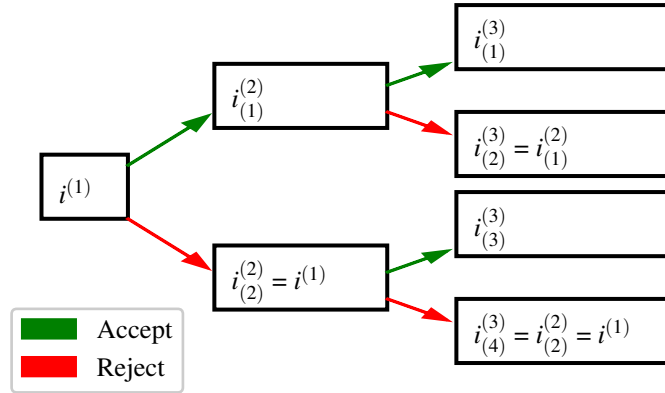


Figure 2.2: An example of an MCMC decision tree across two iterations of a chain. $i_{(l)}^{(k)}$ represents the l^{th} possible set (candidates) of the arterial network parameter scaling terms at the k^{th} step of the chain. Green branches denote the path of the chain taken if a candidate is accepted, while red branches show the path of the chain if a candidate is rejected.

progress. Thus, computational simulations for all the leaf-node parameters can be run simultaneously in-parallel. Then, η steps of the MCMC algorithm can be taken together without any computational overhead by walking the decision tree.

The wall clock time required to complete each decision tree is ideally constant irrespective of η . This is due to the fact that the time required to both propose and walk the tree are assumed to be negligible, while computation of the posterior probability—and so the time required to compute the acceptance ratio—is being parallelised. Thus, the wall clock time required to complete an MCMC chain should decrease linearly with η . It is likely, however, that inefficiencies will cause an increase in wall clock time per decision tree as η increases. Despite this likely reduction in performance, the major limiting factor associated with pre-fetching is the computational resources associated with it. Ideally the entire MCMC chain could be proposed, solved, and walked within a single decision tree. The number of terminal nodes, and so consequently the number of times the posterior probability is being computed in parallel, is equal to 2^η . The computational impracticality of using high values of η can be seen.

2.4 Machine learning classifiers

A model mapping a vector of input measurements, \mathbf{y} , to a discrete output classification, z , can be described as:

$$z = m(\mathbf{y}) \quad z \in \mathcal{C}, \quad (2.14)$$

with

$$\mathcal{C} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}, \dots, \mathcal{C}^{(j)}\}, \quad (2.15)$$

where \mathcal{C} represents the set describing all possible classifications, and $\mathcal{C}^{(j)}$ represents the j^{th} possible classification. For the case of this thesis, the measured inputs and known output classification represent the haemodynamic measurements taken from patients and the corresponding health of those patients respectively. An ML classifier is one possible method that can be used to learn the model of Equation (2.14).

All ML classifiers within this thesis are supervised algorithms. A database of input measurements and corresponding output classifications for a large number of previously observed events can be described as $\mathcal{D} = \{(\mathbf{y}_i, z_i) \mid i = 1..m\}$. Here \mathbf{y}_i and z_i represent the vector of input measurements and the true state classification of event i ; and m represents the total number of previously observed events. A supervised ML classifier splits the available database into a training set and a test set. The purpose of the training set is to allow the ML classifier to analyse the relationship between input measurements and true state output classifications, based on an algorithm specific optimisation criteria, to obtain a pseudo-mapping:

$$\hat{z} = \hat{m}(\mathbf{y}), \quad (2.16)$$

where \hat{z} represents the predicted classification resulting from the input measurements \mathbf{y} , and \hat{m} represents the pseudo-mapping. The purpose of the test set is to evaluate the accuracy of the pseudo-mapping, described by Equation (2.16). The input measurements corresponding to all events within the test set are passed through Equation (2.16) to obtain a prediction of the classification each event belongs to. The accuracy is then assessed by evaluating the discrepancy between the true state and the predicted classification of all events within the test set.

The methods used to create pseudo-mappings, as described by Equation (2.16),

are outlined next.

2.4.1 Naive Bayes' classifier

A Naive Bayes' classification algorithm is a probabilistic multiclass method [172, 171]. Naive Bayes' classifiers predict the health classification of a patient through the use of a conditional probability model. Bayes' theorem can be used to compute the conditional probability of a patient belonging to each health classification, given the input measurements taken from that patient have been observed:

$$p(\hat{z}_i = \mathcal{C}^{(j)} \mid \mathbf{y} = \mathbf{y}_i) = \frac{p(\mathbf{y} = \mathbf{y}_i \mid \hat{z}_i = \mathcal{C}^{(j)}) p(\hat{z}_i = \mathcal{C}^{(j)})}{p(\mathbf{y} = \mathbf{y}_i)}, \quad (2.17)$$

where the posterior probability, $p(\hat{z}_i = \mathcal{C}^{(j)} \mid \mathbf{y} = \mathbf{y}_i)$, represents the probability of the i^{th} VP being assigned a health classification of $\mathcal{C}^{(j)}$, given that the input measurements \mathbf{y}_i have been taken from the VP. By computing the posterior probability of each health classification, the classification that the patient is most likely to belong to can be found, *i.e.*:

$$\hat{z}_i = \arg \max_{\mathcal{C}^{(j)} \in \mathcal{C}} p(\hat{z}_i = \mathcal{C}^{(j)} \mid \mathbf{y} = \mathbf{y}_i). \quad (2.18)$$

In Equation (2.17) $p(\mathbf{y} = \mathbf{y}_i \mid \hat{z}_i = \mathcal{C}^{(j)})$ represents the likelihood of the measurements, given that the patient is predicted to belong to health classification $\mathcal{C}^{(j)}$. To simplify the computation of the likelihood, a naive Bayes' classifier assumes that all inputs are independent. The likelihood term can, therefore, be written as:

$$p(\mathbf{y} = \mathbf{y}_i \mid \hat{z}_i = \mathcal{C}^{(j)}) = \prod_{n=1}^N p(y_i^{(n)} = y_i^{(n)} \mid \hat{z}_i = \mathcal{C}^{(j)}), \quad (2.19)$$

where $p(y_i^{(n)} = y_i^{(n)} \mid \hat{z}_i = \mathcal{C}^{(j)})$ represents the likelihood corresponding to the n^{th} input, and N represents the total number of measurements being taken. To allow for computation of each component within Equation (2.19) an assumption must be made about the distribution of the input measurements. Assuming all input

measurements are normally distributed [139], Equation (2.19) can be written as:

$$p(\mathbf{y} = \mathbf{y}_i \mid \hat{z}_i = \mathcal{C}^{(j)}) = \prod_{n=1}^N \frac{1}{\sigma^{(n,j)} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{y_i^{(n)} - \mu^{(n,j)}}{\sigma^{(n,j)}}\right)^2\right), \quad (2.20)$$

where $\mu^{(n,j)}$ and $\sigma^{(n,j)}$ represents the mean and standard deviation of the n^{th} input measurement given that the i^{th} VP is predicted to belong to health classification $\mathcal{C}^{(j)}$. The mean and standard deviation of each input measurement given each health classification is found empirically through the training data. By splitting the training data in subcategories based on the true state health classification, the mean and standard deviation of each input measurement given each health classification is found.

Within Equation (2.17) $p(\hat{z}_i = \mathcal{C}^{(j)})$ represents the probability of the i^{th} VP belonging to health classification $\mathcal{C}^{(j)}$ irrespective of the input measurements. As with the likelihood term, the parameters describing the statistics of the prior probability for each health classification can empirically be found from the training data:

$$p(\hat{z}_i = \mathcal{C}^{(j)}) = \frac{m^{(j)}}{m}, \quad (2.21)$$

where $m^{(j)}$ represents the number of patients in the training set for which $z_i = \mathcal{C}^{(j)}$, and m represents the total number of patients within the training set.

The final term in Equation (2.17), $p(\mathbf{y} = \mathbf{y}_i)$, represents the probability of the input measurements \mathbf{y}_i occurring irrespective of the health classification of a patient. While the parameters describing the statistics of the input measurements irrespective of the health classification can be empirically found from the training set, this is not required. It can be seen that this term is independent of the health classification, and so constant given a particular vector of values for \mathbf{y} . This term is, therefore, not required when comparing the posterior probability of different health classification for a given patient.

2.4.2 Logistic regression

The LR classifier [194, 86] is a probabilistic binary classification method. Analysis of the relationship between the measured inputs and the known output classification

of the events within the training set allow a partition to be made between positive and negative responses to input variables through the high dimensional input space. A prediction of the classification of a new unseen event, *i.e.* an example of a set of input measurements taken from the test set, is then made through the analysis of the recorded input measurements corresponding to that event in relation to this partition.

Given that patients belong to one of the two classifications, *i.e.* $\mathcal{C} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}\}$, the true state binary health ψ of patients is assigned to all subjects in the training set:

$$\psi_i = \begin{cases} 1 & \text{if } z_i = \mathcal{C}^{(1)} \\ 0 & \text{if } z_i = \mathcal{C}^{(2)}, \end{cases} \quad (2.22)$$

To predict the binary health of a patient an activation function is used. A general equation for an activation function $h(\mathbf{y}_i, \boldsymbol{\theta})$ can be written as:

$$p(\hat{\psi}_i = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}) = h(\mathbf{y}_i, \boldsymbol{\theta}), \quad (2.23)$$

where $p(\hat{\psi}_i = 1 \mid \mathbf{y}_i, \boldsymbol{\theta})$ represents the predicted probability that the i^{th} VP belongs to $\mathcal{C}^{(1)}$, given that the patient specific input measurements \mathbf{y}_i have been observed, and that the vector of measurement specific weightings are described by $\boldsymbol{\theta}$. An activation function is visually shown in Figure 2.3. Typical choices for $h(\mathbf{y}_i, \boldsymbol{\theta})$ are the sigmoid and tanh functions. The sigmoid function is shown below:

$$h(\mathbf{y}_i, \boldsymbol{\theta}) = \frac{1}{1 + \exp(-\boldsymbol{\theta}^T \mathbf{y}_i)}. \quad (2.24)$$

To obtain optimal measurement specific weightings $\boldsymbol{\theta}$, the logistic regression algorithm is trained by minimising the mean error between the predicted probability of VPs belonging to a positive binary classification and the true state classification across the training set, *i.e.*:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \{L(\boldsymbol{\theta}, \mathbf{Y}^{\text{train}}, \boldsymbol{\psi}^{\text{train}})\}, \quad (2.25)$$

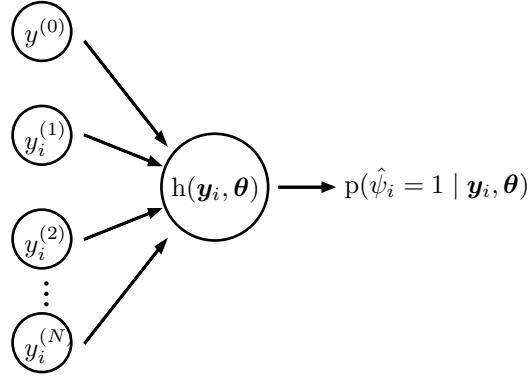


Figure 2.3: A logistic regression algorithm returns a prediction of the probability of an observed set of input data resulting in a positive binary classification using a single activation function. $\{y_i^{(1)}, \dots, y_i^{(k)}\}$ represents the 1st to the N^{th} measurement taken from patient i , and $y^{(0)}$ represents a bias term—and so is patient independent.

with:

$$L(\boldsymbol{\theta}, \mathbf{Y}^{\text{train}}, \boldsymbol{\psi}^{\text{train}}) = -\frac{1}{m} \sum_{i=1}^m \left(\psi_i \log(h(\mathbf{y}_i, \boldsymbol{\theta})) + (1 - \psi_i) \log(1 - h(\mathbf{y}_i, \boldsymbol{\theta})) \right), \quad (2.26)$$

where $L(\boldsymbol{\theta}, \mathbf{Y}^{\text{train}}, \boldsymbol{\psi}^{\text{train}})$ represents the average cost, in this case computed as a log loss, across the training set; $\mathbf{Y}^{\text{train}}$ and $\boldsymbol{\psi}^{\text{train}}$ represent the matrix of input measurements and the vector of the known correct binary classifications for all the m VPs in the training set, respectively; \mathbf{y}_i and ψ_i represents the vector of input measurements and the known true state classification corresponding to the i^{th} VP, respectively; and $\boldsymbol{\theta}$ represents the measurement specific weightings. It can be seen from Equation (2.26) that two different cost profiles are created dependent on the correct binary classification of a particular patient. The numerical minimisation can be carried out using many algorithms such as gradient descent, gradient descent with momentum [161], Nesterov accelerated gradient (NAG) [141], Adadelata [223], and Adam method [109]. All of these algorithms follow a similar procedure:

- Initialising values are assigned to all measurement specific weightings.

- Initial predictions are made using Equation (2.24) and these weightings.
- The cost of these initial predictions are computed through the use of Equation (2.26).
- An update function is used to update the vector of weightings to minimise the error in the predictions made. An update function is described by the general equation:

$$\boldsymbol{\theta}_{s+1} = \boldsymbol{\theta}_s + \Delta\boldsymbol{\theta}_s, \quad (2.27)$$

where $\boldsymbol{\theta}_s$ and $\Delta\boldsymbol{\theta}_s$ represents the vectors of measurement specific weightings and the updates to be made to these weightings at iteration s , respectively.

- The sequential process of predicting the probability of each patient within the training set belonging to a positive binary classification, calculating the cost of these predictions, and updating the weightings based on this cost is repeated until a set standard of convergence is reached.

Post training, the obtained weightings can be used to predict the health classification of new unseen VPs, *i.e.* VPs within the test set, by equation (2.24) through application of a threshold \mathcal{B} , often referred to as the decision boundary, to the predicted probabilities as follows:

$$\hat{z}_i = \begin{cases} \mathcal{C}^{(1)} & \text{if } p(\hat{\psi}_i = 1 | \mathbf{y}_i, \boldsymbol{\theta}) \geq \mathcal{B} \\ \mathcal{C}^{(2)} & \text{otherwise,} \end{cases} \quad (2.28)$$

where \hat{z}_i represents the predicted health classification of the new unseen test VP, $p(\hat{\psi}_i = 1 | \mathbf{y}_i, \boldsymbol{\theta})$ represents the predicted probability returned by the activation function through equation (2.24), and \mathcal{B} represents a chosen decision boundary.

2.4.3 Multi-layer Perceptron

A multi-layer perceptron (MLP) is a form of artificial neural network [140]. An understanding of an MLP can be gained by first examining the structure of an LR algorithm. An LR classifier passes a series of measured inputs into an activation function, which contains pre-trained measurement specific weightings.

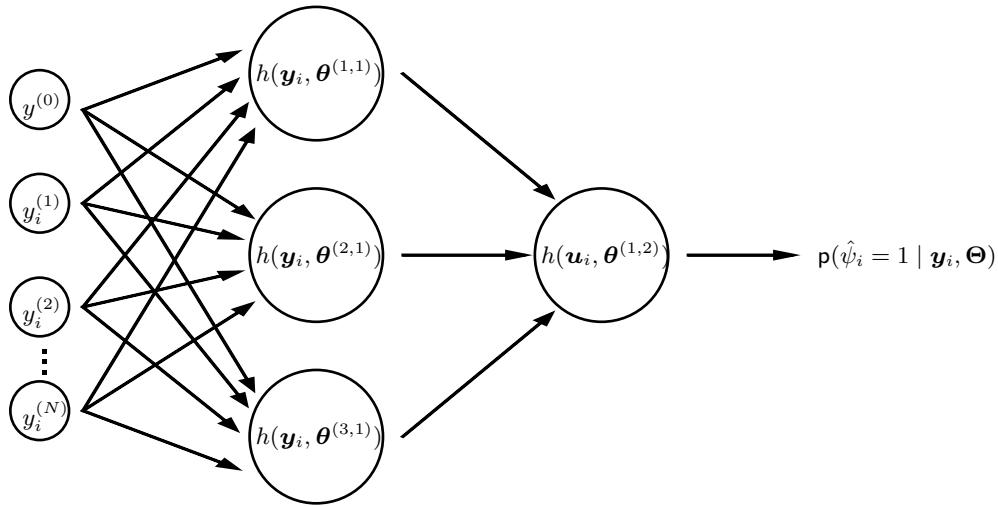


Figure 2.4: A simple MLP classifier, containing one hidden layer consisting of three neurons. $\{y_i^{(1)}, \dots, y_i^{(N)}\}$ represents the 1st to the N^{th} measurement taken from patient i , $y^{(0)}$ represents the bias term, and $\boldsymbol{\theta}^{(i,j)}$ represents the measurement specific weightings associated with the i^{th} parallel activation function within the j^{th} sequential layer. In the above figure \mathbf{u}_i represents the vector of inputs into the output layer, and within the case shown above is equal to $\mathbf{u}_i = \{h(\mathbf{y}_i, \boldsymbol{\theta}^{(1,1)}), h(\mathbf{y}_i, \boldsymbol{\theta}^{(2,1)}), h(\mathbf{y}_i, \boldsymbol{\theta}^{(3,1)})\}$. $\boldsymbol{\Theta}$ represents the matrix of input specific weightings for all activation function within the MLP.

This activation function returns a prediction of the probability of the input data belonging to a positive binary classification. The architecture of the single activation function within a logistic regression classifier is shown in Figure 2.3.

While an LR classifier is simple, and so requires little problem specific optimisation, it is limited by the fact it can only produce linear partitions between positive and negative classifications through the high dimensional input space. A solution to this problem is to create a network of sequential layers containing multiple parallel activation functions. Additional sequential layers of activation functions between the input and output of an MLP are referred to as hidden layers. An example of a simple MLP containing 1 hidden layer, which in turn consists of 3 activation functions—referred to as neurons—is shown in Figure 2.4. MLPs can be constructed, in theory, using any number of hidden layers each consisting of any number of neurons.

It can be seen from Figure 2.4 that while the first hidden layer within an MLP returns an output based on the raw input measurements, all subsequent layer use a transformed version of the input data. The sequential processes of passing the input data through each layer of an MLP allows for the creation of complex inter-measurement features. As with an LR classifier the output layer of an MLP returns a prediction of the probability of a set of inputs belonging to a positive binary classification, given the input measurements and a series of input specific weightings.

A further similarity with LR is that an MLP is trained with the primary objective of minimising the error between the predicted probability and the correct classification across the training set. The complexity of training an MLP is increased from that of LR, however, by the fact that the error in the predictions made is only computed at the output layer. The error in the predictions across the training set must, therefore, be “back propagated” [83] to compute the gradient of the error with respect to each input specific weighting within each neuron of each hidden layer. The weightings with each neuron of an MLP can then be updated using Equation (2.27). An MLP is trained by repeatedly feeding the input measurements forward from the input to the output layer, computing the error between the predicted probabilities returned and the known correct health classification of all patients within the training set, back propagating this error, and then updating the input specific weightings. This processes is repeated until a set standard of convergence is reached.

2.4.4 Support vector machine

A support vector machine (SVM) classifier also forms a partition between positive and negative examples of an event through a high dimensional feature space [103]. In the case of LR and MLP speculative classifications are repeatedly proposed for the training data, and the weights assigned to each input measurement optimised to minimise the error in prediction. In the case of an SVM classifier, a hyperplane is found that maximises the distance between itself and the nearest instances of both classifications.

Given that all patients belong to the same two classifications seen in the examples of LR and MLP, *i.e.* $\mathcal{C} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}\}$, an SVM classifier assigns a true state classification of 1 to all patients belonging to $\mathcal{C}^{(1)}$, and -1 to all patients belonging

to $\mathcal{C}^{(2)}$:

$$\psi_i = \begin{cases} 1 & \text{if } z_i = \mathcal{C}^{(1)} \\ -1 & \text{if } z_i = \mathcal{C}^{(2)}. \end{cases} \quad (2.29)$$

A general equation describing a linear hyper plane is given by:

$$d(\mathbf{y}, \boldsymbol{\theta}, y^{(0)}) = 0, \quad (2.30)$$

where

$$d(\mathbf{y}, \boldsymbol{\theta}, y^{(0)}) = \boldsymbol{\theta}^T \mathbf{y} + y^{(0)}, \quad (2.31)$$

and $\boldsymbol{\theta}$ represents the vector of measurement specific weightings, \mathbf{y} represents the position within the input measurement space, and $y^{(0)}$ represents a bias term. The purpose of an SVM classifier is to find optimum values for $\boldsymbol{\theta}$ and $y^{(0)}$ that best partition positive and negative events. Once an optimum partition has been found, classifications are assigned to new unseen VPs based on the inputs measurements position relative to this separating hyperplane, *i.e.*:

$$\hat{\psi}_i = \begin{cases} 1 & \text{if } d(\mathbf{y}_i, \boldsymbol{\theta}, y^{(0)}) > 0 \\ -1 & \text{if } d(\mathbf{y}_i, \boldsymbol{\theta}, y^{(0)}) < 0, \end{cases} \quad (2.32)$$

where \mathbf{y}_i represents the input measurements belonging to the test patient, and $\hat{\psi}_i$ represents the predicted label assigned to the patient.

To determine the best partition between positive and negative events, the “margin” is used. The nearest instances of each classification to the partitioning hyperplane are referred to as the support vectors (SVs). The margin represents the distant between the hyperplanes tangential to the partition, that pass through the SVs. An example highlighting the SVs and the margin in a simple two dimensional case is shown in Figure 2.5.

To simplify the mathematics associated with the computation of the hyperplane that maximises the margin, the partitioning hyperplane is found in its canonical form with respect to the training data. The partitioning hyperplane is in its canonical form if:

$$\min_{\mathbf{y}_i \in \mathbf{Y}^{\text{train}}} |d(\mathbf{y}_i, \boldsymbol{\theta}, y^{(0)})| = 1. \quad (2.33)$$

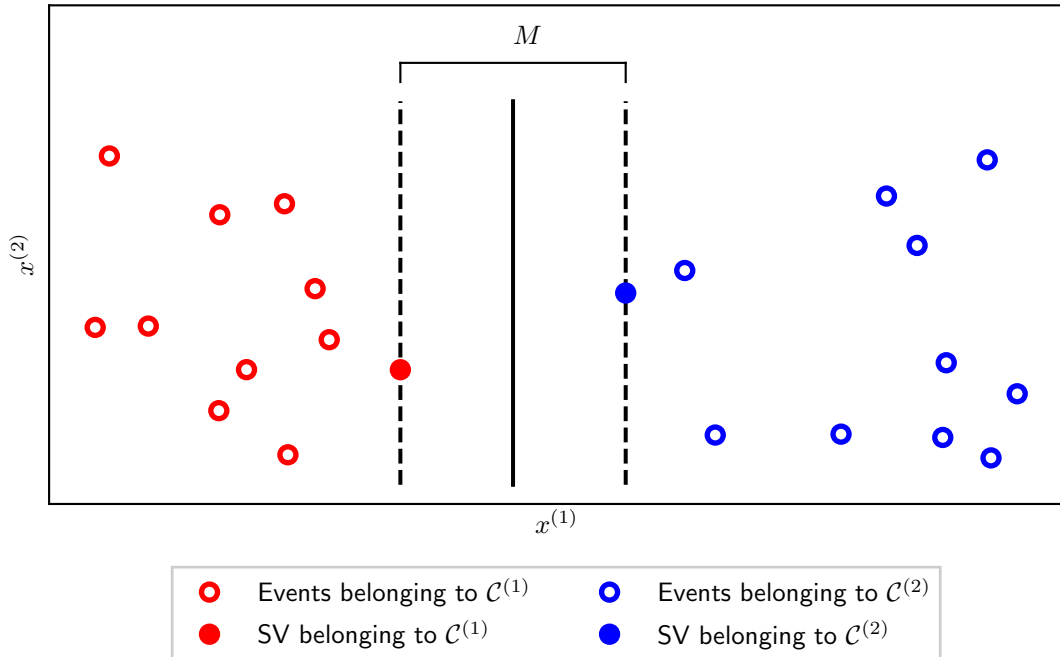


Figure 2.5: Multiple instances of two different classifications are shown. The support vector corresponding to each of the two classifications are highlighted, and the resulting margin is represented by M . The partition between positive and negative events is shown by the solid black line.

The margin associated with the partition is then equal to [102]:

$$M = \frac{2}{\|\boldsymbol{\theta}\|}, \tag{2.34}$$

where:

$$\|\boldsymbol{\theta}\| = \sqrt{\sum_{n=1}^N (\theta^{(n)})^2}, \tag{2.35}$$

and M represents the margin, $\boldsymbol{\theta}$ represents the vector of measurement specific weightings, and $\theta^{(n)}$ represents the n^{th} input measurement weighting. It can be seen that the margin is maximised if $\|\boldsymbol{\theta}\|$ is minimised. The minimisation of $\|\boldsymbol{\theta}\|$ is treated as a quadratic optimisation problem, and the derivation of a solution through the saddle point of the Lagrange function is found in [103].

2.4.5 Random forest

A random forest (RF) classifier comprises of an ensemble of parallel weak decision tree classifiers [123, 23]. A decision tree classification algorithm assigns a predicted classification to a patient based on the “answers” given to a series of pre-determined “questions”.

The training of a decision tree is initialised by obtaining the attribute—a feature of the input measurements—that maximises an evaluation criteria—in this thesis the information gained—when used to partition the full training set. The information gained by partitioning the training set based on attribute c is computed using the equation:

$$\text{IG}(\mathbf{Y}^{\text{train}}, c) = \text{E}(\mathbf{Y}^{\text{train}}) - \text{E}(\mathbf{Y}^{\text{train}} | c), \quad (2.36)$$

where $\text{IG}(\mathbf{Y}^{\text{train}}, c)$ represents the information gained when the data set $\mathbf{Y}^{\text{train}}$ is split using attribute c ; and $\text{E}(\mathbf{Y}^{\text{train}})$ and $\text{E}(\mathbf{Y}^{\text{train}} | c)$ represents the entropy of the data set $\mathbf{Y}^{\text{train}}$ and the entropy when the data is split based on attribute c , respectively. The entropy of a data set can be thought of as the unpredictability of the classifications within the set, and so is inversely related to the homogeneity. The entropy of the training data is computed using the equation:

$$\text{E}(\mathbf{Y}^{\text{train}}) = \sum_{n=1}^N \frac{m^n}{m} \log_2 \frac{m^n}{m}, \quad (2.37)$$

where N represents the length of \mathcal{C} (see Equation (2.14)), m represents the total number of patients within the data set, and m^n represents the number of patients within the data set for which $z_i = \mathcal{C}^{(n)}$. The entropy of the training set given the data is being split based on attribute c is found through the weighted mean of the entropy of the two resulting daughter nodes. By computing the information gain achieved by splitting the data set using all available attributes, the attribute that results in the greatest information gain—and so the greatest homogeneity at the resulting daughter nodes—is found.

Training of a decision tree is progressed by sequentially splitting the training set, with the attribute that maximises the information gain at each node in each generation of the tree being found. This process of repeatedly splitting the data based on the optimum attribute to maximise information gain is repeated until a

stopping criteria is met. To assign a non-probabilistic classification to each terminal node, the classification with the highest probability of occurrence is found. When classifying new data, in this case determining the health of new patients, the tree is walked from the root node to the appropriate terminal node.

A major issue with decision trees is that they are very sensitive to the data they are trained on, *i.e.* they have high variance. To overcome this issue a random forest classification algorithm trains a series of decision trees employing a bootstrap aggregation method. Bootstrap sampling refers to the process of uniformly sampling with replacement. In the case of an RF algorithm a series of decision trees are trained, each using a different subset of the training data created through bootstrap sampling. When testing new unseen patients, the health classification returned by each tree within the RF is found. The results of each tree within the forest are then aggregated by computing the classification most frequently predicted. Generally more decision trees within an RF will result in higher accuracy classification, up to an asymptotic limit. This increase in accuracy, however, comes at the cost of increased computational time. It is therefore important to consider the number of trees used when training a random forest algorithm.

2.4.6 Gradient boosting classifier

In a similar manner to RF, a gradient boosting (GB) classifier [67, 57] is an ensemble classifier. The output of a series of weak decision trees are combined to achieve higher accuracy results. In the case of RF a series of decision trees are created independently, and a classification assigned to test patients by averaging the results across these trees. A GB classifier differs from this by instead creating classifiers within the ensemble sequentially. Each new classifier within the ensemble is trained with respect to the error of all previous classifiers.

A GB classifier is a binary ML classifier, and so true state binary classifications are assigned to VPs through Equation (2.22). Training of a GB classifier is initialised by computing the probability of a patient belonging to a positive true state classification, irrespective of the recorded input measurements. This initial naive probability $p(\hat{\psi}_i = 1)$ is found using the equation:

$$p(\hat{\psi}_i = 1) = \frac{m^{(\psi_i=1)}}{m}, \quad (2.38)$$

where $m^{(\psi_i=1)}$ represents the number of patients within the training set who have a positive true state classification, and m represents the total number of patients within the training set. The residual between the naive prediction and the correct binary classification is computed and recorded for all patients within the training set using the equation:

$$\chi_i = \psi_i - p(\hat{\psi}_i = 1), \quad (2.39)$$

where χ_i represents the residual associated with patient i . A decision tree (see Section 2.4.5) is created to predict the residual associated with each patient in the training set, based on the input measurements. A probability correction factor γ is then computed for each terminal leaf within this decision tree using the equation:

$$\gamma = \frac{\sum_{i=1}^{m^{\text{leaf}}} \chi_i}{\sum_{i=1}^{m^{\text{leaf}}} p(\hat{\psi}_i = 1) \times (1 - p(\hat{\psi}_i = 1))}, \quad (2.40)$$

where m^{leaf} represents the number of patients in the appropriate leaf node. The initial naive probability of patients having a positive binary classification is updated by including a scaled version of the probability correction factor corresponding to the terminal leaf to which the patient belongs, *i.e.*:

$$p(\hat{\psi}_i = 1 \mid \mathbf{y}_i) = p(\hat{\psi}_i = 1) + \alpha\gamma, \quad (2.41)$$

where α represents the learning rate. The residual associated with each patient in the training set is updated based on the new predicted probability, and the processes of creating a decision tree to predict the residual, calculating the probability correction factor, and incorporating this into the prediction is repeated. This sequential process can in theory be repeated any number of times.

The predicted probability of disease associated with test patients is computed by first finding the terminal leaf to which the patients belongs for all decision trees within the ensemble. The predicted probability is then computed through the summation of the the scaled probability correction factors associated with these terminal leaves, *i.e.*:

$$p(\hat{\psi}_i = 1 \mid \mathbf{y}_i) = p(\hat{\psi}_i = 1) + \sum_{n=1}^N \alpha\gamma^{(n)}, \quad (2.42)$$

where $p(\hat{\psi}_i = 1 \mid \mathbf{y}_i)$ represents the predicted probability of patient i belonging to a positive binary classification given the input measurements \mathbf{y}_i have been recorded, $p(\hat{\psi}_i = 1)$ represents the initial naive probability irrespective of the recorded measurements, N represents the number of decision trees within the ensemble, and $\gamma^{(n)}$ represents the probability correction factor associated with the n^{th} decision tree.

In a similar manner to an LR classifier, the output of a GB is not a direct prediction of the classification a patient belongs too. Instead a prediction of the probability of a patient belonging to a positive binary classification is found. To turn this predicted probability into a output classification a decision boundary is applied, as is shown in Equation (2.28).

2.4.7 Motivation for the choice of methods

While the six aforementioned classification methods have all been previously described as ML classifiers (for ease of description), MLP classifiers can be considered a deep learning method. Other examples of deep learning methods—which are not employed in this thesis—include convolutional neural network, and recurrent neural networks. The purpose of this thesis is to perform an initial exploratory investigation into the possibility of using ML classifiers to detect different forms of arterial disease. Focus is, therefore, on uncovering patterns and behaviours—such as which haemodynamic measurements are particularly informative—rather than optimisation to achieve increasingly higher accuracy. With consideration for this objective, it is not feasible to perform extensive optimisation and analysis on every single ML classifier trained and tested. Thus, the ML methods used are chosen based on their “robustness”—*i.e.* insensitivity to the hyper-parameters used and unsusceptibility to problems such as over-fitting—relative to more complex deep learning methods. It is difficult to make an *a priori* prediction of which modelling approaches (tree-based, kernel-based, Bayesian, or neuron-base) are best suited to haemodynamic classification. Thus, the methods chosen encompass a range of probabilistic and non-probabilistic applications of different modelling approaches (see Table 2.1)—to allow for an *a posteriori* evaluation of which approaches are preferential for haemodynamic problems—while fulfilling the aforementioned characteristics. Along with the five ML methods, the one deep learning method (MLP) is also employed for comparison. It is *a priori* expected that multi-layer

Modelling approach	Non-probabilistic	Probabilistic
Tree-based	RF	GB
Kernel-based	SVM	
Bayesian		NB
Neuron-based		LR, MLP

Table 2.1: The four different modelling approaches and how each classification method aligns with these approaches.

perceptron classifiers will not perform to their full potential in this thesis, as they are more reliant on complex hyper-parameter optimisation and monitoring for overfitting than the five ML methods. The use of multi-layer perceptron will, however, provide some, albeit limited, comparison of ML and deep learning methods.

Chapter 3

Three vessel network proof-of-concept study

3.1 Introduction

The aim of this proof-of-concept (PoC) study is to carry out an initial investigation into the potential of using ML classification algorithms to predict the presence of arterial disease, using haemodynamic measurements. The likely considerations and limitations that are involved with both creating and exploiting synthetic data sets are also assessed. Thus, priority is on quick and inexpensive creation of the virtual patient database (VPD) over the physiological realism of the resulting data sets. This PoC is being carried out to create a template, that can then be expanded on to increase the complexity, and consequently the physiological realism, in subsequent chapters of this thesis.

A simple three vessel system representing the aortic bifurcation, as outlined in [20], is used as the basis for the create of virtual patients (VPs). The resulting VPD consists of healthy and stenosed patients. It is chosen to complete the PoC using stenosed patients, rather than aneurysm, as the majority of previous work has focused on the detection of aneurysms. While previously studies have already shown that the presence of an aneurysm produces measurable biomarkers within flow-rate profiles [183], to the authors' knowledge no comparable work has yet been completed for stenoses.

To create the VPD, *a priori* distributions are first constructed for the parameters describing the arterial network of VPs. Random realisations are sampled from these distributions, and the physics based model—as outlined in Chapter 2—is solved to obtain the corresponding pressure and flow-rate profiles. Similar one-dimensional models have been used to compute the pressure and flow-rate profiles associated with patients with the presence of an aneurysm in [33], [96] and [183]; and stenosis in [21] and [96]—suggesting the suitability of this model for use on both healthy and unhealthy patients. Finally, “hard” filters, *i.e.* the direct imposition of bounds on

the range of pressure profiles, are applied to the VPD to reduce the occurrence of physiological unrealistic VPs. This virtual population is then used to train and test a series of ML classifiers—also outlined in Chapter 2—to detect arterial disease, and test their performance. Focus is on uncovering why some ML methods perform better than others and which measurements (and their combinations) are more informative.

The majority of work in this chapter is published in: *Jones G., Parr J., Nithiarasu P., Pant S. A proof of concept study for machine learning application to stenosis detection. Medical and Biological Engineering and Computing. 2021 August <http://dx.doi.org/10.1007/s11517-021-02424-9>.*

3.2 Virtual patient database

3.2.1 Motivation and procedure

To train and test ML classifiers a large database of haemodynamic measurements taken from a comprehensive cohort of patients is required. The corresponding correct arterial health of these patients is also required. As opposed to using measurements taken from real patients, VPs are created using a physics based model of pulse wave propagation. As mentioned in Chapter 1, this VP approach has several advantages:

- **Expense:** creating VPs is relatively inexpensive.
- **Class imbalance:** creating VPs allows for the control of the distribution of different disease states.
- **Measurement availability:** using VPs allows for measurements of pressure and flow-rate to be taken at any location within the arterial system.

This final benefit is particularly important for this PoC where feasibility of the ML approach is being assessed. A primary purpose of this study is to gain an initial understanding of the patterns between the measurements and classification accuracy. The availability of both pressure and flow-rate measurements throughout the arterial network allows an in-depth *a priori* analysis of the importance of both different measurement types, *i.e.* pressure vs flow-rate, and different measurement locations to be carried out.

To create a VPD the following four steps are sequentially completed:

1. **Topology and parameterisation:** the topology and parameterisation of the arterial network of VPs is chosen. It is computationally unpragmatic to model the entire arterial system, and so a subnetwork of interest is identified. Prior and subsequent vessels to this subnetwork are lumped into appropriate boundary conditions. While creating VPs with a larger subnetwork of the arterial system, containing more vessels, increases the physiological realism of the resulting synthetic data set, it also increases the dimensionality and computational time associated with the creation of VPs. This is undesirable for a PoC analysis. The parameterisation of the arterial network being modelled must balance control over the random realisations of arterial networks that can be created, with the ability to create variability across the resulting VPD. Strict imposition of patterns and behaviours through the parameterisation of the arterial network reduces the occurrence of VPs with physiologically unrealistic pressure and flow-rate profiles, however also decreases the variability in arterial networks seen across the VPD.
2. **Probability distributions:** appropriate distributions for the parameters describing the arterial networks of VPs, across the resulting VPD, are chosen. It is not possible to obtain exact distributions from which VPs arterial networks should be sampled. Instead, an estimation of the distributions is made. The complexity of this estimation is likely to be proportional to the physiological realism of the resulting data set. The expense of creating the VPD must therefore be carefully balanced against the required physiological realism.
3. **Sampling to create the VPD:** random realisations of VPs arterial networks are sampled from the distributions described above. For each VP, the physics based model of pulse wave propagation is solved to obtain the corresponding pressure and flow-rate profiles. The representation of the pressure and flow-rate profiles being taken from VPs, which are used as input measurements into ML classifiers, is chosen to balance the capturing of information with the required dimensionality.
4. **Post processing:** post processing to create a usable data set from the raw

VPD is carried out. Filters are applied to the VPD to remove VPs with physiologically unrealistic pressure and flow-rate profiles.

Each of these aspects are described next.

3.2.2 Topology and parameterisation

Healthy subjects

This section presents the topology and parameterisation of the arterial network for the creation of VPs. Once parameterised, random realisations of network parameters can be sampled to create VPs. In this chapter, the network of interest is the abdominal aorta bifurcating into the two common iliacs. A pre-existing model for this is taken as the reference network from [20]. This is shown in Figure 3.1, where the three vessels (abdominal aorta and two common iliacs) are represented in 1D while suitable boundary conditions are imposed at the inlet and outlets. The significant truncation of the arterial network of VPs results in a reduction in the physiological realism of the resulting VPD, however a significant reduction in the dimensionality of VPs arterial networks is achieved.

It can be seen that each VP contains one time domain inlet flow-rate, three arterial vessels, and two terminal Windkessel models. Both the time domain inlet flow-rate profile and the terminal Windkessel models are described within Chapter 2. For each arterial vessel, the model also requires the assignment of the length, the wall thickness, the reference diameter, and the Young's modulus.

As mentioned within Chapter 2, the inlet volumetric flow-rate, $Q_{\text{inlet}}(t)$, is described using a Fourier series (FS). It is suggested in [180] that pressure profiles can be described using a FS truncated at the 6th harmonics. To test if this is true for flow-rate profiles, two flow-rate profiles are synthesised from the same FS, however truncated at the 5th and 6th harmonic respectively. These profiles are visually compared to the exact profile from which the FS coefficients are computed, as shown in Figure 3.2. Figure 3.2 shows that while there are noticeable differences between the exact profile and the FS synthesised up to the 4th order, no differences can be seen by eye between the expansion truncated at the 5th order. Thus, the

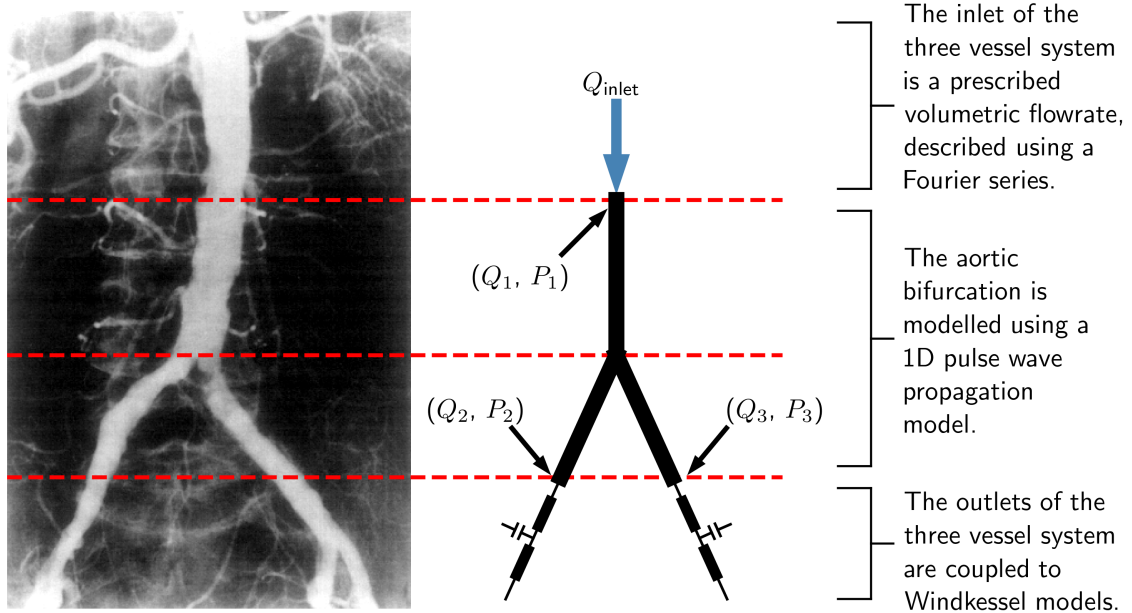


Figure 3.1: Coupling of the inlet and outlet boundary conditions to the numerical model are shown. How this model represents the aortic bifurcation is also shown through comparison to an angiogram [26].

time domain inlet flow-rate profile is described by:

$$Q_{\text{inlet}} = \{a_0 = 0, b_0, a_1, b_1, \dots, a_5, b_5\}, \quad (3.1)$$

where a_n and b_n represent the n^{th} sine and cosine FS coefficients respectively—resulting in the time domain inlet flow-rate profile requiring specification of 11 coefficients.

As a significantly truncated version of the arterial system is modelled to create VPs, the length of vessels within the network are relatively short. The length of the aorta and two iliacs within the reference model taken from [20] are 8.6cm and 8.5cm respectively. The short lengths of vessels within the arterial network results in there likely being very little variation of the baseline arterial vessel properties, *i.e.* excluding the variance introduced due to disease, along their lengths. It is, therefore, assumed that the baseline properties of all the three vessels are constant along their lengths.

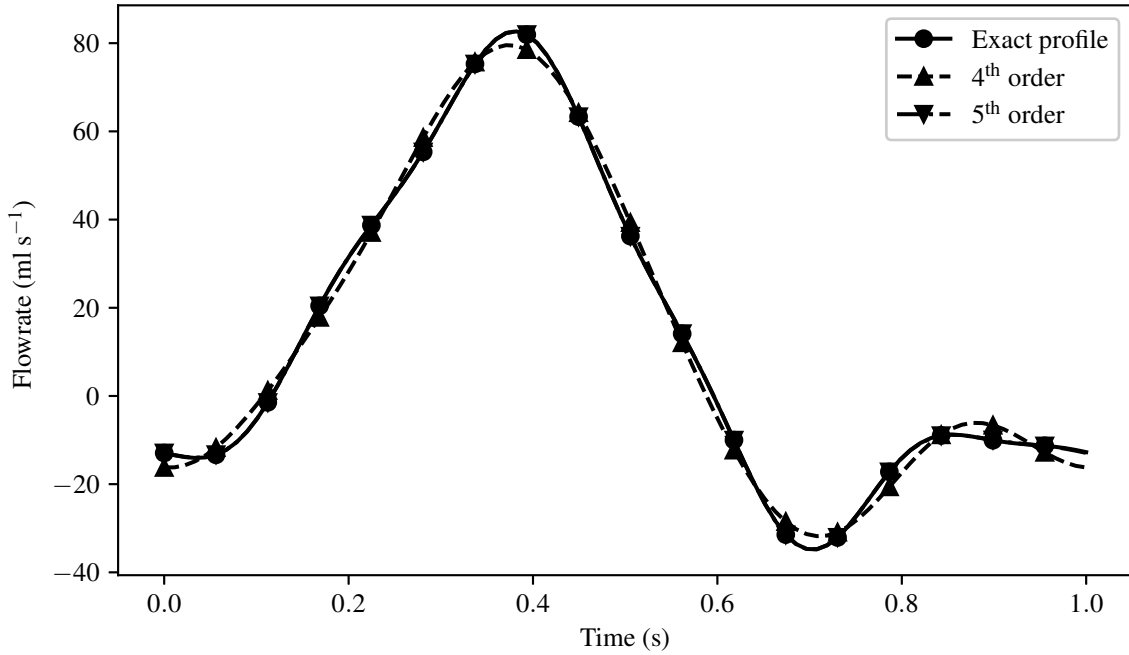


Figure 3.2: A comparison of the flow-rate profiles produced when expanding the same FS, however truncated at different orders.

To impose geometric and mechanical symmetry on the lower extremities, the two common iliac arteries are assumed to share baseline properties. This symmetry, however, is not extended to the terminal Windkessel model parameters. To complete the parameterisation of the arterial network of healthy VPs, it is assumed that $P_{\text{ext}} = 0$, and that the cardiac period of all VPs is the same. The parameterisation of the network, thus, requires specification of the following 25 parameters:

- **Six geometric properties:** the two common iliac arteries require a single length, a reference area, and a wall thickness. These properties are also required for the aorta.
- **Two mechanical properties:** the Young's modulus of the aorta and the common iliacs needs to be specified.
- **Six terminal boundary parameters:** each of the Windkessel models requires two resistances and a compliance.

- **11 FS coefficients:** the time domain inlet flow-rate profile is described using a FS truncated at the 5th order.

Disease subjects

For an ML classifier to be trained to distinguish between healthy and unhealthy patients, examples of both classifications are required within the VPD. A parameterisation must, therefore, be chosen to describe stenosed arterial vessels. For simplicity all VPs are limited to having a maximum of one diseased vessel. To create a change in the reference area of a diseased vessel a normalised map of each vessels area is produced. Both the length and cross sectional area of the vessel is normalised between 0 and 1. This map, for a 60% stenosis, is shown in Figure 3.3, where the x-axis represents the reference position along the length of the vessel and the y-axis represents the reference cross sectional area. For healthy vessels the normalised reference cross sectional area is constant and equal to 1. For unhealthy vessels a cosine curve is used to create a change in area. This cosine curve is scaled to create variation in location and severity of disease between patients. To scale the cosine curve three parameters are used. These parameters are the severity, the start location, and the end location of the disease represented by \mathcal{S} , o , and e respectively. The normalised cross sectional area (A_n) of a diseased vessel at a spatial location described by x_n is equal to:

$$A_n = \begin{cases} \left(1 - \frac{\mathcal{S}}{2}\right) + \frac{\mathcal{S}}{2} \cos\left(\frac{2(x_n - o)\pi}{e - o}\right) & \text{for } o \leq x_n \leq e \\ 1 & \text{otherwise} \end{cases} \quad (3.2)$$

Thus, in addition to the 25 parameters for the description of healthy subjects, three more parameters are required for specification of disease. Random realisations of these parameters are sampled and the physics based model of pulse wave propagation solved to produce each VP.

3.2.3 Probability distributions

Ideally the distribution of both arterial network parameters and the resulting pressure and flow-rate profiles should be representative of those measured in a real population. Since one-dimensional arterial network parameters are generally either

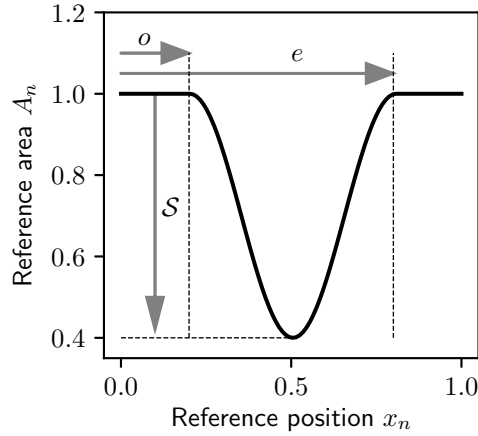


Figure 3.3: An example of a 60% stenosis is shown. This disease is created with a start location $o = 0.2$ and an end location $e = 0.8$.

expensive and invasive to obtain or non-physical (so cannot be directly measured), their exact distributions are not known. Thus, *a priori* distributions are assumed for both healthy and diseased virtual subjects, as described next.

Healthy subjects

A priori distributions are assumed for the arterial network parameters, based on values reported in literature [20]. It is assumed that across a large population all parameters required to describe VPs arterial networks, excluding disease parameters, are independent and normally distributed. The mean value for each of these parameters is taken from [20] and the standard deviation is set to be 20% of the mean, as summarised in Table 3.1. VPs are assigned disease so that the VPD consists of 50% healthy patients, and there is an equal number of aortic, first iliac, and second iliac stenoses VPs.

Disease subjects

In addition to the parameters described for healthy patients above, a diseased patient is characterised by three more parameters—disease severity, start location, and end location—which are assigned uniform distributions based on physical considerations. To impose a minimum possible length for stenoses a fourth parameter, the reference location of the disease (represented by r_e), is introduced. The four parameters

Parameter	Mean		Standard deviation	
	Aorta	Iliac	Aorta	Iliac
Length	8.6cm	8.5cm	1.72cm	1.7cm
Wall thickness	1.03mm	0.72mm	0.21mm	0.14mm
Reference diameter	1.72cm	1.2cm	0.344cm	0.24cm
Young's modulus	500kPa	700kPa	100kPa	140kPa
R_1	-	$6.81 \times 10^7 \text{ Pa s m}^{-3}$	-	$1.36 \times 10^6 \text{ Pa s m}^{-3}$
R_2	-	$3.10 \times 10^9 \text{ Pa s m}^{-3}$	-	$6.20 \times 10^8 \text{ Pa s m}^{-3}$
C	-	$3.67 \times 10^{-10} \text{ m}^3 \text{ Pa}^{-1}$	-	$7.33 \times 10^{-11} \text{ m}^3 \text{ Pa}^{-1}$

Table 3.1: Arterial network parameter mean and standard deviation values. The mean values for each parameter are taken from [20], and the standard deviation set to 20% of the mean

required to describe disease are sequentially sampled from uniform distributions within the following bounds:

$$\text{Constraints : } \begin{cases} 0.2 \leq r_e \leq 0.8 \\ 0.1 \leq o \leq r_e - 0.05 \\ r_e + 0.05 \leq e \leq 0.9. \\ 0.5 \leq \mathcal{S} \leq 0.9 \end{cases} \quad (3.3)$$

3.2.4 Sampling to create the VPD

Random realisations of arterial networks are generated by drawing samples from the assumed distributions outlined above, and the pressure and flow-rate waveforms across the network associate with each are computed using the physics based model of pulse wave propagation. The physical system is numerically solved using a discontinuous Galerkin scheme [6]. This scheme is chosen as a pre-existing solver is available that has been successfully validated against a 3D model of blood flow through stenosed arterial vessels [21].

Representation of haemodynamic profiles

The output of the pulse wave propagation model is the pressure and flow-rate at all temporal and spatial locations. These vectors of pressure or flow-rate at any

location can be used directly as measurement inputs to the ML algorithms, *i.e.*:

$$\mathbf{q}(x) = \{q(x, t_0), q(x, t_0 + \Delta t), q(x, t_0 + 2\Delta t), \dots, q(x, t_0 + k\Delta t)\}, \quad (3.4)$$

where:

$$k = \frac{T}{\Delta t}, \quad (3.5)$$

and $\mathbf{q}(x)$ represents the vector describing the pressure or flow-rate profile at spatial location x ; $q(x, t)$ represents the scalar pressure or flow-rate at the spatial and temporal position x and t respectively; t_0 represents the initial temporal position; Δt represents the change in time between measurements; k represents the number of measurements taken; and T represents the cardiac period. Using this representation of pressure and flow-rate profiles is, however, not ideal as an input into an ML classifier. To fully capture the information held within the pressure and flow-rate profiles Δt must be small, and so k large. This results in the subsequent dimensionality of input measurements being provided to ML classifiers being large. Furthermore, as severity of stenosis increases, resulting in additional non-linearities in the model, the time step Δt for a stable solution becomes very small. As pressure and flow-rate profiles are periodic it seems natural to represent the time domain haemodynamic profiles using a FS representation. Using this representation allows the pressure and flow-rate profiles to be described to a high level of completeness in much fewer dimensions.

The assumption that all arterial network parameters are independent and normally distributed is physiologically incorrect. To correct for this assumption, post simulation filters are applied to discard non-physiological patients. This is described next.

3.2.5 Post processing

Through random sampling, there is a chance that VPs are assigned combinations of arterial network parameters that result in physiologically unrealistic pressure and flow-rate profiles. Thus, to remove these VPs from the VPD, a post simulation filter is applied. “Hard filters” are applied to VPs—*i.e.* ranges within which pressure profiles must fall are directly imposed. To create limits for this post simulation filter it is first important to understand normal pressure ranges within the arterial

3.3. AVAILABLE PRESSURE AND FLOW-RATE MEASUREMENTS

system being replicated. It is stated in [192] that for patients who have undergone endovascular repair on an aortic aneurysm the range of systolic and diastolic pressures can be expected to be between 126mmHg–199mmHg and 60mmHg–95mmHg respectively. To allow the full range of possible pressure waveforms present in a real population to be expressed in the VPD these ranges are made more liberal. The three conditions of the post simulation filter are:

$$\text{Filters: } \begin{cases} \max(\mathbf{P}_{\text{inlet}}) < 225\text{mmHg} \\ \min(\mathbf{P}_{\text{inlet}}) > 25\text{mmHg} \\ \max(\mathbf{P}_{\text{inlet}}) - \min(\mathbf{P}_{\text{inlet}}) < 120\text{mmHg} \end{cases}$$

where $\mathbf{P}_{\text{inlet}}$ represents the vector describing the time domain pressure profile at the inlet of the system. Using the VPD created through the methodology described above, the ability of ML classifiers to distinguish between healthy and unhealthy VPs is assessed.

3.3 Available pressure and flow-rate measurements

To use the VPD outlined above to assess the accuracy of ML classifiers, limitations are imposed on the available pressure and flow-rate measurement. While pressure and flow-rate profiles can be obtained at any location within the arterial network, measurements are limited to the inlet and two outlets of the system, shown in Figure 3.1 by P_1 , Q_1 , P_2 , Q_2 , P_3 , and Q_3 respectively. These measurements should be sufficient to study the affect of measurement location and type on classification accuracy, while maintaining a feasible maximum input dimensionality into ML classifiers. How ML classifiers are used to predict the health of VPs from these haemodynamic measurements is outlined next.

3.4 Machine learning set up

A model mapping a vector of input features, \mathbf{y} , to a discrete output classification, z , can be described as:

$$z = m(\mathbf{y}) \quad z \in \mathcal{C}, \quad (3.6)$$

with

$$\mathcal{C} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}, \dots, \mathcal{C}^{(j)}\}, \quad (3.7)$$

where \mathcal{C} represents the set describing all possible classifications, and $\mathcal{C}^{(j)}$ represents the j^{th} possible classification. The input features, \mathbf{y} , represents a vector describing a user defined combination of the haemodynamic measurements [$Q_1, Q_2, Q_3, P_1, P_2, P_3$] in FS form, and the output classification represent the corresponding health of the VP. The methodologies used to both create models described by Equation (3.6) and then subsequently assess their performances are explained next.

3.4.1 Test/train split

The VPD is split into two parts: training set and testing set. The training set is used for learning in the ML algorithms and is set to two-thirds of the size of the VPD. The remaining one-third of the VPs comprise the test set, which is used to assess the accuracy of the ML algorithms on previously unseen data, *i.e.* the data not used while training.

3.4.2 Standardisation of input data

As the FS coefficients describing the haemodynamic profiles vary by several orders of magnitude, it is important to standardise this data before it is provided to ML classifiers. Without this standardisation, it is likely that the information imparted by different coefficients will be weighted according to their order of magnitude, restricting the maximum information that can be extracted. In this study the data set is transformed to standard score form (referred to as Z-score standardisation [136]), based on the statistics of the training set. The raw VPD is described by an $m \times d$ matrix:

$$\mathbf{Y} = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \\ y_{m1} & & y_{mn} \end{bmatrix},$$

where each row describes a different VP—and so m represents the total number of VPs within the VPD—and each column describes a different FS coefficient—and so n represents the total number of FS coefficients and must be a multiplication of 11. The standard score form of the j^{th} coefficient corresponding to the i^{th} VP, $\mathcal{Z}(y_{ij})$, is equal to:

$$\mathcal{Z}(y_{ij}) = \frac{y_{ij} - \bar{y}_j^{\text{train}}}{\sigma_j^{\text{train}}}, \quad (3.8)$$

where \bar{y}_j^{train} and σ_j^{train} represent the mean and standard deviation of the j^{th} coefficient across the training set, described in Section 3.4.1. Equation 3.6 can be written to include this standardisation as:

$$z_i = \tilde{m}(\mathcal{Z}(\mathbf{y}_i)), \quad (3.9)$$

where \mathbf{y}_i represents a singular row (corresponding to a singular VP) of matrix \mathbf{Y} , and \tilde{m} represents the standardised version of model m .

3.4.3 Machine learning algorithms

Within Chapter 2 six different ML classification methods are outlined. Here ML classifiers are created using four of these different methods. These are logistic regression (LR), support vector machine (SVM), naive Bayes' (NB), and random forest (RF). Two characteristics that can be used to distinguish between different ML methods are if they are capable of producing linear or non-linear partitions between different classifications, and if they return a probabilistic or non-probabilistic output prediction. These four ML methods are chosen as they encompass all four combinations of classifier characteristic behaviours, as shown in Table 3.2. Another attractive feature of these methods is that they all require very little problem specific optimisation. Before ML classifiers are trained and tested using each of the four different methods, preliminary tests are carried out using an LR method. An LR method is used for these initial tests as it is computationally inexpensive. Once

	Capable of linear partitions	Capable of non-linear partitions
Probabilistic	LR	NB
Non-probabilistic	SVM with linear kernel	SVM with radial basis function kernel RF

Table 3.2: The four major classifier behaviour characteristics, and how each classification method aligns with these characteristics.

an initial understanding of the VPD has been gained further classifiers are trained using the other three ML classification methods.

3.4.4 Required size of the VPD

An important consideration in the creation of the VPD is its size – how many VPs are sufficient for the ML algorithms to be applied? Here *a priori* evaluation of the required size of the VPD is presented, while *a posteriori* analysis is found in Section 3.5.1. A common rule of thumb within ML is that to train a classifier at least 10 examples of each possible classification are required per input dimension, known as events per variable or EPV [211]. To be able to determine how many VPs are required, the maximum number of input dimensions provided to an ML classifier must be understood. Obtainment of pressure and flow-rate measurements is restricted to the inlet and two outlets of the system (as outlined in Section 3.3). Thus, the maximum number of input dimensions into ML classifiers is 66 (each measurement is described by 11 FS coefficients and all six measurements taken).

An estimate to the number of VPs required within the VPD is calculated by assuming the minimum EPV of any one health classification must be 12, to be on the conservative side of the rule of 10. It is chosen that two thirds of VPs within the VPD are used for training classifiers, and the remaining one third used for testing. From this, it is calculated that the VPD requires 1,188 ($3/2 \times 12 \times 66$) VPs with disease in each of the three vessels. Since a balanced data set is desired, the number of healthy patients required is 3,564 (1118×3). This results in the EPV of healthy patients being 36. With repeated sampling of arterial networks and application of post simulation filters, the VPD of desired size is created.

3.4.5 Classifier configurations

To use the VPD to train and test ML classifiers, precise classification objects are outlined. Classification configurations are then constructed to allow for the generalised ML classification methods outlined in Chapter 2, to be used to create ML classifiers capable of meeting these objectives. The objectives and configurations of classifiers can be split into two general categories. These two categories are binary classifiers and multiclass classifiers. Binary ML classifiers are trained to predict the outcome of Equation (3.9) when the output classification may belong to one of two possible outcomes, *i.e.* $\mathcal{C} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}\}$. In contrast, when more than two classes are present, multiclass classifiers are necessary.

Binary configurations

Binary classifiers are created using one of two different configurations.

- **Entire network binary configuration**

The first configuration of binary classifiers used are entire network binary classifiers (ENBCs). The purpose of ENBCs is to predict the health of a VPs entire arterial network, *i.e.* irrespective of the vessel in which the disease is located. When creating ENBCs VPs with no disease present within their arterial network are assigned to the first discrete output classification, $\mathcal{C}^{(1)}$, while all other VPs are assigned to the second discrete output classification, $\mathcal{C}^{(2)}$. The assignment of true state classifications to VPs when creating ENBCs is described by:

$$z_i = \begin{cases} \mathcal{C}^{(1)} & \text{if no disease is present,} \\ \mathcal{C}^{(2)} & \text{else,} \end{cases} \quad (3.10)$$

where z_i represents the true state classification of the i^{th} VP.

- **Individual vessel binary configuration**

The second configuration of binary classifiers are individual vessel binary classifiers (IVBCs). The purpose of IVBCs are to predict if there is a stenosis present within a particular vessel of a VP's arterial network. When creating IVBCs an arterial vessel of interest must be isolated, and VPs with

disease present within this vessel are assigned to the first discrete output classification, $\mathcal{C}^{(1)}$. All other VPs are assigned to the second discrete output classification, $\mathcal{C}^{(2)}$. The assignment of true state classifications to VPs when creating IVBCs is described by:

$$z_i = \begin{cases} \mathcal{C}^{(1)} & \text{if disease is present within} \\ & \text{vessel } \mathcal{V} \\ \mathcal{C}^{(2)} & \text{else,} \end{cases} \quad (3.11)$$

where \mathcal{V} represents the arterial vessel for which the binary health is being predicted. It is chosen to include all VPs that do not have disease in vessel \mathcal{V} within classification $\mathcal{C}^{(2)}$ —rather than just healthy VPs—as this allows for a more complete analysis of the haemodynamic differences between different disease states. Multiclass ML classifiers are discussed next.

Multiclass ML configurations

Multiclass classifiers predict the outcome of Equation (3.9) when the output may belong to more than two different classifications. The purpose of multiclass classifiers is to predict if there is a stenosis present within a VP's arterial network, and if so which vessel does that disease occur within. Thus, four different classifications exist:

$$\mathcal{C} = \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}, \mathcal{C}^{(3)}, \mathcal{C}^{(4)}\}, \quad (3.12)$$

where $\mathcal{C}^{(1)}$, $\mathcal{C}^{(2)}$, $\mathcal{C}^{(3)}$, and $\mathcal{C}^{(4)}$ represents no disease present, disease present within the aorta, disease present within the first iliac, and disease present within the second iliac respectively. It is found through analysis of binary classification behaviours (Section 3.5.2) that LR and SVM classifiers consistently achieve higher accuracy classification than NB and RF classifiers. Thus, multiclass classifiers are only created using these two methods. However, LR and SVM methods are both inherently binary—only naturally capable of distinguishing between two classes. In order to be used as multiclass classifiers they can be adopted through strategies such as one-versus-all [170] and one-versus-one (OVO) [174]. These are described next.

- **One-versus-all configuration**

An OVA strategy [170] trains multiple instances of binary classifiers, each designed to predict the probability of a separate classification problem. These probabilities are then combined to make a multiclass prediction.

In the case of this PoC, the OVA classifier trains four instances of a binary classifier. Each binary classifier prescribes a correct binary health classification of 1 to all VPs belonging to the corresponding possible classification. All other patients, irrespective of which of the other three classifications they belong too, are assigned a correct binary health classification of 0:

$$\psi_i^{(j)} = \begin{cases} 1 & \text{if } z_i = \mathcal{C}^{(j)} \\ 0 & \text{otherwise.} \end{cases}, \quad j \in \{1, 2, 3, 4\}, \quad (3.13)$$

where $\psi_i^{(j)}$ represents the correct binary health classification of the i^{th} VP for the j^{th} instance of a binary classifier. To assign a predicted multiclass classification to a new subject, the predicted probability of producing a positive binary health classification is found for all the four binary classifiers. The classification that corresponds to the highest predicted probability is then selected as the multiclass prediction, *i.e.*:

$$\hat{z}_i = \begin{cases} \mathcal{C}^{(1)} & \text{if } p\left(\psi_i^{(1)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(1)}\right) = \max\left(p\left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)}\right)\right) \\ & \text{for } j \in \{1, 2, 3, 4\} \\ \mathcal{C}^{(2)} & \text{if } p\left(\psi_i^{(2)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(2)}\right) = \max\left(p\left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)}\right)\right) \\ & \text{for } j \in \{1, 2, 3, 4\} \\ \dots & \\ \mathcal{C}^{(4)} & \text{if } p\left(\psi_i^{(4)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(4)}\right) = \max\left(p\left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)}\right)\right) \\ & \text{for } j \in \{1, 2, 3, 4\} \end{cases} \quad (3.14)$$

where \hat{z}_i represents the predicted classification of the i^{th} VP, $p\left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)}\right)$ represents the probability of the i^{th} VP being predicted to have a positive binary health classification for the j^{th} instance of a classifier within the ensemble, \mathbf{y}_i represent the vector of measurements for the test

patient, and $\theta^{(j)}$ represent the measurement specific weightings for the j^{th} classifier.

- **One-versus-one configuration**

An OVO strategy [174] creates binary classifiers for all the pairs of the classes. Thus, if n total classes exist, then $n(n - 1)/2$ binary classifiers are created. The most frequent class predicted among these binary classifiers is then used as the multiclass prediction.

For the PoC problem, the OVO strategy creates six instances of a binary classifier. Each binary classifier is designed to distinguish between two different classes. Thus, the binary classifier created to distinguish between classifications $\mathcal{C}^{(j)}$ and $\mathcal{C}^{(k)}$ uses:

$$\psi_i^{(j,k)} = \begin{cases} 1 & \text{if } z_i = \mathcal{C}^{(j)} \\ 0 & \text{if } z_i = \mathcal{C}^{(k)}. \end{cases}, j, k \in \{1, 2, 3, 4\}, j \neq k \quad (3.15)$$

When predicting the classification of an unseen test VP, a voting scheme is applied. The input measurements taken from the test VPs are passed through each of the six instances of a binary classifier, and the predicted classifications recorded. The classification that occurs most frequently is selected as the multiclass prediction.

It is found that while both LR classifiers employing an OVA method and SVM classifiers employing an OVO method achieve high aortic, first iliac, and second iliac classification accuracy, they produce very low healthy VP classification accuracy (see Section 3.5.4). To rectify the low healthy VP classification accuracies a custom probabilistic configuration (CPC) is developed, as described next.

- **Custom probabilistic configuration**

In the case of both OVA and OVO methodologies, no initial assumption is made about the classification of VPs before the obtainment of predictions from the different instances of binary classifiers within the ensemble. In the case of probabilistic binary ML methods, *i.e.* LR, an alternative CPC is to assign all VPs a health classification corresponding to ‘no disease’ before running

any binary classifiers. This strategy treats ‘no disease’ as the opposite to the three other possible classifications a VP may belong to. The binary classifiers employed in the CPC are identical to the OVA, with the exception that the classifier for ‘no disease’ is omitted. Thus, as opposed to four binary classifiers in the OVA strategy, this strategy uses only three binary classifiers—each pertaining to diseased aorta, first iliac, and second iliac respectively. The assignment of true state binary health classifications to VPs for the three binary classifiers are:

$$\psi_i^{(j)} = \begin{cases} 1 & \text{if } z_i = \mathcal{C}^{(j)} \\ 0 & \text{otherwise.} \end{cases}, \quad j \in \{2, 3, 4\}. \quad (3.16)$$

Note that $j = 1$ for ‘no disease’ classification is not included. To predict a multiclass classification for test VPs, the vessel that produces the highest probability of being diseased among the three binary classifiers is first found. The default multiclass classification is ‘no disease’ unless the highest probability of disease occurring is greater than a prescribed threshold (decision boundary), in which case the test VP is predicted to have disease in the arterial vessel with this highest probability, *i.e.*

$$\hat{z}_i = \begin{cases} \mathcal{C}^{(1)} & \text{if } \max \left(\text{p} \left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)} \right) \right) < \mathcal{B} \quad \text{for } j \in \{2, 3, 4\}, \\ \mathcal{C}^{(2)} & \text{if } \text{p} \left(\psi_i^{(2)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(2)} \right) = \max \left(\text{p} \left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)} \right) \right) \\ & \text{for } j \in \{2, 3, 4\} \\ & \text{and } \text{p} \left(\psi_i^{(2)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(2)} \right) \geq \mathcal{B}, \\ \dots & \\ \mathcal{C}^{(4)} & \text{if } \text{p} \left(\psi_i^{(4)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(4)} \right) = \max \left(\text{p} \left(\psi_i^{(j)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(j)} \right) \right) \\ & \text{for } j \in \{2, 3, 4\} \\ & \text{and } \text{p} \left(\psi_i^{(4)} = 1 \mid \mathbf{y}_i, \boldsymbol{\theta}^{(4)} \right) \geq \mathcal{B}, \end{cases} \quad (3.17)$$

where \mathcal{B} represents the threshold (decision boundary).

As opposed to the classical OVA, where the classification with highest predicted probability, irrespective of the magnitude of this probability, is

	VP belongs to classification $\mathcal{C}^{(j)}$	VP does not belongs to classification $\mathcal{C}^{(j)}$
VP predicted to belong to classification $\mathcal{C}^{(j)}$	$\mathcal{C}^{(j)}$ True positive (TP)	$\mathcal{C}^{(1)}$ False positive (FP)
VP predicted to not belong to classification $\mathcal{C}^{(j)}$	$\mathcal{C}^{(j)}$ False negative (FN)	$\mathcal{C}^{(j)}$ True negative (TN)

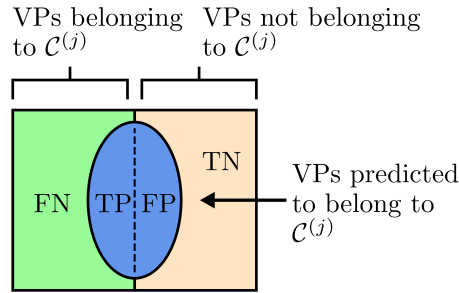
Table 3.3: The process of determining whether a VP is classified correctly or incorrectly, for a specific discrete classification $\mathcal{C}^{(j)}$, is outlined above.

chosen, CPC requires a minimum certainty of disease being present to be met before the default hypothesis ‘no disease’ can be overridden. It is not possible to create multiclass classifiers in this manner using non-probabilistic methods, such as SVM.

3.4.6 Quantification of results

Two different methods are used to quantify and compare the results achieved by different classifiers. The first, and most intuitive, of these is to compute the sensitivity and specificity of classification across the test set. Determination of whether a VP is classified correctly or incorrectly can be achieved by comparison against the true state, see Table 3.3. The proportion of VPs belonging to a classification that are correctly classified, *i.e.* the sensitivity (S_e), is computed using the equation $S_e = TP / (TP + FN)$, while the proportion of VPs not belonging to a classification that are correctly classified, *i.e.* the specificity (S_p), is compute using the equation $S_p = TN / (TN + FP)$. The relationships between the TP, FN, FP, TN, S_e , and S_p with respect to the class $\mathcal{C}^{(j)}$ are shown within Figure 3.4.

In the case of multiclass classifiers, assessment of the accuracy of classification requires provision of the sensitivity and specificity corresponding to each discrete classification. Multiclass classifiers are created to predict which one of four different classifications VPs belong to, thus the full accuracy of multiclass classifiers requires eight different numbers (four sensitivities and four specificities). While quantifying the accuracy of ML classifiers through the sensitivity and specificity of each classification is simple and easily understood, the description of results through two different numbers per classification can make comparison of different classifiers



$$\text{Recall } (\mathcal{R}) = \text{Sensitivity } (S_e) = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Precision } (\mathcal{P}) = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Specificity } (S_p) = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$F_1 = \frac{2\mathcal{P}\mathcal{R}}{\mathcal{P} + \mathcal{R}}$$

Figure 3.4: The differences between sensitivity, specificity, recall, and precision. Each of these relationships to the TP, FN, FP, and TN classification accuracies is also included.

difficult.

A more complex, however easier to compare, method for quantifying the accuracy of ML classifiers is to use the F score [181]. The F score produces a single quantitative score allowing for easy comparison. Higher values of F score imply a better classification. To calculate the F score the precision (\mathcal{P}) and recall (\mathcal{R}) of each discrete classification are calculated. A visual explanation of F score, precision, and recall is included within Figure 3.4. Precision is the proportion of patients predicted to belong to a classification, who do in fact belong to that classification ($\text{TP}/(\text{TP} + \text{FP})$). The recall is the portion of patients belonging to a classification who are correctly classified, thus identical to the sensitivity ($\text{TP}/(\text{TP} + \text{FN})$). The difference between precision and recall can be seen in the denominator. The F score combines the precision and recall as follows:

$$F = \frac{(\delta^2 + 1)\mathcal{P}\mathcal{R}}{\delta^2\mathcal{P} + \mathcal{R}}, \quad (3.18)$$

where δ represents a hyper parameter. Values of δ above 1 give preference to recall, while values under 1 give preference to precision. Although there is a preference to recall in the proposed application of the classifiers, $\delta = 1$ is used to get a general

sense of how classifiers perform without any bias. As $\delta = 1$ is being used, the F score is referred to as F_1 score and is the harmonic mean of precision and recall.

While the F_1 score balances the effect of precision and recall, it does not balance the effect of the sensitivity and specificity. Given a situation in which there is an equal number of healthy and unhealthy VPs, an ENBC which correctly predicts 80% of healthy VPs ($\mathcal{R} = S_e = 0.8$) and 20% of unhealthy VPs ($S_p = 0.2$) will achieve an F_1 score of 0.61. An ENBC that correctly predicts the health of 20% of healthy VPs ($\mathcal{R} = S_e = 0.2$) and 80% of unhealthy VPs ($S_p = 0.8$), however, will achieve an F_1 score of 0.28, despite the fact that the total number of VPs who have been correctly classified is unchanged. The importance of using both the F_1 score and the sensitivity and specificity in combination can be seen from the example highlighted above.

3.4.7 ML implementation

It is possible to simplify the process of training and testing different ML classifiers by using available ML packages. These packages are created to allow for quick application of ML methods to different problems. It is important, however, to ensure the use of these pre-made packages does not have a detrimental effect on the results achieved. The deliberate non-intensive nature of these pre-made packages restricts the problem specific optimisation of the classifiers. To examine the affect of using pre-made ML packages, in this case the Scikit-learn package [156], on the accuracy of results achieved a series of LR classifiers are created using both a custom written implementation and the Scikit-learn implementation of the LR method.

The Scikit-learn package offers 5 different solvers that can be used to train an LR algorithm. The custom implementation of an LR method has been created to use one of two different hypothesis functions—sigmoid or tanh—and one of five different update functions—gradient descent, gradient descent with momentum [161], Nesterov accelerated gradient (NAG) [141], Adadelta [223], and Adam [109]. This results in the custom implementation being capable of running ten different LR set ups (each of the two different hypothesis functions being combined with each of the five different update functions). To compare the accuracy of the 15 different unique LR algorithms—the five Scikit-learn and ten custom implementations—each is used to train and test an ENBC, using the measurements of pressure and flow-

rate taken from VPs at all three locations. For each LR set up, five instances of a classifier are created. Each of these instances uses a different random subset of VPs for training and testing the classifier. The average performance of these five instances is then computed, reducing the affect of the VPs selected for training and testing on the results achieved. This is referred to as five fold validation.

The highest accuracy results achieved by the Scikit-learn package produced an average sensitivity for classification $\mathcal{C}^{(1)}$, *i.e.* the percentage of healthy VPs classified correctly, of 0.727, and an average specificity for classification $\mathcal{C}^{(1)}$, *i.e.* the percentage of unhealthy VPs classified correctly, of 0.522 across the five folds of the VPD. These accuracies correspond to an F_1 score of 0.581. The highest accuracy results achieved by the custom implementation used a tanh activation function and a Adam update function. This LR set up produced an average sensitivity for classification $\mathcal{C}^{(1)}$ of 0.729, and an average specificity for classification $\mathcal{C}^{(1)}$ of 0.525 across the five folds. These accuracies correspond to an F_1 score of 0.583, identical to the result achieved by the Scikit-learn package when evaluated to two decimal places. These results suggest using the Scikit-learn package to train future ML classifiers will have little to no affect on the accuracy of classification achieved. Due to the significant increase in ease of implementation, and so consequently decrease in time to implement, the majority of future ML classifiers are created using the Scikit-learn package. The custom LR implementation is only used when there is a clear benefit to doing so.

3.4.8 Hyper-parameter optimisation

The architecture of LR, NB, and SVM classifiers can all be considered to be problem independent. While these three algorithms are able to undergo varying levels of problem specific optimisation, the underlying structure of the classifier cannot be changed. The architecture of RF classifiers, however, is dependent on the specific problem. The architecture choices for the classifiers and associated hyper-parameter optimisation is described next.

LR, SVM, and NB

For LR, the ‘LIBLINEAR’ solver offered by the Scikits-learn [156] package is chosen. In the case of SVM, a kernel is typically chosen to map the input measurements to a

higher order feature space [94]. All SVM classifiers use a radial basis function kernel [184]. In the case of NB, the distribution of input measurements across the data set is chosen to be Normal [139].

Random forest

In the case of RF the number of trees within the ensemble and the maximum depth of each tree can be optimised for a specific problem. To optimise these two hyper-parameters, a grid search is carried out. A grid is constructed by discretising the number of trees within the ensemble between 10 and 400 at intervals of 10, and discretising the maximum depth of each tree between 20 and 200 at intervals of 10. An ENBC is created using an RF method and every permutation of number trees within the ensemble and maximum depth of each tree, over five folds of the VPD, and using all six measurements. The hyper-parameters describing the architecture that produces the highest F_1 score is found, and this combination of hyper-parameters is then chosen for all subsequent classifiers. It is found that the highest F_1 score achieved is recorded for the RF classifier trained and tested using 20 trees with a maximum depth of 60.

It is unlikely that a single architecture will consistently produce the maximum accuracy results achievable when varying the input measurements provided. Due to the exploratory stance of this study—with priority being given to uncovering patterns and behaviours rather than optimisation for maximum accuracy—it is computationally impractical to perform a grid search for every classifier. It is instead assumed that the differences in optimum architectures when using different input measurements is minor—fine-tuning rather than systemic restructuring—and so reasonable representations of achievable accuracies can be produced using a fixed architecture. Thus, all future RF classifiers are trained and tested using the optimum hyperparameters found and outlined above. The likely affect of using a single architecture for all RF classifiers created will be considered within the later chapters of this thesis.

3.5 Results and discussion

The results of the application of ML classifiers to the VPD are presented in the subsequent section. A brief analysis on the statistics of the VPD is found in Appendix A. Hitherto, all FS have been described by their sine-cosine coefficients. There are, however, two forms of a FS. While these two forms of a FS require the same dimensionality to represent a pressure or flow-rate profile, they do not necessarily capture the same information about the haemodynamic profiles as one another. A comparison of the use of these two forms as input measurements into ML classifiers is presented in Appendix B. It is found that using the sine-cosine form produces the highest accuracy classification, and so all profiles are described by their sine-cosine coefficients.

3.5.1 *A posteriori* analysis of the size of the VPD

While an *a priori* estimation of the required size of the VPD has been made by calculating the EPV, this can be checked more thoroughly by training and testing a series of ML classifiers with successively larger numbers of available VPs. To fully evaluate the adequacy of the number of VPs within the VPD, classifiers must be trained and tested under the circumstances with the lowest number of points (*i.e.* the lowest number of VPs) in the highest dimensional space (*i.e.* the most pressure and flow-rate measurements) possible. Thus, this assessment is made for the case with the largest input dimensionality—the pressure and flow-rate measurements at all three available locations. Furthermore, classifiers must be trained and tested with the lowest number of VPs belonging to a single classification—*i.e.* the lowest number of VPs belonging to either $\mathcal{C}^{(1)}$ or $\mathcal{C}^{(2)}$ possible—and thus classifiers must be trained to predict the health of each vessel individual. As the VPD has been created so that there is an equal number of healthy and unhealthy VPs, for any given number of available VPs an ENBC will have half of the number of available VPs belonging to $\mathcal{C}^{(1)}$ and half belonging to $\mathcal{C}^{(2)}$. On the contrary, three series of IVBCs are created (as described in Section 3.4.5), each predicting the health of a different vessel. This results in each instance of an IVBC having 5/6 of the available VPs belonging to a negative binary classification, however only 1/6 of the number of available VPs belonging to a positive binary classification. By empirically showing

there is an adequate number of VPs to train and test classifiers within this extreme situation, it is reasonable to assume there is an adequate number of VPs to train and test ENBCs.

Due to the class imbalance present, *i.e.* there are significantly more VPs belonging to $\mathcal{C}^{(2)}$ than $\mathcal{C}^{(1)}$, a multiplier is applied to the cost of VPs belonging to $\mathcal{C}^{(1)}$ when training IVBCs. Without this multiplier IVBCs are biased towards VPs belonging to $\mathcal{C}^{(2)}$. The weighting w applied to the cost of the prediction of VPs belonging to $\mathcal{C}^{(1)}$ for each classifier is calculated by assigning a ratio ι to the effective number of VPs belonging to classifications $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$:

$$\iota = \frac{w * m^{(1)}}{m^{(2)}}, \quad (3.19)$$

where $m^{(1)}$ and $m^{(2)}$ represent the number of VPs belonging to classes $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$, respectively. The corresponding cost (loss) function of the LR classifier (as described by Equation (2.26) in Chapter 2) is modified to include the weight w as:

$$L(\boldsymbol{\theta}, \mathbf{Y}, \boldsymbol{\psi}) = -\frac{1}{m} \sum_{i=1}^m \left(w\psi_i \log(\text{h}(\mathbf{y}_i, \boldsymbol{\theta})) + (1 - \psi_i) \log(1 - \text{h}(\mathbf{y}_i, \boldsymbol{\theta})) \right), \quad (3.20)$$

where $L(\boldsymbol{\theta}, \mathbf{Y}, \boldsymbol{\psi})$ represents the average cost across the data set; \mathbf{Y} and $\boldsymbol{\psi}$ represents the matrix of input measurements and the vector of the known correct binary classifications taken from all patients within the data set respectively; \mathbf{y}_i and ψ_i represents the vector of input measurements and the known correct binary classification corresponding to patient i respectively; $\boldsymbol{\theta}$ represents the measurement specific weightings; and m represents the number of patients within the data set. When $\iota = 1$ is used, VPs belonging to $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ have the potential to contribute equally to the total cost of prediction across the training set. If $\iota > 1$ is used, bias is given towards VPs belonging to $\mathcal{C}^{(1)}$, and $\iota < 1$ gives bias towards VPs belonging to $\mathcal{C}^{(2)}$. Unless stated otherwise, $\iota = 1$ is used.

For successively increasing number of VPs, three IVBCs corresponding to disease in the three vessels are trained and tested, over five folds of the VPD. The average F_1 scores achieved across the training and test sets with increasing numbers of VPs are shown in Figure 3.5.

Figure 3.5 shows that both training and test accuracies are low when a small

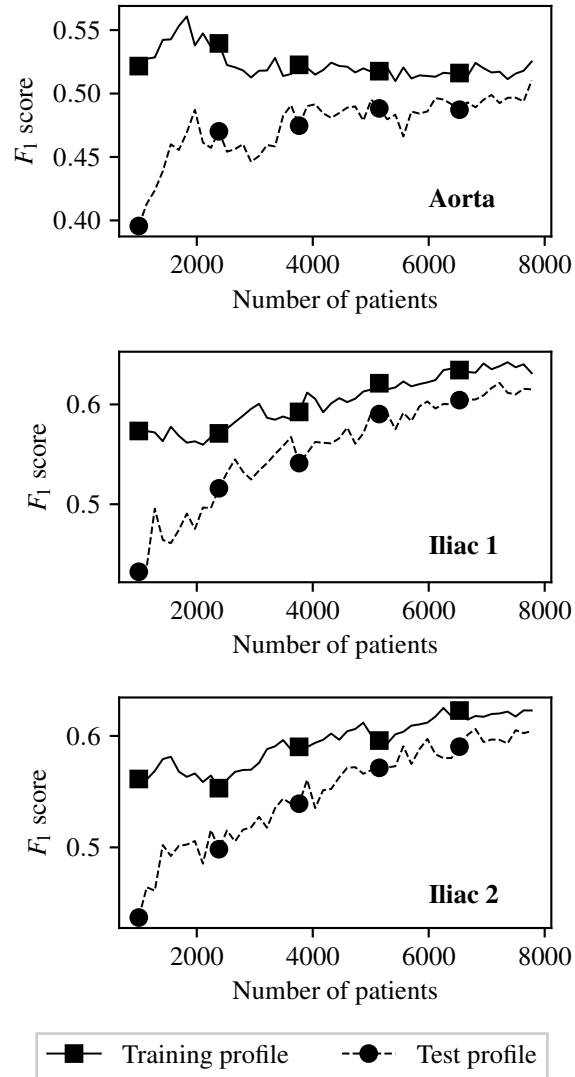


Figure 3.5: The results of the analysis of the adequacy of the size of the VPD, when using pressure and flow-rate measurements at all the three locations, are shown above. The first, second, and third rows show the results of aortic, first iliac, and second iliac classification respectively.

proportion of the VPD is made available to ML classifiers. This suggests that the classifiers being trained are underfitting the training data, *i.e.* low variance but high bias. The classifiers trained can neither fit the training data nor generalise to the test data. As the number of available VPs increases the behaviour of classifiers differs between the aorta and two common iliacs. In the case of the aorta, the training accuracy remains relatively constant, while the test accuracy increases. In the case of the two common iliac classifiers, both the training and test accuracy increase. These behaviours suggest the classifiers are fitting the training data better, and as a consequence are better able to classify test patients. Initially, between 1,000 and 5,000 available VPs, the changes made to the partition between VPs belonging to $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ through the input measurement space are significant, and so there are large jumps in change to the training and test accuracies. As the number of available VPs continuous to increase the partition between healthy and unhealthy patients through the input measurement space begins to converge to an optimum solution. This causes the changes to the training and test accuracies to reduce, and eventually flatten off. Figure 3.5 suggests that beyond 7,000 VPs the VPD contains enough VPs to train and test ML classifiers. This is shown by the fact that the training and test accuracies of each vessel are consistent for the final several numbers of available VPs, and so the partitions between healthy and unhealthy patients are no longer changing.

3.5.2 ENBC results

There are 63 possible combinations of input measurements that can be provided to an ML classifier from the three locations at which pressure and flow-rate are measured. A combination search is performed—for every combination of input measurements, an ENBC is trained and then subsequently tested using each of the four different classification methods. The average F_1 score, sensitivity, and specificity of healthy classification accuracy for each input measurement combination and classification method across five folds of the VPD are recorded. Combinations of interest are then further analysed. The full tables of results are shown in Appendix C. The F_1 score achieved by each ML method and combination of input measurements are visually shown in Figure 3.6.

Like for like input measurement comparison.

To gain a better understanding of what difference in F_1 score can be considered insignificant, classifiers that should theoretically achieve identical accuracies are compared. Exploiting the symmetrical structure of the arterial network (see Figure 3.1), classifiers that use symmetrical measurements can be identified. These are referred to as like-for-like measurements; two examples of such measurements are shown in Figure 3.7. Any discrepancy between the F_1 scores achieved by classifiers trained using like-for-like input measurement combinations is therefore introduced due to training and statistical errors.

There are 24 possible cases of like-for-like input measurement pairs. The discrepancy in the F_1 score achieved by the two classifiers within each of these pairs is computed when using each of the four different classification methods. It is found that NB classifiers show significantly greater magnitudes in the discrepancies of F_1 scores produced than any of the three other methods. The maximum discrepancy in F_1 scores produced when using a NB method is equal to 0.18. This large discrepancy points to something beyond statistical and training errors and is, therefore, most likely related to the unsuitability of the NB method for this problem. It is, therefore, decided to exclude the results achieved by the NB method from all subsequent analysis carried out. The histograms of the discrepancies in the F_1 score between like-for-like input measurement combinations produced when using the remaining three ML methods are shown in Figure 3.8.

Figure 3.8 shows that the discrepancy in F_1 scores between like-for-like input measurement combinations follow a very similar pattern for both the LR and RF classification methods. For both of these methods it can be seen that the majority of the 24 combinations produce a discrepancy in F_1 score of less than 0.005. There is then a clear decay in the number of occurrences as the F_1 score discrepancy increases. 20 of the 24 LR pairs, and 16 of the 24 RF pairs achieved a discrepancy of less than 0.01. When looking at the F_1 discrepancies of SVM classifiers, there appears to be no real decay in the number of occurrences as the F_1 discrepancy increases, and instead a relatively consistent number of SVM pairs produce F_1 discrepancies between 0 and 0.025.

The maximum discrepancy in F_1 scores between like-for-like input measurement combinations is equal to 0.0231. This discrepancy in F_1 score is measured between

two pairs of input measurements when using an SVM method. The firsts of these two pairs is (Q_3, P_3) and (Q_2, P_2) . When training a SVM classifier using (Q_3, P_3) the sensitivity and specificity is equal to 0.71 and 0.51 respectively. When training a SVM classifier using (Q_2, P_2) the sensitivity and specificity is equal to 0.74 and 0.47 respectively. The second pair of input measurements producing a discrepancy in F_1 score of 0.0231 is (Q_3, P_1) and (Q_2, P_1) . When training SVM classifiers using (Q_3, P_1) and (Q_2, P_1) the sensitivities and specificities are equal to 0.76 and 0.50; and 0.8 and 0.46 respectively. It can be seen that in the case of both pairs of input measurements highlighted above, while there are some differences in the sensitivities and specificities produced, the differences in accuracies are relatively low and the behaviours of each of the two classifiers are relatively consistent.

From Figure 3.8 and the aforementioned analysis, a difference in F_1 score of more than 0.01 between two LR, SVM, or RF classifiers trained using different input measurements can be considered to significant and likely due to the behaviour of the classifiers. It is important to remember, however, that a difference in F_1 score of approximately 0.025 is required to fully rule out the possibility that patterns are due to training or statistical errors.

Effect of number of input measurements

Appendix C and Figure 3.6 show that there is a correlation between the number of input measurements used to train and test ML classifiers and the F_1 score achieved. To investigate this further the average F_1 score achieved by of all classifiers trained using one to six input measurements is found for each of the three different classification methods. The maximum and minimum F_1 score are also recorded and shown in Figure 3.9. It can be seen that as the number of input measurements increases, the average F_1 score achieved by all classification methods also increases. The increase in F_1 score is most noticeable for the SVM method. For the LR and RF classification methods, the average F_1 score achieved when using 1 input measurement is approximately 0.5, representing naive classification ($S_e + S_p = 1$). The average F_1 score achieved by SVM classifiers trained using 1 input measurement is marginally better than naive classification. This finding that the average F_1 score increases as the number of input measurements increases is expected as the discriminatory information increases, on average, as more measurements are made

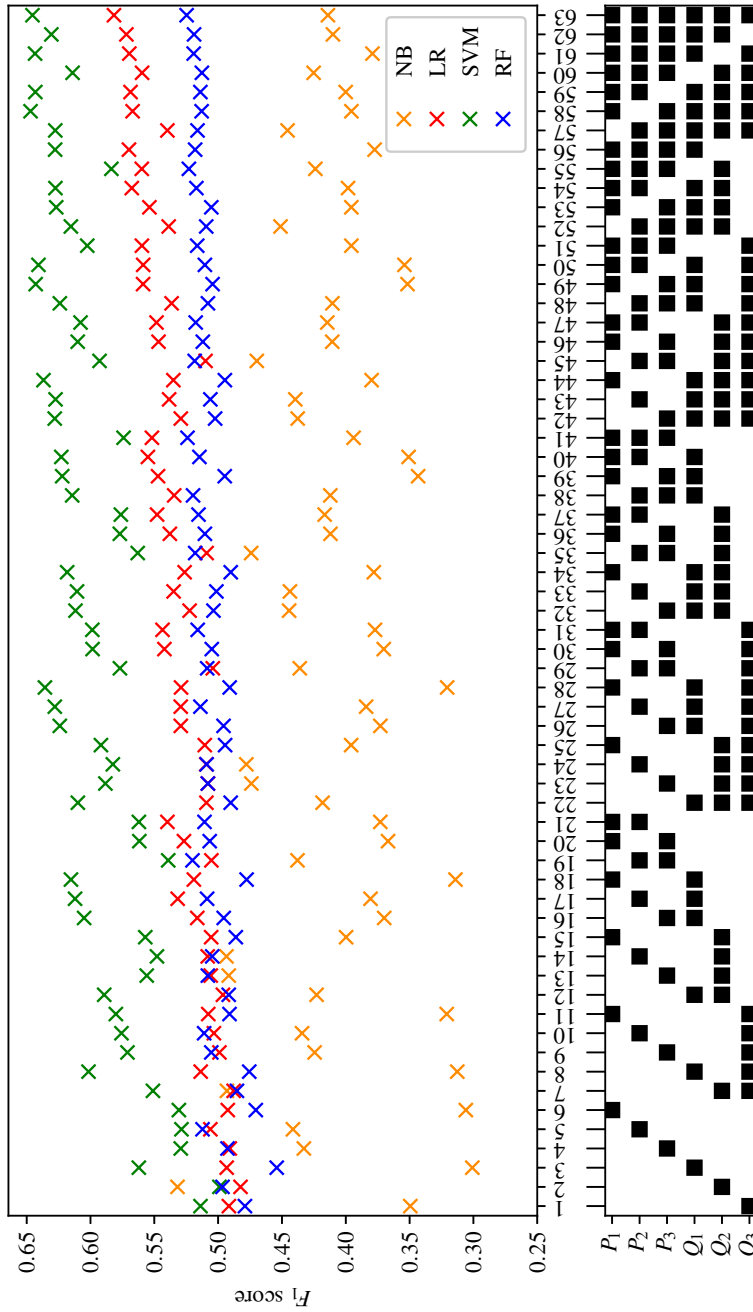


Figure 3.6: The F_1 scores achieved by the ENBCs employing the NB, LR, SVM, and RF methods when providing each combination of input measurements are shown within the upper subplot. The lower subplot indicated which discrete measurements are present within a combination. Measurements included within each combination are highlighted with a black square.

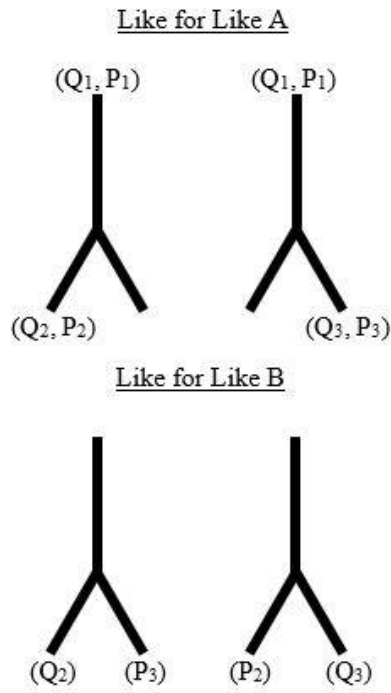


Figure 3.7: Two examples of like for like input measurements that should result in identical classification accuracies.

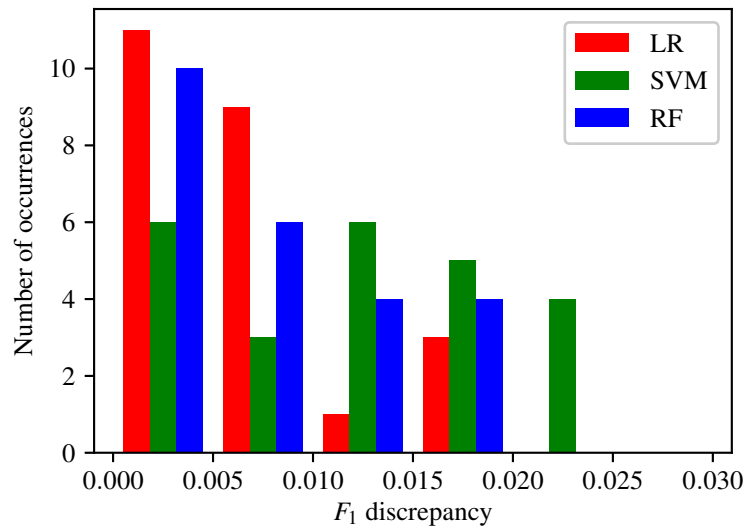


Figure 3.8: Histograms of the discrepancy between the F_1 score of like-for-like classifiers, when employing the LR, SVM, and RF classification methods.

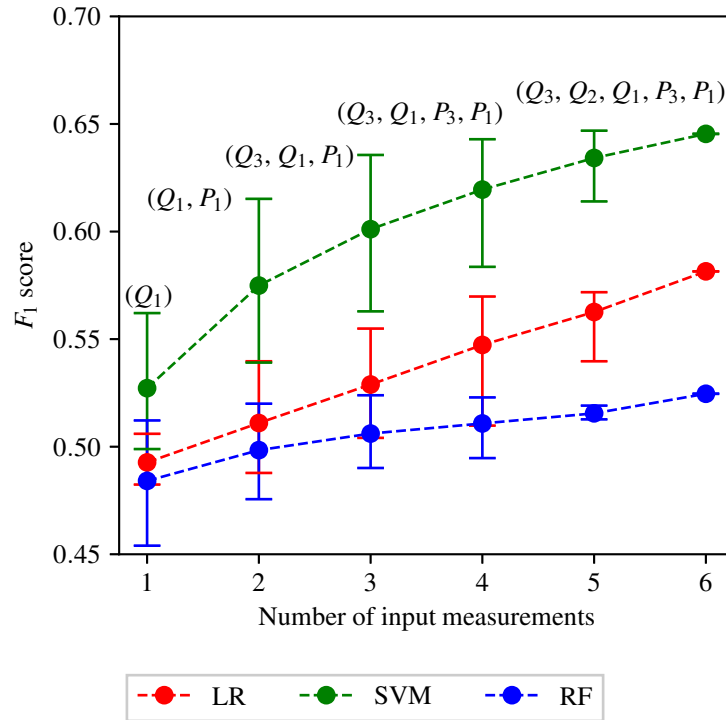


Figure 3.9: The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits. The combination of input measurements that produces the highest F_1 score is identified for the SVM method when using each number of input measurements, up to five input measurements.

available.

Observing the range of maximum to minimum F_1 scores in Figure 3.9 it is seen that as the number of input measurements increases, the range of F_1 scores decreases. An interesting pattern to note is that while the average and minimum F_1 score achieved increases when increasing the number of input measurements between four and six, the maximum remains relatively constant. The maximum and minimum F_1 scores when using four to six measurements are shown in Table 3.4, along with the corresponding sensitivities and specificities. Table 3.4 shows that the maximum accuracy of classification—assessed by the F_1 score, sensitivities, and specificities—vary insignificantly between four, five, and six measurements. Thus, the analysis shows that similar levels of accuracies can be achieved by using four measurements

compared to the case when all six measurements are used, if these measurements are chosen judiciously.

Importance of inlet pressure and flow-split

A further pattern noticed within the tables in Appendix C is that classifiers trained using P_1 generally perform better than those that do not use P_1 . To analyse this further, the F_1 score of classifiers trained with and without P_1 are separated and plotted in Figure 3.10. For LR and SVM classifiers, a clear improvement of $\Delta F_1 \approx 0.05$ is observed when P_1 is included. This behaviour is expected, in part due to design. There are a total of 32 combinations of input measurements that include the use of P_1 , and 31 combinations of input measurements that exclude the use of P_1 . The classifier trained using all six input measurements, and five of the six trained using five input measurements contain P_1 . Only one classifier trained using five input measurements does not include P_1 . It has previously been shown in Figure 3.9 that, generally, classifiers trained using more input measurements achieve higher accuracy classification results. There is, therefore, some expected skewing towards higher F_1 scores in favour of classifiers trained with P_1 . This expected behaviour is further amplified by the fact that only one combination of input measurements consists of a single input measurement and contains P_1 . This compares to five combinations that consist of a single input measurement and exclude P_1 . This results in an expectation of more low scoring classifiers without P_1 .

Figure 3.10 shows that in the case of LR, only 11 of the 32 classifiers trained using P_1 achieve an F_1 score of less than 0.54. This compares to all 31 LR classifiers trained without P_1 achieving an F_1 score of less than 0.54. In the case of SVM classifiers, only 1 combination of input measurements containing P_1 achieves an F_1 score of less than 0.54. This compares to 5 combinations of input measurements that do not contain P_1 that achieved an F_1 score of less than 0.54. When the threshold for comparison is increased to 0.6 it is found that 20 of the 32 SVM classifiers trained with P_1 exceed this threshold, compared to 14 of the 31 trained without P_1 exceeding this threshold.

Similar analysis is performed for the inclusion and exclusion of the measurement Q_1 , shown in Figure 3.11. Figure 3.11 shows a high degree of consistency with Figure 3.10 and the aforementioned analysis. Thus, measurements of pressure and flow-rate

Number of input measurements	Importance	Combination	F_1 score	Sensitivity	Specificity
4	Maximum	(Q_3, Q_1, P_3, P_1)	0.6429	0.7994	0.5688
	Minimum	(Q_2, P_3, P_2, P_1)	0.5836	0.8059	0.4920
5	Maximum	$(Q_3, Q_2, Q_1, P_3, P_1)$	0.6469	0.8115	0.5683
	Minimum	$(Q_3, Q_2, P_3, P_2, P_1)$	0.6140	0.7947	0.5340
6	-	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.6454	0.8050	0.5694

Table 3.4: The combinations of input measurements that produce the maximum and minimum F_1 scores when providing four, five, and six input measurements to ENBCs are shown in the table above. The corresponding sensitivities and specificities are also included.

at the inlet of the system appear to be particularly informative in differentiating between healthy and unhealthy patients.

Another observation made by observing the highest scoring SVM classifiers in Figure 3.9 is that the best performing classifiers include P_1 , Q_1 , and a secondary flow-rate to allow for determination of the flow-split between the left and the right iliacs. For example, when three measurements are used, the best combination is (Q_3, Q_1, P_1) , which would enable the flow split to be known through mass conservation (noting that compliance of the arteries is relatively small) in addition to P_1 . This observation bears similarity to the classical inverse problem analysis presented in [148], where the authors show that in order to find the parameters of any arterial network, the inlet pressure and flow-splits to all the outlets should be known.

Linear versus non-linear partitions

Comparing the results achieved by LR and SVM classifiers in all previous analysis, it can be seen that SVM classifiers consistently achieve higher accuracy results than LR classifiers. When using all six input measurements the LR and SVM classifiers achieve sensitivities and specificities of 0.73 and 0.52; and 0.80 and 0.57, respectively. Similarly, the F_1 scores for LR and SVM classifiers are 0.58 and 0.65, respectively. All SVM classifiers trained up to this point have mapped the input measurements provided to a higher order feature space through the use of radial basis function. The fact that the accuracy of SVM classifiers are consistently higher than LR classifiers suggests that the partition between healthy and unhealthy VPs through the pressure and flow-rate measurement space is non-linear. To test the hypothesis that the increase in accuracy seen in SVM classifiers is due to this higher order mapping, an SVM classifier is trained and tested with a linear kernel. If the increased accuracy achieved by previous SVM classifiers is in fact due to the need for a higher order partition through the input space, the SVM classifier trained and tested using the linear kernel will perform comparably to the LR classifiers. If, however, the increase in accuracy seen in SVM classifiers, when compared to LR classifiers, is unrelated to the higher order partition between healthy and unhealthy patients, the SVM classifier trained and tested using the linear kernel will perform comparably to the previous SVM classifiers.

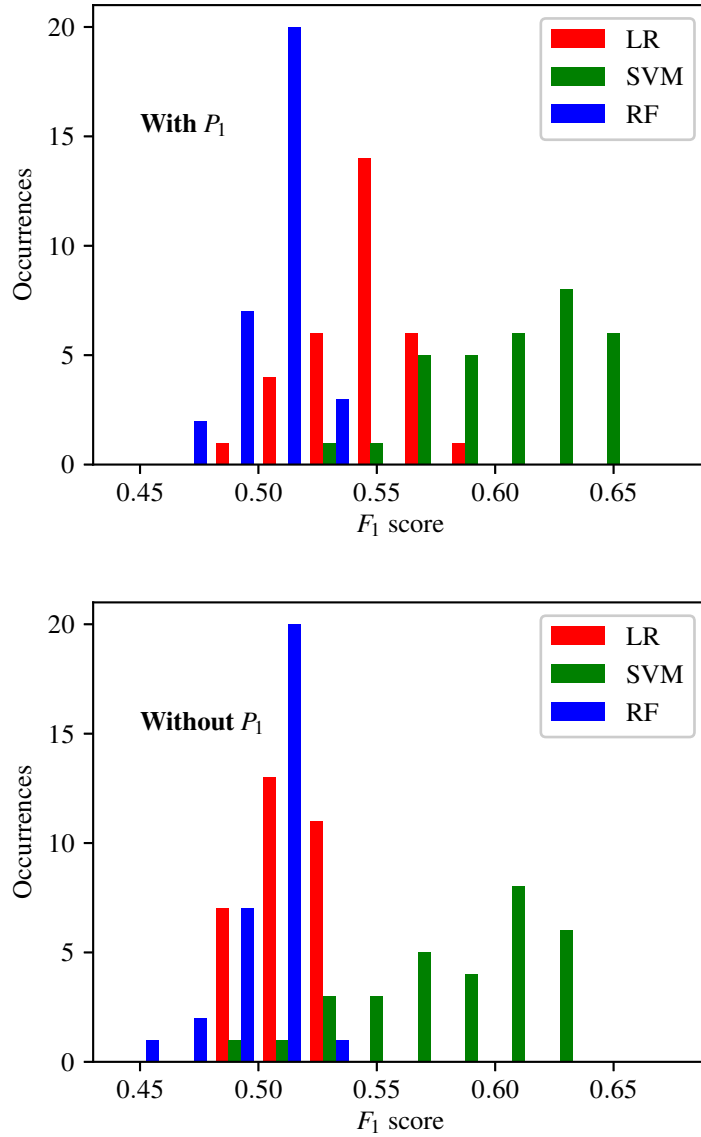


Figure 3.10: The histograms of the F_1 scores achieved by each of the three different classification methods are shown for all input measurement combinations that include P_1 in the upper plot, and exclude P_1 in the lower plot.

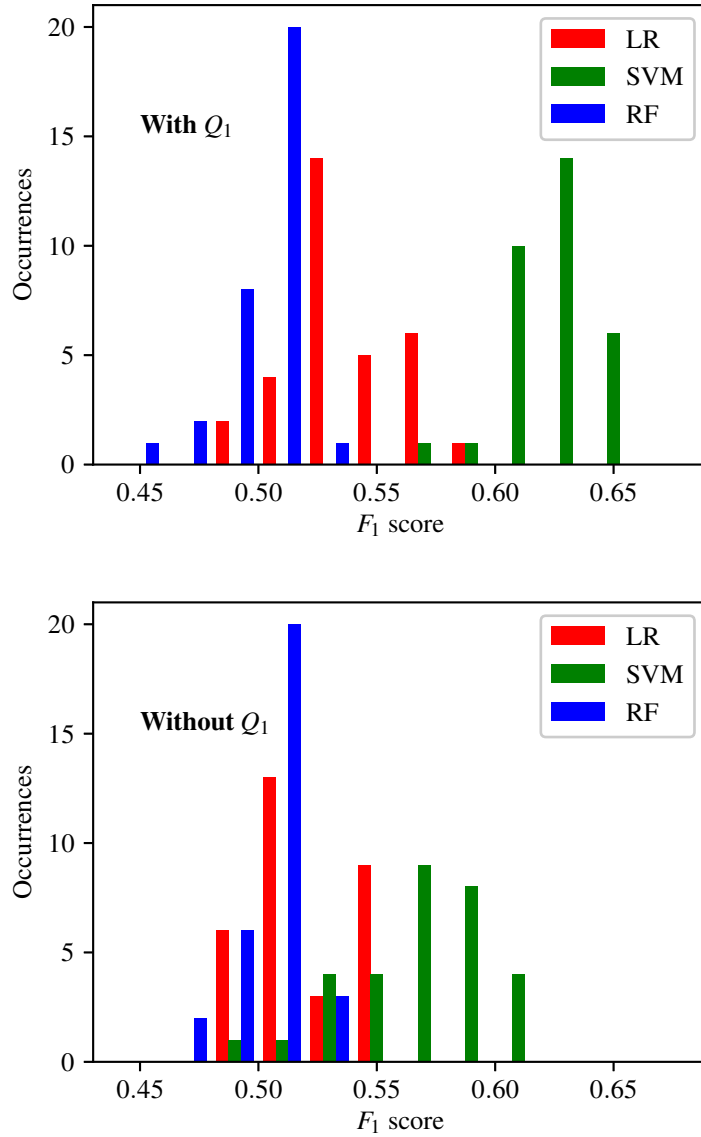


Figure 3.11: The histograms of the F_1 scores achieved by each of the three different classification methods are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.

It is found that an SVM classifier trained using all six pressure and flow-rate measurements and a linear kernel produces an average sensitivity and specificity of 0.85 and 0.42 respectively over five folds of the VPD. This corresponds to an F_1 score of 0.53. The corresponding F_1 scores achieved for LR and radial basis function SVM are 0.58 and 0.65, respectively. The non-linear SVM outperforms the linear SVM and LR (also linear), thus demonstrating that a non-linear mapping is beneficial in discerning between healthy and diseased states.

Effect of disease severity

Here the effect of disease severity on the accuracy of classification is investigated. This analysis is performed using an SVM classifier employing an RBF kernel and an LR classifier, both using pressure and flow-rate measurements at all the three locations.

A scatter plot of the predicted probability returned by the LR classifier against the severity of disease (*i.e.* diseased VPs) for False Negatives and True Positives is shown in Figure 3.12 (left), while a histogram of the predicted probabilities for all the healthy VPs (*i.e.* zero severity) is shown in Figure 3.12 (right). Contrary to intuitive reasoning, which suggests that higher severity of stenosis should be easier to detect, no trends are observed in Figure 3.12, with classification accuracy being independent of the severity. This suggests that the variability in the pressure and flow-rate waveforms induced by the boundary conditions representing physiology before and after the anatomical network (see Figure 3.1) is large and can overshadow the variability induced by stenosis severity alone.

Since the SVM classifiers do not predict a probability of disease, but instead a direct classification of the health of a subject, histograms of the distributions of True Positives and False Negatives across the range of severities for diseased subjects are considered to assess the effect of stenosis severity. These are shown in Figure 3.13. For the healthy subjects, 224 False Positives and 1044 True Negatives are recorded. Similarly to the LR results above, the SVM results do not show a strong trend of severity affecting classifier performance.

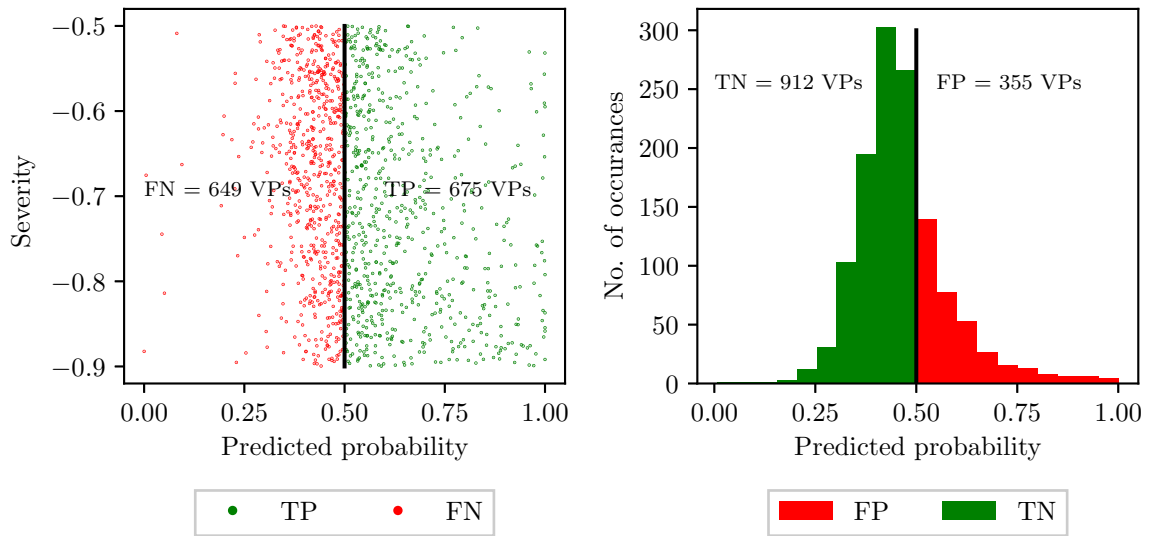


Figure 3.12: Logistic Regression: predicted probability of disease against stenosis severity for diseased patients (left) and histogram of predicted probability of disease for healthy subjects (right). TP: True Positive; FN: False Negative; FP: False positive; and TN: True Negative.

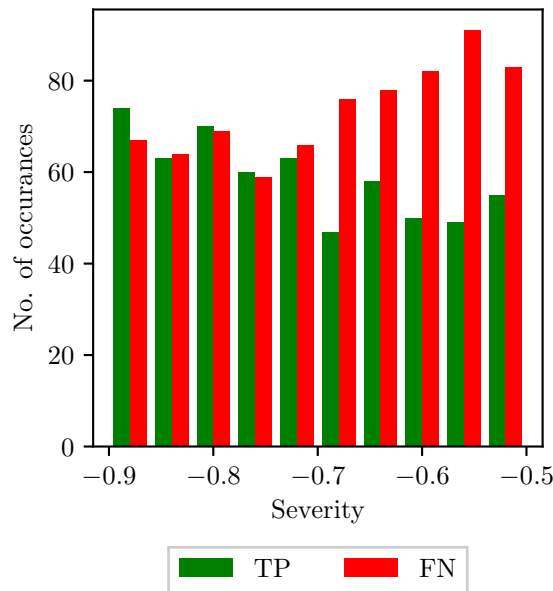


Figure 3.13: SVM with an RBF kernel: histograms of True Positives (TP) and False Negatives (FN) against stenosis severity.

3.5.3 IVBC results

Following an identical procedure to that employed for the ENBC combination search, three IVBC combination searches—one for disease classification in each of the three vessels—are performed using the LR and SVM methods. It is chosen to limit the IVBC combination searches to these two classification methods due to the higher computational expense, and the fact that these two methods have shown consistently higher accuracy results. The full tables of results for the IVBC combination searches are presented in Appendix D. The average, minimum, and maximum F_1 score achieved when using one to six input measurements are shown in Figure 3.14. There is a good agreement between the overall behaviour seen across the IVBC and ENBC (as shown in Figure 3.9) combination searches. These similarities include:

- The average and minimum F_1 score achieved continuously increases when increasing the number of input measurements from one to six.
- The maximum F_1 score initially increases rapidly and reaches an asymptotic limit between two and four input measurements.
- The SVM method consistently produces higher accuracy results than the LR method.

For the SVM configurations corresponding to maximum F_1 scores, the sensitivities, specificities, and the combination of measurements is shown in Table 3.5. Table 3.5 shows that the combinations of input measurements that produce the highest F_1 scores in the two common iliacs are not only identical, but also symmetrical (with the same input measurements being taken from the right and left sides). While the combinations of input measurements that produce the highest F_1 scores differ from the ENBC results (see Table 3.4), a similarity between the two is that the best performing classifiers include a pressure measurement and a combination to determine the flow-split. In Section 3.5.2 it is hypothesised that the combination of pressure at the inlet and flow-split may be particularly informative. Table 3.5, however, seems to suggest that it may be the pressure within the diseased vessel and the flow-split that best captures the presence of a stenosis.

Comparing Tables 3.4 and 3.5 also shows that IVBCs, owing to their more granular characterisation of diseases states, lead to higher F_1 scores, sensitivities,

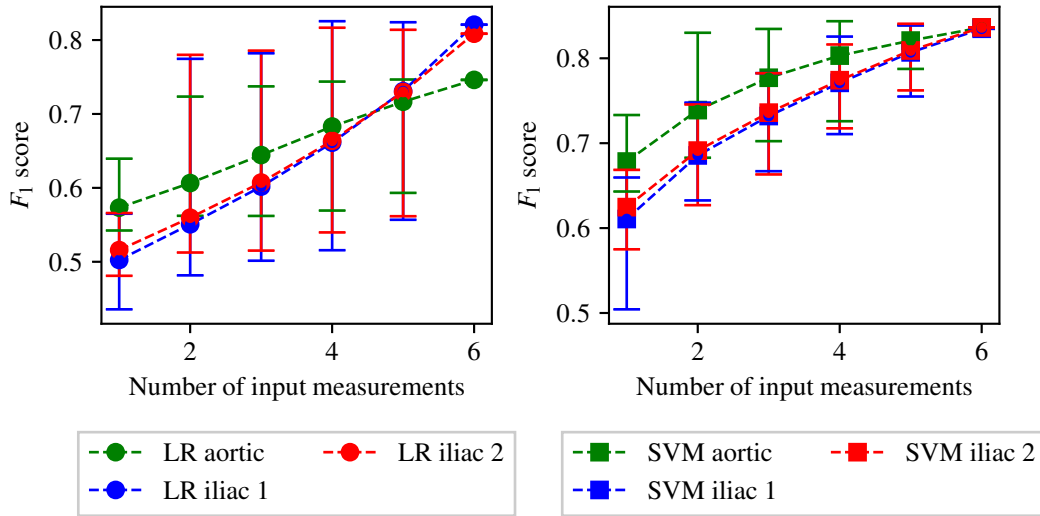


Figure 3.14: The average, maximum, and minimum F_1 score achieved by all IVBCs when providing different numbers of input measurements to detect disease in each of the three vessels are shown. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.

and specificities, relative to the ENBCs. Neither of them are, however, good at pointing to the precise vessel that is diseased in the network. Note that even if an IVBC classifier has perfect accuracy it does not lead to knowledge of the precise diseased vessel; for example, the aortic IVBC classifier only determines whether disease is in the aorta, and considers both healthy and diseased iliac vessel patients together in one class (see Section 3.4.5). When knowledge of not only the presence of disease but also the precise location is required, multiclass classifiers are necessary, and their results are presented next.

3.5.4 Multiclass analysis

Results of the multiclass configurations are presented here. Unlike the ENBC and IVBC classifier results presented above, here the goal is also to determine which vessel the disease is located within. A combination search is carried out with multiclass classifiers. Due to the increased computational expense, this combination search is not completed using all four classification methods used in the ENBC search. Instead the multiclass combination search is only carried out using LR and

No. of input measurements	Vessel	Combination	F_1 score	Sensitivity	Specificity
4	Aorta	(Q_3, Q_2, Q_1, P_1)	0.8437	0.8893	0.7814
	Iliac 1	(Q_3, Q_2, P_3, P_2)	0.8256	0.8439	0.7996
	Iliac 2	(Q_3, Q_2, P_3, P_2)	0.8163	0.8303	0.7961
5	Aorta	$(Q_3, Q_2, Q_1, P_2, P_1)$	0.8391	0.8775	0.7862
	Iliac 1	$(Q_3, Q_2, Q_1, P_3, P_2)$	0.8387	0.8333	0.8464
	Iliac 2	$(Q_3, Q_2, Q_1, P_3, P_2)$	0.8407	0.8406	0.8409
6	Aorta	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8363	0.8734	0.7847
	Iliac 1	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8348	0.8255	0.8479
	Iliac 2	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8364	0.8276	0.8488

Table 3.5: The combinations of input measurements that produce the maximum F_1 scores when providing four, five, and six input measurements to IVBCs employing the SVM method are shown. The corresponding sensitivities and specificities are also included.

SVM methods. These two classification methods are chosen for two reasons:

- LR and SVM classifiers have repeatedly performed better than all other classification methods. While the accuracy of classification is not a primary objective of this PoC, higher accuracy results are likely to emphasise any patterns and behaviours in the results achieved.
- Using these two methods allows for comparison of linear and non-linear partitions between healthy and unhealthy VPs. This characteristic of classifiers is highlighted as being important within the ENBC analysis.

Initially LR and SVM classifiers employ the OVA and OVO strategies, respectively. The full table of results of the multiclass combination searches are shown in Appendix E.

Comparison of the accuracy of different classifications

Looking at the tables within Appendix E, it is seen that for both classification methods the sensitivity of $\mathcal{C}^{(1)}$, *i.e.* the sensitivities of healthy VP classification, is consistently approximately 0, while the specificity of $\mathcal{C}^{(1)}$ is approximately 1. While the accuracy of classification of healthy VPs is poor the aortic disease, first iliac

disease, and second iliac disease classification accuracies are much higher. When using a LR method and pressure and flow-rate measurements at all three available locations, the sensitivity and specificity of classification of $\mathcal{C}^{(2)}$, $\mathcal{C}^{(3)}$, and $\mathcal{C}^{(4)}$ are equal to 0.6421 and 0.8513; 0.8459 and 0.7295; and 0.8248 and 0.7255 respectively. The fact that aortic disease, first iliac disease, and second iliac disease classification accuracies are high, while healthy classification accuracies are low, suggests that it is easier to differentiate between the pressure and flow-rate profiles taken from VPs with disease in different vessels than it is to differentiate between the pressure and flow-rate profiles taken from VPs with and without the inclusion of a stenosis. This finding suggests that while the inclusion of a stenosis does create biomarkers within pressure and flow-rate profiles, these biomarkers are not consistent between different arterial vessels.

Using the sensitivities and specificities achieved, the F_1 scores are computed. Due to the fact that the sensitivity of healthy classification is consistently close to 0, computation of the F_1 score associated with this classification is not possible. Instead the F_1 score of $\mathcal{C}^{(2)}$, $\mathcal{C}^{(3)}$, and $\mathcal{C}^{(4)}$ classification are computed for each combination of input measurements and classification method. These results are included within Appendix E.

A major difference that can immediately be seen across the binary and multiclass input combination searches is that in the case of the ENBC and IVBC searches SVM classifiers consistently perform better than LR. The opposite is seen to be true in the multiclass combination search. When using all six input measurements the multiclass LR classifier achieves F_1 scores of 0.7171, 0.7993, and 0.7857 for aortic, first iliac, and second iliac disease classification respectively. When using all six input measurements the SVM classifier achieves F_1 scores of 0.5811, 0.6284, and 0.6024 for aortic, first iliac, and second iliac disease classification respectively.

Importance of number of input measurements

As a clear relationship has previously been seen between the number of input measurements used and the F_1 score achieved, the average F_1 score of $\mathcal{C}^{(2)}$, $\mathcal{C}^{(3)}$, and $\mathcal{C}^{(4)}$ classification achieved by all combinations of inputs using one to six measurements are computed and plotted in Figure 3.15. As in the case of binary classification, the maximum and minimum F_1 score achieved when using each

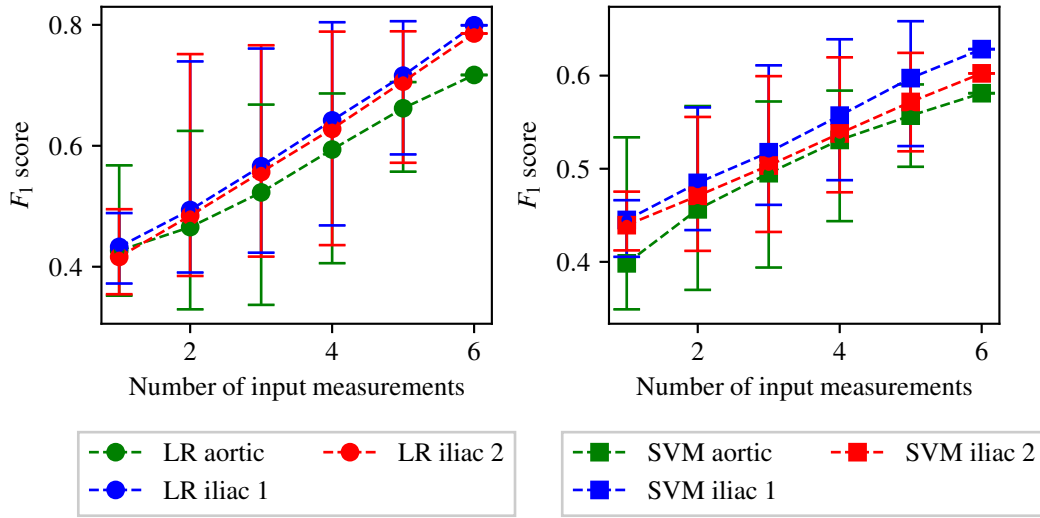


Figure 3.15: The average, maximum, and minimum F_1 score achieved by all classifiers trained within the multiclass combination search using different numbers of input measurements. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.

number of input measurements is also found and included within this comparison.

Figure 3.15 shows a high degree of similarity to the results achieved in the ENBC and IVBC searches. When using a single input measurement the average F_1 score achieved for aortic, first iliac, and second iliac disease classification is close to 0.5—representing naive classification—for both classification methods. As the number of input measurements increases, so does the average F_1 score associated with classification of each vessel. In a very similar manner to what is seen in the case of binary classification, the maximum F_1 score achieved does not appear to significantly increase between four and six input measurements. In the case of LR classifiers, the maximum F_1 score achieved when using four input measurements is equal to 0.69, 0.80, and 0.79 for aortic, first iliac, and second iliac disease classification respectively. The maximum F_1 score achieved when using five input measurements and a LR method is equal to 0.71, 0.81, and 0.79 for aortic, first iliac, and second iliac disease classification respectively. The combination of input measurements producing each of the above listed F_1 scores, and the corresponding sensitivities and specificities are shown in Table 3.6. Comparing the sensitivities and specificities shown in Table 3.6, it can be seen that there is very little difference between the maximum accuracy

No. of measurements	Vessel	Combination	Sensitivity	Specificity
4	Aorta	(Q_1, P_1, P_2, P_3)	0.6004	0.8513
	First iliac	(Q_3, Q_2, P_3, P_2)	0.8360	0.7572
	Second iliac	(Q_3, Q_2, P_3, P_2)	0.8130	0.7517
5	Aorta	$(Q_3, Q_2, P_3, P_2, P_1)$	0.6279	0.8475
	First iliac	$(Q_3, Q_2, P_3, P_2, P_1)$	0.8592	0.7273
	Second iliac	$(Q_3, Q_2, P_3, P_2, P_1)$	0.8329	0.7226

Table 3.6: The combination of input measurements that produces the highest F_1 score for aortic, first iliac, and second iliac disease classification when using four and five input measurements and employing the LR method are shown above. The corresponding sensitivities and specificities are also shown.

results achieved using four and five input measurements.

Importance of pressure at the inlet of the system

It is found in the binary input combination searches that classifiers trained using the input measurement P_1 , generally, achieved higher accuracy results than those that do not use P_1 . To test if this is also true for the multiclass search, the F_1 scores of all input combinations that include and exclude the measurement P_1 are separated and plotted for each classification. As LR outperforms SVM within the multiclass combination search, it is chosen to complete this analysis on only the LR method. The histograms of the F_1 scores of aortic, first iliac, and second iliac accuracy when including or excluding P_1 are shown in Figure 3.16.

Figure 3.16 shows that there is no real difference between the accuracy of first and second iliac disease classification when either including or excluding the input measurement P_1 . For first iliac classification 22 of the 32 input combinations that include P_1 achieve an F_1 score of less than 0.6. This compares to 23 of the 31 input combinations that exclude P_1 that achieve an F_1 score of less than 0.6. In the case of second iliac disease classification 24 of the 32 input combinations that include P_1 , and 23 of the 31 input combinations that exclude P_1 achieve an F_1 score of less than 0.6. Better distinction can be made between the aortic classifiers trained with and without P_1 . Only 9 of the 32 input combinations that include P_1 achieve an aortic

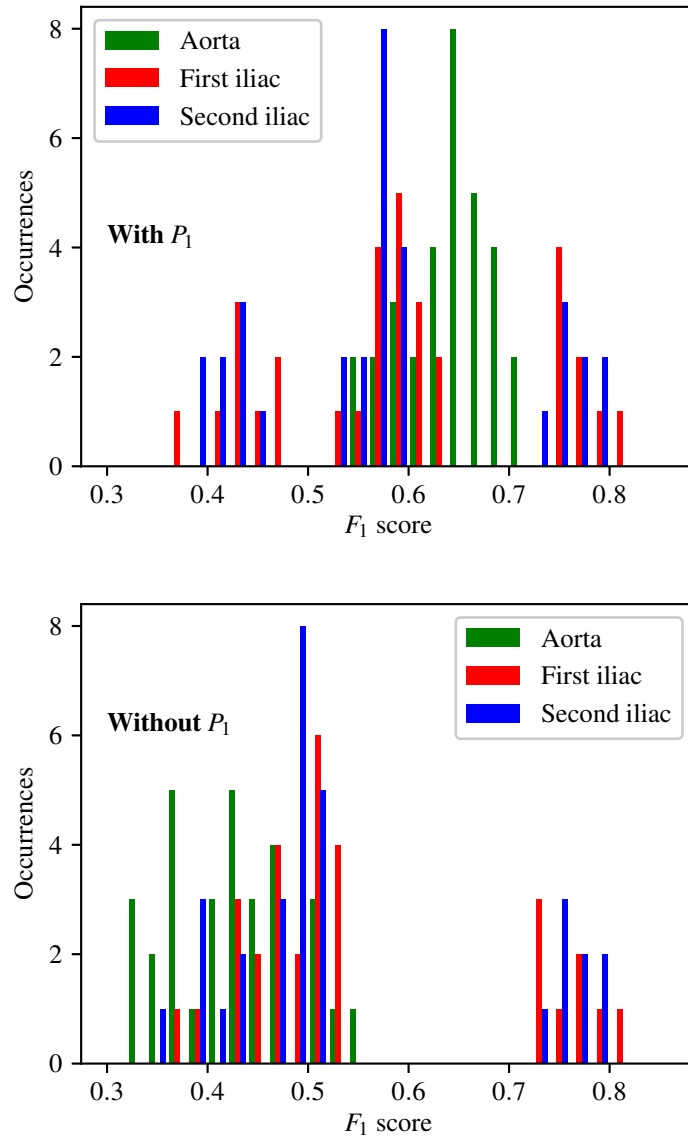


Figure 3.16: The histograms of the F_1 scores achieved for aortic, first iliac, and second iliac classification are shown for all input measurement combinations that include P_1 in the upper plot, and exclude P_1 in the lower plot. This analysis is shown for the results produced by the LR method.

F_1 score of less than 0.6. This compares to all 31 of the classifiers trained without P_1 that achieve an aortic F_1 score of less than 0.6.

The analysis carried out on the multiclass combination search has shown that similar relationships exist between the haemodynamic measurements used to train and test ML classifiers and the accuracy of classification achieved to those seen in the ENBC and IVBC combination searches. Currently, however, the performance of multiclass classifiers is severely limited by the low accuracy of healthy classification. To try and mitigate the low healthy classification accuracies, a CPC is trained and tested next.

3.5.5 Custom probabilistic configuration analysis.

It is seen in Section 3.5.4 that due to the variability in biomarkers introduced by disease in different vessels, the accuracy of classification of healthy VPs when using an OVA or OVO method is very low. It is possible to improve the multiclass results achieved by LR classifiers by instead training and testing CPC classifiers. The configuration of each instance of an LR classifier within a CPC ensemble are explained within Section 3.4.5. Rather than running a full grid search using CPC multiclass classifiers, a single classifier is trained and tested. This classifier uses the measurements of pressure and flow-rate at all three available locations. This analysis is carried out to gain an understanding of the differences in behaviour of OVA, OVO, and CPC multiclass configurations. The results achieved by the CPC classifier are shown in Table 3.7. The results achieved by the OVA, and OVO classifiers using pressure and flow-rate measurements at all three locations are also included in Table 3.7 to allow for comparison.

When comparing OVA and OVO against CPC, the highest improvement is seen for the sensitivity of healthy classification, an increase to $\sim 50\%$ for CPC compared to $\sim 0\%$ in OVA and OVO. For the aorta and iliacs, a re-balancing of sensitivities and specificities is observed relative to OVA and OVO—an increase in specificity is accompanied by a decrease in sensitivity, with their averages relatively unchanged. Overall, Table 3.7 shows that the CPC achieves its purpose of improving the classification accuracy for healthy (‘no disease’) class without significantly compromising other classification accuracies.

When creating CPC multiclass classifiers, preference can be given to healthy

Configuration	Healthy		Aorta		First iliac		Second iliac	
	Sen.	Spec.	Sen.	Spec.	Sen.	Spec.	Sen.	Spec.
One-versus-all (OVA)	0.056	0.986	0.642	0.851	0.846	0.729	0.825	0.725
One-versus-one (OVO)	0.120	0.916	0.493	0.798	0.584	0.726	0.550	0.725
Custom probabilistic config (CPC)	0.496	0.832	0.581	0.882	0.745	0.867	0.722	0.860

Table 3.7: Multiclass accuracies achieved by OVA, OVO, and CPC configurations when using pressure and flow-rate at all three location.

or unhealthy VPs by adjusting the decision boundary \mathcal{B} in equation (3.17)—*i.e.* the certainty required to override the default classification that a VP has no disease present. Reducing the certainty required to change the classification a VP is assigned to, *i.e.* lowering the decision boundary, creates bias towards unhealthy VPs as the CPC classifier is more willing to override the default classification that a VP is healthy. Increasing the decision boundary will require more certainty to classify a VP as diseased, giving bias toward healthy VPs, as the CPC is less willing to override the default classification that a VP is healthy.

To analyse the aforementioned affect of the decision boundary used on the classification of VPs, receiver operating characteristic (ROC) curves [4] are plotted. ROC curves are obtained by plotting the true positive rates against the false positive rates of each classification when different decision boundaries are applied. By recording a series of discrete true-positive/false-positive points for various decision boundaries, a curve is fitted that shows the characteristics of the accuracy of each classification across all possible decision boundaries. A complete ROC curve must start at the point representing a true positive and false positive rate of 0, *i.e.* no VPs are predicted to belong to the classification being examined, and must end at the point representing a true positive and false positive rate of 1, *i.e.* all VPs are predicted to belong to the discrete classification being examined. A naive classifier, achieving an accuracy of 50%, will produce a straight line between these two points, and so the area under the curve (AUC) is equal to 0.5. A perfect classifier ascends vertically along the the y-axis between the points (0, 0) and (0, 1), then transverses the x axis between the points (0, 1) and (1, 1). This will result in a perfect AUC score of 1. The point (0, 1) represents a perfect classifier, as 100% of positive VPs are correctly classified, while 0% of negative patients are incorrectly classified.

Within the context of the multiclass CPC, when a decision boundary of 1 is applied, all VPs are classified as healthy, and so the ROC curves of aortic disease, first iliac disease, and second iliac disease classification all begin at the true positive and false positive position (0, 0). When a decision boundary of 1 is applied, the true positive and false positive position of healthy classification is (1, 1), *i.e.* all VPs are being assigned to the classification of no disease present. When the decision boundary is set to be 0 all VPs are classified as having disease in one of the three vessels, and so the healthy classification will reach the point (0, 0). A complete ROC curve can, therefore, be obtained for healthy classification accuracy. When

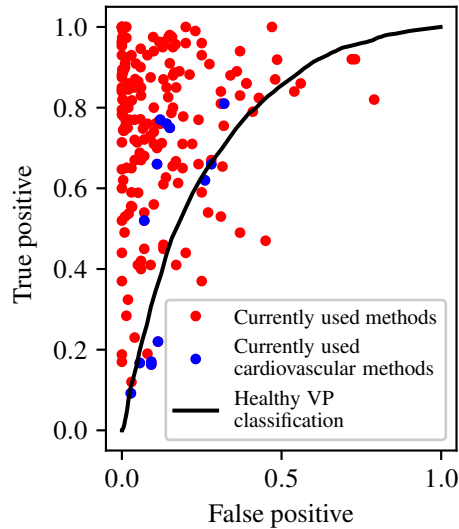


Figure 3.17: The ROC curve of healthy VP classification within the CPC ensemble trained and tested using pressure and flow-rate measurements at all three available measurement locations are shown. The accuracy of classification is compared to 193 current screening methods (general methods shown in red, and cardiovascular methods shown in blue), recorded in [132] and [7].

the decision boundary is set to be 0, while all VPs are classified as having disease in one of the three vessels, this does not necessarily ensure that the true positive and false positive classification accuracy is equal to 1 for disease classification in each individual vessel. Complete ROC curves can, therefore, not be plotted for aortic, first iliac, and second iliac disease classification.

The ROC curve of healthy VP classification accuracy is plotted against the reported true positive and false positive rates of 193 current screening methods, recorded in [132] and [7], and is shown in Figure 3.17. The current screening methods are not necessarily cardiovascular related (*i.e.* this analysis does not compare the results achieved here to current comparable methods), and instead Figure 3.17 gives a comparison of the results achieved against the general landscape of clinically used methods. While this will not allow for an evaluation of classification accuracy in each vessel, it will allow for evaluate of the overall bias given to healthy or unhealthy VPs. Figure 3.17 shows that the ROC curve of healthy VP classification follows a desirable profile. The AUC of the ROC curve is computed to be 0.75. An AUC of between 0.7 and 0.9 can be considered as moderate accuracy [62]. The overall

correct classification of healthy VPs by the CPC outperforms approximately 20 of the current methods.

The ROC curve of classifiers created in this PoC study can not be fairly compared to current screening methods, as the affects of simplifications such as only using a simple three vessel system, and limiting the number of diseased vessels to one are not understood. However, Figure 3.17 provides some indication of how the results achieved in this PoC study compare to currently used screening methods. Overall, the results, despite simplifications and assumptions used in this study, are encouraging and point towards the potential of increased classification accuracies when larger networks and more sophisticated ML or deep learning algorithms are used.

3.6 Conclusions

This PoC has shown encouraging results that a stenosis within a simple three vessel arterial network can produce consistent biomarkers within pressure and flow-rate measurements, which can be detected by ML algorithms. Four ML algorithms—NB, LR, SVM, and RF—are analysed within the frameworks of both binary and multiclass classifiers. It is shown that among the four ML methods employed, the LR and SVM perform significantly better than NB and RF. These two methods provide the further advantage that they both require little to no problem specific optimisation. Within the binary analysis it is seen that classification methods capable of producing non-linear partitions between healthy and unhealthy VPs, such as radial basis function SVM, appear to be well suited to distinguishing between healthy and unhealthy VPs. This behaviour, however, is not seen in the case of multiclass classification. More work needs to be carried out in understanding the importance of non-linear partitions between positive and negative examples of classifications on the accuracy.

A key relationship shown is that the accuracy of classification is generally proportional to the number of input measurements provided. This finding seems intuitive, as the provision of more input measurements likely results in more information being provided to ML classifiers. The correlation between the number of input measurements provided to an ML classifier and the accuracy of results achieved

may suggest a future limiting factor to using pressure and flow-rate measurements to detect arterial disease. Within a clinical application, the need for more pressure or flow-rate measurements will likely result in an increase in the cost of diagnosis. This being said, it has been shown that if the input measurements are chosen judiciously high accuracy results are achieved when using as few as four input measurements. In this context, the importance of the inlet pressure (or pressure in the diseased vessel) and the flow-split measurement (particularly when including inlet flow-rate) is highlighted.

Overall, it is shown that the approach of using a VPD for assessment of ML algorithms for disease classification and screening holds substantial potential. One of the primary purposes of the PoC study is to create an advisory template that can be expanded on within future studies. A detailed procedure for both the generation and exploitation of synthetic data sets is presented. This procedure can now be expanded on in subsequent chapters to increase the complexity, and consequently physiological realism, of the VPD. The modular methodology used to create the VPD within this PoC allow for the topology and parameterisation of VPs arterial networks; the distribution of arterial network parameters; and description of diseased vessels to all be easily changed to create more physiological realistic synthetic data sets. This PoC must now be expanded upon through the application of more sophisticated ML algorithms trained on a larger comprehensive arterial network. This assessment forms the primary body of the remainder of this thesis.

3.7 Limitations

Several simplifications and assumptions are made during both the creation of the VPD and the training and testing of ML classifiers. These are likely to affect the accuracy of classification achieved within this study. Some of these major limitations are:

- The arterial network, containing only three vessels, is small. It is not clear whether this aids or hinders classification. On the one hand, due to small nature of the network, the signals are less diffused, and on the other hand specific features which may be the result of unique reflections in certain anatomical locations is not accounted for. However, the small arterial network

does allow for a preliminary analysis which, with encouraging results, points towards exploration in larger networks.

- The distribution of all arterial network parameters across the VPD are described using independent distributions. These simple distributions ignore the complex inter-parameter relationships likely seen within real arterial networks. The simplification of the distribution of arterial network parameters likely results in a wider range of pressure and flow-rate profiles across the VPD, making distinction between healthy and unhealthy VPs more difficult. This may be potentially solved by first determining the probability distributions through an inverse problem approach, for example a Markov-chain Monte Carlo method.
- This study is completed without significant consideration for clinical requirements. For example which measurements are really obtainable easily, and what range of stenosis severities should a ML classifier be able to detect? These questions are best explored on a larger network.

Chapter 4

Creation of a physiologically realistic virtual patient database

4.1 Introduction

In Chapter 3 the potential of using ML classifiers to predict the presence of stenosis, using only pressure and flow-rate measurements, is shown. This study, however, is limited by several factors:

- The PoC study is completed on a very small arterial network, containing only three vessels.
- The distribution of all parameters describing the arterial networks of VPs, excluding disease parameters, across the VPD are assumed to be independent and normally distributed.
- The PoC study is completed without significant consideration for clinical requirements, *i.e.* what range of stenosis severities should a ML classifier be able to detect. There is also no consideration given to clinical limitations, *i.e.* what measurements are available inexpensively.

To be able to more comprehensively assess the potential of using ML classifiers to detect arterial disease, a new VPD is required. To address the issues outlined above it must be ensured that the new VPD is representative of a real population, on both an individual patient level and across the database. This study presents a probabilistic framework for creation of such a VPD based on a previously proposed arterial network [20]. This network is referred to as the initial network and is a reduced version of the anatomically detailed arterial network (ADAN) presented in [18] [17].

In a similar manner to Chapter 3, the physics based model of pulse wave propagation is used to compute the pressure and flow-rate profiles corresponding to random realisations of VPs arterial networks. Unlike in the PoC, however, overtly stated *a priori* distributions are not assumed for the parameters describing

VPs. Instead prior distributions are constructed based on known physiological restrictions and geometrical constraints. The affect of literature based measurements of pressure and flow-rate are incorporate into these prior distributions—through the employment of a Bayesian framework—creating a physiologically realistic posterior distribution. An appropriate sampling method, in this case a Markov chain Monte-Carlo (MCMC) method [69, 71], is used to draw random realisations of VPs arterial networks from this posterior distribution. A recent study [36] has published a VPD created by varying six cardiovascular properties by ± 1 SD from their age-specific mean values for subjects of each age decade from 55 to 75 years. The approach taken here differs from [36] as literature reported measurements are used to create an informed distribution from which random realisations are sampled, as opposed to using the literature reported measurement to dictate each individual realisation. The approach taken by [36] will result in each VP being highly physiologically likely, however significantly less variability across the resulting VPD.

This chapter begins by comparing the discontinuous Galerkin (DCG) implementation, used to numerically solve the governing system of equations of the physics based model in the previously completed PoC, to a more computationally efficient subdomain collocation (SDC) implementation [30]. While adoption of the SDC implementation has the potential to significantly reduce the computational time required to create the new VPD, it is important to ensure the use of this implementation does not have any adverse affects on the pressure and flow-rate profiles produced.

Next an understanding of both where disease is likely to occur within the arterial network, and where time varying pressure and flow-rate profiles can be obtainable is gained. Using this information, the topology of the initial network is optimised by removing vessels peripheral to both measurement obtainment and disease occurrence locations, however ensuring this has minimal impact on the upstream pressure and flow-rate profiles. The parameterisation used to describe the arterial network of VPs is chosen to minimise the dimensionality associated with each VP, while imposing required behaviours on the geometrical and mechanical properties of the network. These imposed behaviours are carefully considered to maximise the variability in the realisations of arterial networks, and so consequently pressure and flow-rate profiles, produced across the final VPD.

Prior distributions are constructed for the parameters describing the arterial

networks of VPs by examining and understanding the known physiological restrictions and geometrical constraints each is subject to. The prior distributions prescribed are uninformative or weakly informative and so primarily act as a regularising term, while the literature based measurements incorporated into the posterior distribution predominantly control the shape. The literature based measurements being incorporated into the posterior distribution are carefully selected to update this distribution with regards to particular behaviours of the pressure flow-rate profiles produced across the final VPD.

Finally VPs are sampled from the posterior distribution through the use of a Markov chain Monte-Carlo (MCMC) method [69, 71]. The pressure and flow-rate profiles corresponding to VPs are computed using the SDC implementation of the physics based model. This chapter culminates in the *a posteriori* behaviour of both the arterial network parameters and the resulting pressure and flow-rate profiles across the new VPD being examined.

The majority of work in this chapter is published in: *Jones G, Parr J, Nithiarasu P, Pant S. A physiologically realistic virtual patient database for the study of arterial haemodynamics. International Journal for Numerical Methods in Biomedical Engineering. 2021 May 10:e3497.*

4.2 Key criteria of the new VPD

The arterial network used as a basis for the creation of new VPs is taken from [20], and is based on a version of the ADAN network [18] [17]. This network is referred to as the initial network. The new VPD is created with preference being given to the variability of VPs produced. Behaviours and relationships that are expected to be seen in the arterial networks of VPs, however are not required geometrically or physiologically (for example symmetry between opposite extremities) are not forcibly imposed. Instead it is expected that the incorporation of adequate literature based measurements into the distribution of arterial network parameters will introduce and impose required behaviours. The new VPD is designed to meet the following five key criteria:

1. **Consistency with initial network.** A realisation of a VP's arterial network that results in geometric and mechanical properties comparable to

the initial network must produce consistent pressure and flow-rate profiles at key locations. By ensuring a comparable realisation produces pressure and flow-rate profiles consistent with the initial network, it can be shown that no intra- or inter-measurement details are being lost through simplifications.

2. **Minimisation of dimensionality.** The dimensionality required to describe each VP must be minimised. Reductions to the dimensionality of VPs arterial networks, however, must not come at the cost of the first key criteria.
3. **Minimisation of computational time.** The computational time required to produce each VP must be minimised. As with the reduction of the dimensionality of VPs, reductions in the computational time required to produce VPs must not come at the cost of the first key criteria.
4. **Freedom to produce variability.** Stringent inter-parameter relationships must be avoided. Inclusion of these relationships is only permitted if the resulting reduction in the dimensionality, or the reduction in the likely frequency at which unrealistic VPs occur significantly outweighs the reduction in variability imposed.
5. **Physiological realism.** It must be ensured that the new VPD is representative of a real cohort, on both an individual VP and population-wide level. This is to be achieved through the incorporation of both literature reported measurements, and known physiological restrictions and geometrical constraints into the VPD. The known physiological restrictions and geometrical constraints must be carefully balanced against the fourth key criteria.

4.3 One-dimensional solver

In the previously completed PoC a discontinuous Galerkin (DCG) implementation is used to numerically solve the governing system of equations of the physics based model of pulse wave propagation, used to compute the pressure and flow-rate profiles corresponding to random realisations of VPs arterial networks. This implementation is chosen as a pre-existing solver is available that has successfully been validated

against a 3D model of blood flow through stenosed arterial vessels [21]. The drawback of this numerical implementation, however, is its computational expense and so consequently the large time required to produce each VP. Before completion of this secondary study a more computationally efficient solver is made available. This solver implements the subdomain collocation (SDC) method. This solver and the SDC implementation are outlined in [30].

Before adopting the SDC implementation, it is important to compare results achieved by the DCG and SDC implementations. To compare these two implementations the aortic bifurcation network set out in [20]—as used in PoC—is solved using each for a healthy, stenosed, and aneurysm VP. Both the stenosed and aneurysm VPs are created with their respective disease affecting the aorta, with the vessel area profiles created using the method presented in the PoC chapter. Both diseases are created with a start and end location of 20% and 80% of the vessel length respectively. The stenosed VP has the inclusion of a 60% reduction in area, while the aneurysm VP has a 150% increase in area.

To allow for fair comparison of the two methods a mesh independence study is performed on each, presented in Appendix F. The healthy, stenosed, and aneurysm VPs are solved using each implementation when applying the optimised meshes detailed in Appendix F. The percentage discrepancy between the pressure and flow-rate profiles recorded within each VP when using each of the two implementations are computed using the error matrices given in [58]:

$$E_{P,\text{avg}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{P_{\text{SDC},i} - P_{\text{DCG},i}}{P_{\text{DCG},i}} \right|, \quad E_{Q,\text{avg}} = \frac{1}{N} \sum_{i=1}^N \left| \frac{Q_{\text{SDC},i} - Q_{\text{DCG},i}}{\max(\mathbf{Q}_{\text{DCG}})} \right|, \quad (4.1)$$

$$E_{P,\text{sys}} = \frac{\max(\mathbf{P}_{\text{SDC}}) - \max(\mathbf{P}_{\text{DCG}})}{\max(\mathbf{P}_{\text{DCG}})}, \quad E_{Q,\text{sys}} = \frac{\max(\mathbf{Q}_{\text{SDC}}) - \max(\mathbf{Q}_{\text{DCG}})}{\max(\mathbf{Q}_{\text{DCG}})}, \quad (4.2)$$

$$E_{P,\text{dias}} = \frac{\min(\mathbf{P}_{\text{SDC}}) - \min(\mathbf{P}_{\text{DCG}})}{\min(\mathbf{P}_{\text{DCG}})}, \quad E_{Q,\text{dias}} = \frac{\min(\mathbf{Q}_{\text{SDC}}) - \min(\mathbf{Q}_{\text{DCG}})}{\max(\mathbf{Q}_{\text{DCG}})}, \quad (4.3)$$

where $E_{P,\text{avg}}$, $E_{Q,\text{avg}}$, $E_{P,\text{sys}}$, $E_{Q,\text{sys}}$, $E_{P,\text{dias}}$, and $E_{Q,\text{dias}}$ represent the average pressure and flow-rate discrepancy; the systolic pressure and flow-rate discrepancy;

and the diastolic pressure and flow-rate discrepancy respectively. In the above equations P_{SDC} , Q_{SDC} , P_{DCG} , and Q_{DCG} represent the pressure and flow-rate profiles associated with the SDC and DCG implementations respectively. For each time point i the pressure produced by the SDC and DCG implementations are described by $P_{\text{SDC},i}$ and $P_{\text{DCG},i}$ respectively; and the corresponding flow-rate measurements are described by $Q_{\text{SDC},i}$ and $Q_{\text{DCG},i}$ respectively. Finally N represents the number of discrete time points within the cardiac period.

The pressure profile percentage discrepancies are shown Table 4.1, and the flow-rate in Table 4.2. A visual comparison of the pressure and flow-rate profiles produced by each of the two implementations are shown in Figures 4.1, 4.2, and 4.3 for the healthy, stenosed, and aneurysm VPs, respectively. Figures 4.1, 4.2, and 4.3 show very little visual discrepancy between the pressure and flow-rate profiles produced when using each of the two implementations. This is further supported by Tables 4.1 and 4.2. In the case of pressure, the maximum errors are consistently seen in diastole. The maximum diastolic pressure error recorded in the healthy, stenosed, and aneurysm VPs are equal to -0.701%, -0.630%, and -1.333%, respectively. In the case of the flow-rate, the systolic and diastolic discrepancies are very similar. The maximum diastolic flow-rate error recorded in the healthy, stenosed, and aneurysm VPs are equal to 2.320%, 2.379%, and 2.117%, respectively. The corresponding maximum systolic flow-rate errors are equal to -2.355%, -2.466%, and 2.210%, respectively. While there are some discrepancy between the pressure and flow-rate profiles produced using each of the two implementations, the maximum error does not exceed 2.5%.

The wall clock time require to solve the healthy, stenosed, and aneurysm VPs using each of the two implementations are shown in Table 4.3. Table 4.3 shows significant reduction in wall clock time required to solve the physics based model when employing the SDC implementation, relative to the DCG implementation. Using the SDC implementation reduces the wall clock time required to solve the healthy, stenosed, and aneurysm VPs by 99.69%, 99.50%, and 99.49% respectively. Due to the large decrease in wall clock time and the relatively low discrepancies in pressure and flow-rate profiles produced, all future VPs are solved using the SDC implementation.

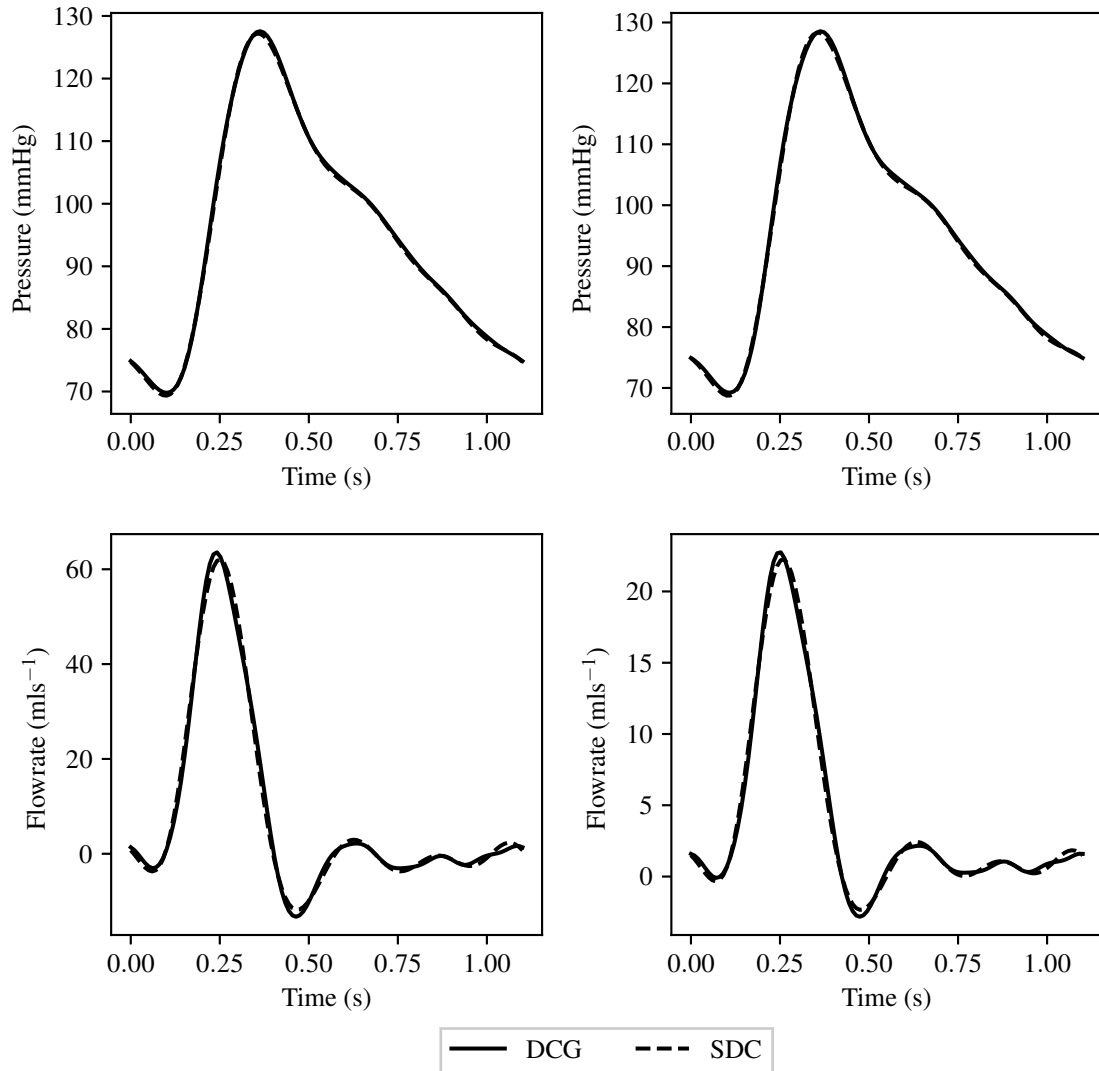


Figure 4.1: Comparison of the pressure (upper row) and flow-rate (lower row) profiles produced at the distal end of the abdominal aorta (left column) and first common iliac (right column), using each of the two numerical methods to solve the healthy VP.

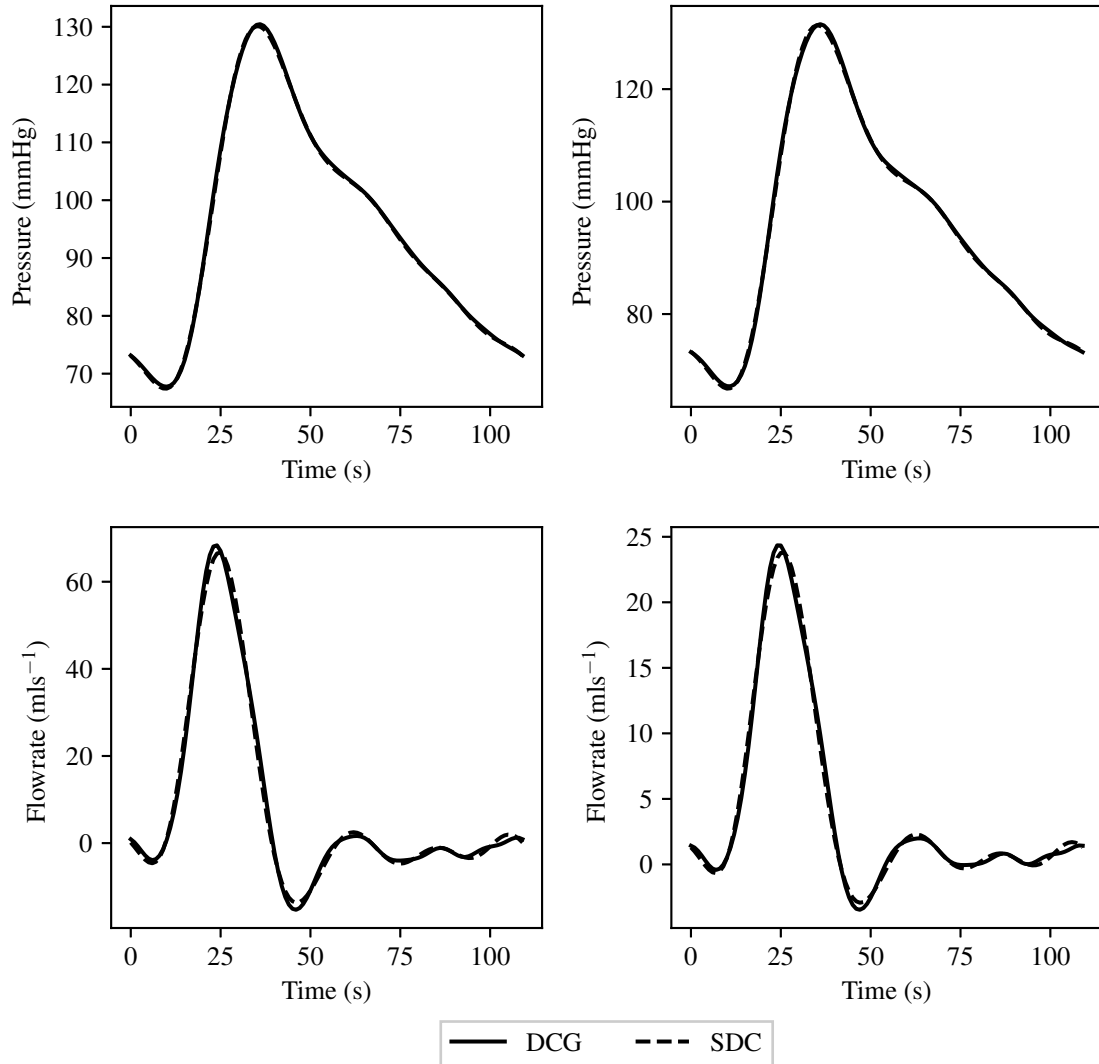


Figure 4.2: Comparison of the pressure (upper row) and flow-rate (lower row) profiles produced at the distal end of the abdominal aorta (left column) and first common iliac (right column), using each of the two numerical methods to solve the stenosed VP.

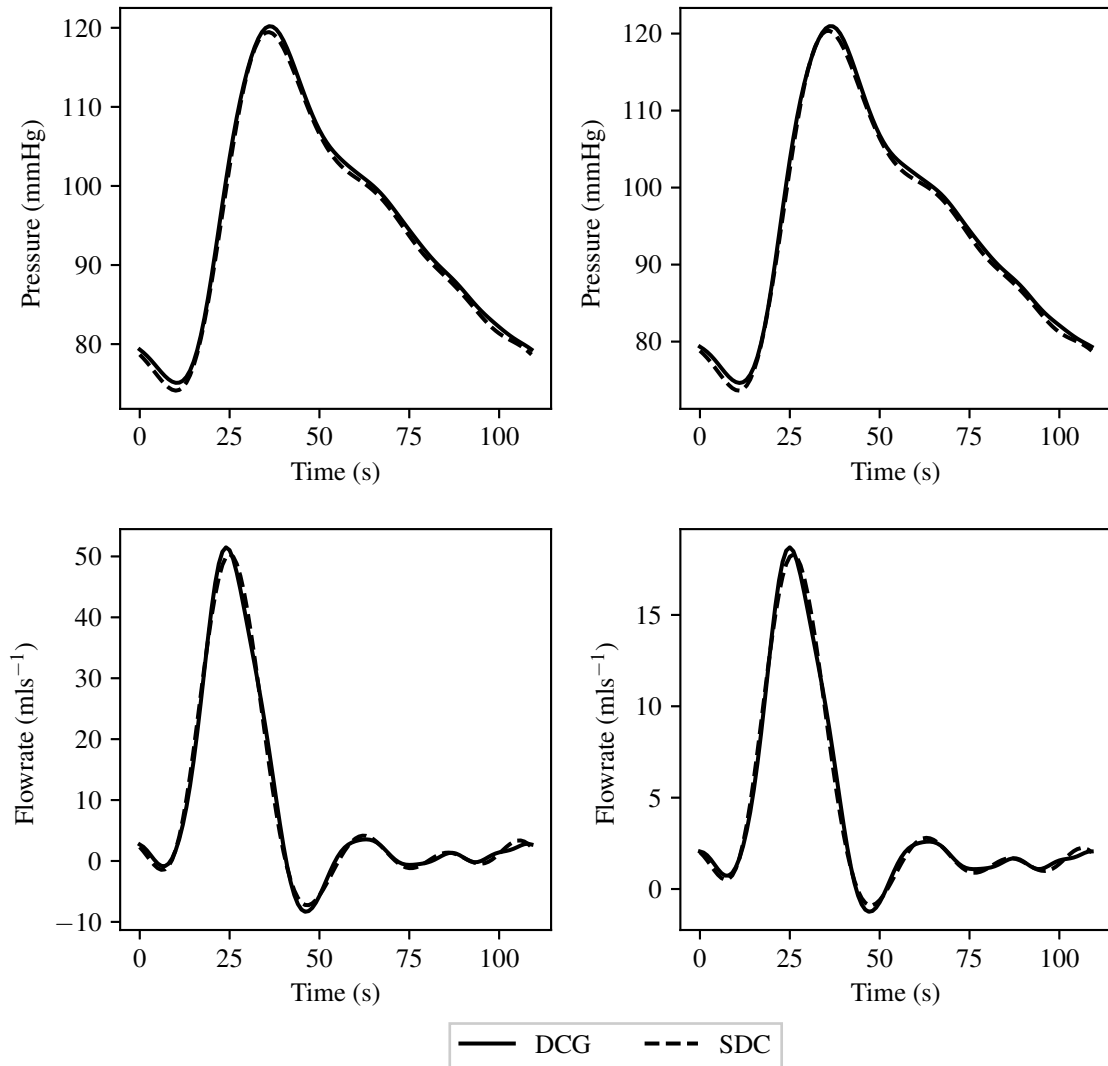


Figure 4.3: Comparison of the pressure (upper row) and flow-rate (lower row) profiles produced at the distal end of the abdominal aorta (left column) and first common iliac (right column), using each of the two numerical methods to solve the aneurysm VP.

Health	Location	Average discrepancy (%)	Systolic discrepancy (%)	Diastolic discrepancy (%)
Healthy	Aorta	0.302	-0.320	-0.637
	Right iliac	0.341	-0.191	-0.701
Stenosed	Aorta	0.263	-0.267	-0.502
	Right iliac	0.318	-0.146	-0.630
Aneurysms	Aorta	0.748	-0.619	-1.312
	Right iliac	0.719	-0.495	-1.333

Table 4.1: Percentage discrepancies between the pressure profiles produced when using the SDC and DCG implementations to solve the healthy, stenosed and aneurysm VPs, based on the error metrics presented in [58].

Health	Location	Average discrepancy (%)	Systolic discrepancy (%)	Diastolic discrepancy (%)
Healthy	Aorta	1.456	-2.355	2.320
	Right iliac	1.402	-2.210	2.141
Stenosed	Aorta	1.489	-2.466	2.379
	Right iliac	1.443	-2.351	2.208
Aneurysms	Aorta	1.423	-2.210	2.117
	Right iliac	1.359	-2.104	1.907

Table 4.2: Percentage discrepancies between the flow-rate profiles produced when using the SDC and DCG implementations to solve the healthy, stenosed and aneurysm VPs, based on the error metrics presented in [58].

Case	Time (s)	
	DCG	SDC
Healthy	8379.53	25.78
Stenosed	8170.71	40.42
Aneurysm	8389.02	42.14

Table 4.3: Wall clock time required to obtain the pressure and flow-rate profiles corresponding to each of the three VPs when using both implementations to solve the physics based model.

4.4 Arterial network topology

4.4.1 Initial network

The initial network contains the largest 56 vessels within the human arterial system, described by 77 arterial vessel segments. The topology of this network is shown in Figure 4.4, and each vessel within the network identified in Table 4.4. This network is described by:

- **38 vessel segments with constant reference radii (r_0) and vessel wall mechanical properties (β) along their lengths.** The description of each of these segments requires three parameters: the length L , reference radius r_0 , and mechanical property β of the segment.
- **39 vessel segments with linearly varying r_0 and β along their lengths.** The description of each of these segments requires specification of five parameters: the length L ; the reference radii r_0 at the proximal and the distal ends of the segment; and β at the proximal and the distal ends of the segment.
- **31 Windkessel models.** Every Windkessel model at the outlets requires 2 resistances and a compliance.
- **One inlet flow-rate profile.** The inlet flow-rate is parameterised by a 5th order FS (see Chapters 2 and 3). This requires specification of 11 coefficients.

Thus, to describe a VP through the direct specification of all the above parameters results in the total dimensionality of 413 per VP. This high dimensionality is

problematic leading to increased complexity in creating VPs [116]. In what follows, methods to reduce the dimensionality of VPs description are presented. These are primarily related to either reducing the network or employing a parsimonious parameterisation. Since the primary purpose of the VPD is deployment of ML methods for screening of stenosis and aneurysm through easily acquirable measurements, it is important that the reduction in dimensionality does not compromise i) the locations where disease and measurements are possible, and ii) the precision and variability of measurements at such locations. Thus, before any effort is made to reduce the dimensionality, the aforementioned locations must be identified.

4.4.2 Important locations

A review of literature is carried out to understand both where disease is likely to occur and where measurements can be obtained. The latter are restricted to locations at which continuous profiles can be recorded non-invasively. While arterial disease can occur in a large number of vessels within the network, for this study vessels with only high prevalence stenosis or aneurysm are considered.

Locations of measurements

Based on literature the following locations where measurements can be taken are identified:

Pressure in the radial and common carotid arteries: continuous non-invasive arterial blood pressure profiles can be obtained in the radial and common carotid arteries using applanation tonometry [3, 146]. This method measures blood pressure by applying a mild pressure to the artery being examined. This pressure causes the artery to flatten as both vessels are located above a bone—the radius bone and cervical vertebra for the radial and common carotid artery, respectively. The opposing pressure being applied by blood flow is then measured using a strain gauge. A more complete explanation of this method can be found in [52] and [54]. Applanation tonometry is already clinically used [118]. The right and left radial and common carotid arteries are identified in Figure 4.4 as vessels 8 and 22; and 5 and 14, respectively.

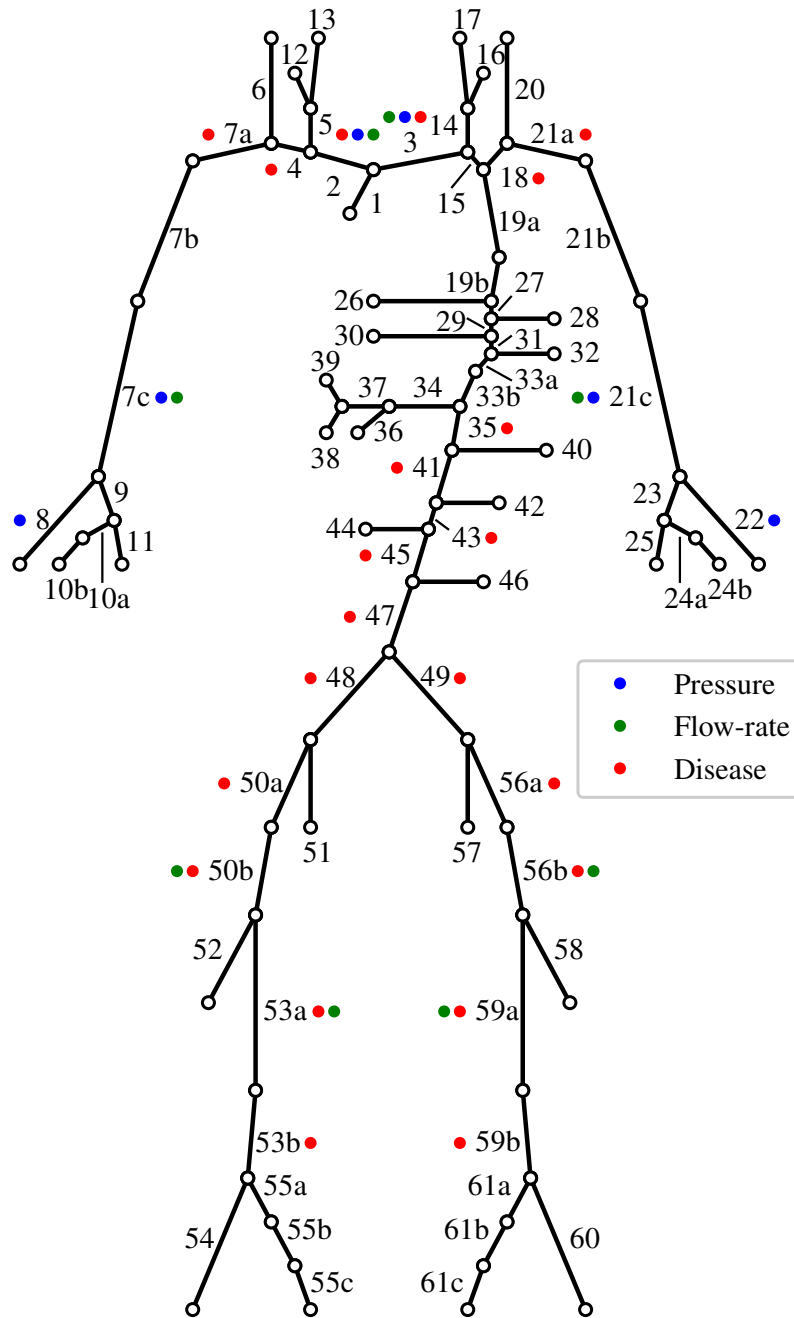


Figure 4.4: The connectivity of the initial network, taken from [20]. At the inlet (free end) of vessel 1, a volumetric flow-rate is specified and at all outlets (free ends of the terminal vessels), a Windkessel model is specified. Locations at which pressure and flow-rate can be measured; and disease is likely to occur are also highlighted, see Section 4.4.2.

4.4. ARTERIAL NETWORK TOPOLOGY

Reference number	Vessel name	Reference number	Vessel name
1	Aortic arch I	34	Celiac trunk
2	Brachiocephalic trunk	35	Abdominal aorta I
3	Aortic arch II	36	Common hepatic
4	Subclavian R I	37	Splenic I
5	Common carotid R	38	Left gastric
6	Vertebral R	39	Splenic II
7a	Subclavian R II	40	Superior mesenteric
7b	Axillary R	41	Abdominal aorta II
7c	Brachial R	42	Renal L
8	Radial R	43	Abdominal aorta III
9	Ulnar R I	44	Renal R
10a	Common interosseous R	45	Abdominal aorta IV
10b	Posterior interosseous R	46	Inferior mesenteric
11	Ulnar R II	47	Abdominal aorta V
12	External carotid R	48	Common iliac R
13	Internal carotid R	49	Common iliac L
14	Common carotid L	50a	External iliac R
15	Aortic arch III	50b	Femoral R I
16	External carotid L	51	Internal iliac R
17	Internal carotid L	52	Profunda femoris R
18	Subclavian L I	53a	Femoral R II
19a	Aortic arch IV	53b	Popliteal R I
19b	Thoracic aorta I	54	Anterior tibial R
20	Vertebral L	55a	Popliteal R II
21a	Subclavian L II	55b	Tibiofibular trunk R
21b	Axillary L	55c	Posterior tibial R
21c	Brachial L	56a	External iliac L
22	Radial L	56b	Femoral L I
23	Ulnar L I	57	Internal iliac L
24a	Common interosseous L	58	Profunda femoris L
24b	Posterior interosseous L	59a	Femoral L II
25	Ulnar L II	59b	Popliteal L I
26	Posterior intercostal R 1	60	Anterior tibial L
27	Thoracic aorta II	61a	Popliteal L II
28	Posterior intercostal L 1	61b	Tibiofibular trunk L
29	Thoracic aorta III	61c	Posterior tibial L
30	Posterior intercostal R 2		
31	Thoracic aorta IV		
32	Posterior intercostal L 2		
33a	Thoracic aorta V		
33b	Thoracic aorta VI		

Table 4.4: The 56 arterial vessels, described by 77 segments, within the initial network, taken from [20]. The numbers assigned to each vessel within this table align with those in Figure 4.4.

Pressure in the brachial arteries: It is possible to estimate continuous blood pressure at the brachial arteries through reconstruction of finger arterial pressure [77]. This method involves taking non-invasive continuous pressure measurements in the finger, and then using a numerical model to compute the estimated pressure in the brachial artery. Although the use of a model to estimate brachial pressure will introduce errors, these recreated brachial pressure profiles have been shown to meet the requirements set by the association for the advancement of medical instrumentation [77, 76]. The right and left brachial arteries are identified in Figure 4.4 as vessels 7c and 21c, respectively.

Flow-rate in the brachial, carotid, and femoral arteries: Using Doppler ultrasound techniques it is possible to measure blood velocity in both the upper and lower extremities. Doppler ultrasound computes blood velocity from the frequency shift between high frequency sound waves sent from a probe placed above an artery, and the returning reflected wave. Doppler ultrasound has been shown to work on the brachial [27], common carotid [144], and femoral [164] arteries. The first and second right and left femoral arterial segments are identified in Figure 4.4 as vessels 50b and 53a; and 56b and 59a, respectively.

While Doppler ultrasound can be used to determine arterial vessel area, and so consequently the volumetric flow-rate can be computed, use of these measurements must be carefully considered. There are two common methods used to estimate volumetric flow-rate using Doppler ultrasound [72]:

- The velocity profile measurement method records the velocity at a series of points, and then sums the flow-rate contribution of each across the vessel area. This method is reliant on high spatial resolution.
- The uniform insonification method finds the average velocity across the vessel, and then multiplies this by the vessel area. This method can not be used when vessels are very small, or very large.

Both of the above methods rely on knowing information about the vessel area. Measuring vessel area using Doppler ultrasound often introduces errors due to the fact that:

- While the time varying diameter of an artery can be recorded, it is not possible to measure both the diameter and the blood velocity during the same cardiac

period [56].

- Computation of the cross sectional area is often based on the assumption that the vessel being examined is cylindrical, and so any variations to this shape induces error. Computing the cross sectional area from the diameter of the vessel also results in any errors in the measurement being increased due to the diameter being squared.

Locations of disease

Four of the most common forms of arterial disease are carotid artery stenosis (CAS), subclavian artery stenosis (SAS), peripheral arterial disease (PAD, a form of stenosis), and abdominal aortic aneurysm (AAA). The locations of each are outlined next.

CAS: the carotid arteries consist of the common carotid, external carotid, and internal carotid segments. While the narrowing of an arterial vessel can occur within any of the three carotid segments [53], it is chosen to limit its occurrence to the common carotid arteries. The right and left common carotid arteries are identified in Figure 4.4 as vessels 5, and 14 respectively.

SAS: The initial network splits the subclavian arteries into two segments. The right and left instances of the first and second subclavian artery segments are identified in Figure 4.4 as vessels 4 and 7a; and 18 and 21a respectively.

PAD: The third frequent form of stenosis is PAD, referring to the stenosis of any peripheral vessel. Isolating arterial vessels that are likely to experience stenosis at a high prevalence in patients with PAD is more difficult relative to CAS or SAS. Both CAS and SAS have short and relatively easy to define lists of possible vessels they can affect. PAD, on the other hand, can cover a large range of different vessels depending on the definition prescribed. Allowing stenosis to be created in a very high number of vessels will likely introduce difficulties when training and testing ML classifiers. As has been seen in Chapter 3, the haemodynamic biomarkers introduced vary dependent on the location of disease. Allowing disease to be created in a high number of vessels will likely result in more diverse pressure and flow-rate profiles, and so a more challenging classification problem. While it is important to ensure ML classifiers are able to detect all major forms of stenosis, within the scope of this thesis limitations must be imposed on the number of vessels in which disease can

occur. Studies into PAD primarily focus on the lower extremities [115, 39, 1] and, thus, it is assumed that patients with PAD have the location of stenosis limited to the legs. PAD within the lower extremities is split into two categories—proximal and distal—in [39]. Proximal PAD is considered to affect the common femoral, while distal PAD is considered to affect the posterior tibial and dorsalis pedis. In study [39] it is stated that 77% of patients with any form of PAD have proximal PAD, either in isolation or in combination with the occurrence of distal PAD. Within [1] the number of subcategories is increased to three: aorta-iliac, femoral-popliteal, and infragenicular. It is reported in study [1] that 94.2% of patients with PAD have aorta-iliac or femoral-popliteal stenosis either in isolation or in combination with one of the other three sub-categories. It is chosen, based on studies [39] and [1], that the vessels in which PAD can occur will be restricted to the common iliacs; external iliacs; first and second femoral segments; and the first popliteal segments. The right and left instances of each of these vessels are identified in Figure 4.4 as vessels 48 and 49; 50a and 56a; 50b and 56b; 53a and 59a; and 53b and 59b respectively.

AAA: The most common form of arterial aneurysm is AAA, with a prevalence of 4.8% [122]. The initial network splits the abdominal aorta into five segments. These five segments are identified in Figure 4.4 as vessels 35, 41, 43, 45, and 47 respectively.

4.4.3 Network reduction

In this section, the reduction of the network by removing vessels while preserving the disease and measurement locations is presented. It is important to ensure that the removal of any vessels does not have a significant impact on the upstream pressure and flow-rate profiles. Network reduction can be performed by removing terminal vessels and merging them into the lumped parameter (Windkessel) boundary conditions [58]. This process is adopted in this study and summarised in Figure 4.5.

In [58], the accuracy of reduced networks is computed through comparison of pressure and flow-rate measurements taken at the aortic root and thoracic aorta from the full and reduced networks. It is seen that when peripheral vessels are removed there is little discrepancy between the pressure and flow-rate profiles produced by the full and reduced networks. As more central vessels are removed, for example

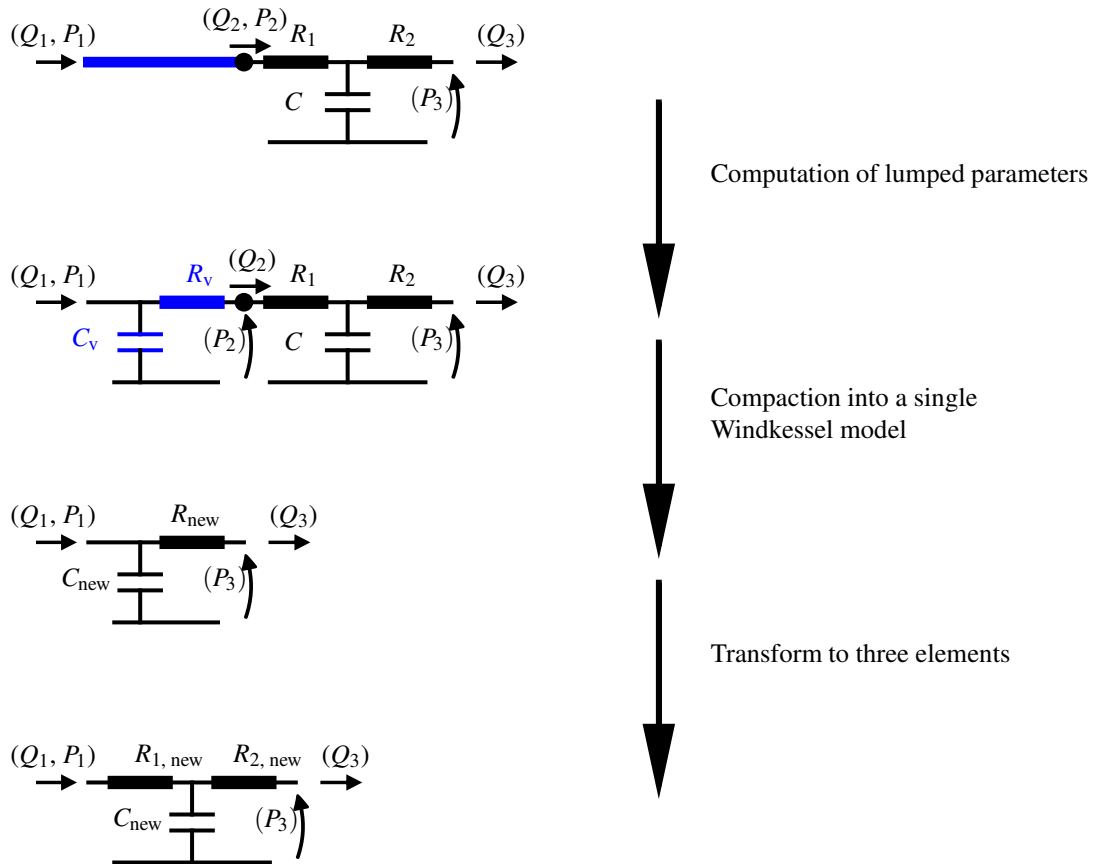


Figure 4.5: The process of incorporating peripheral 1D vessels into the 0D terminal boundary Windkessel model parameters is outlined above. In the above figure Q_1 and P_1 represent the flow-rate and pressure at the proximal end of the vessel that is being removed; Q_2 and P_2 represent the flow-rate and pressure at the distal end of the vessel that is being removed; Q_3 and P_3 represent the the flow-rate and pressure at the outlet of the system; R_1 , R_2 , and C represent the resistances and compliance of the original terminal Windkessel model; R_v , and C_v represent the viscous resistance and compliance of the vessel being removed; and R_{new} and C_{new} represent the resistance and compliance of the new 2 element Windkessel model after the incorporation of the 1D vessel.

the abdominal aorta, the pressure and flow-rate discrepancies between full and reduced networks increases. Peripheral vessels are smaller and have lower flow-rate than central vessels. This means that—relative to the large volumetric flow-rate of the proximal aorta—errors in central vessel pressure and flow-rate profiles are likely to cause more significant affects than the same percentage discrepancy, however in pressure and flow-rate profiles at peripheral locations. As pressure and flow-rate measurements are being taken from VPs within several bifurcation of terminal boundaries, discrepancies in low flow-rate peripheral vessels do not have the same opportunity to be diluted. It is, therefore, important to ensure there are no significant differences to the pressure and flow-rate measurements taken from the full 77 segment network and any proposed reduced network at all measurement locations.

Vessels within the initial network that are distal to both disease occurrence and measurement obtainment locations are identified. The following three groups of vessels eligible for removal are identified:

- **Group 1:** The first and second splenic segments; the left gastric; and the common hepatic (identified in Figure 4.4 as vessels 37 and 39; 38; and 36).
- **Group 2:** The common interosseous, the posterior interosseous, and the second ulnar segment (identified in Figure 4.4 as vessels 10a and 24a; 10b and 24b; and 11 and 25).
- **Group 3:** The second popliteal segment, the anterior tibial, the tibiofibular trunk, and the posterior tibial (identified in Figure 4.4 as vessels 55a, 54, 55b, and 55c on the right side; and 61a, 60, 61b, and 61c on the left side).

Three possible reduced networks are proposed, each with the removal of a single group of vessels (group 1, 2, or 3 in isolation). The pressure and flow-rate profiles produced by each reduced network are compared against the initial network at all measurement locations (see Section 4.4.2). If the maximum error induced by the removal of a group of vessels is less than 2%—computed using the error matrices in Section 4.3 and [58], and relative to the reference network—the full group of vessels are omitted from the arterial network. Otherwise, the vessel segments at the proximal end of the group are re-introduced into the reduced network. The process of re-introducing vessels if the error exceeds 2% is iteratively repeated until

4.4. ARTERIAL NETWORK TOPOLOGY

the maximum number of vessels that can be removed from each group individually without exceeding the imposed allowable error is found. Finally these vessels from the three groups are removed simultaneously, to ensure there is no significant amplifying affect on the errors produced. The results of this process are shown in Figures 4.6–4.11 and Tables 4.5–4.10.

The first group of vessels to be trialled for removal from the initial network consists of the first and second splenic segments; the left gastric; and the common hepatic arteries. This group of arteries forms a sub-network within the aortic region, connected across two bifurcations and resulting in three terminal boundaries. It is expected that the removal of these four vessels will produce low errors at all measurement locations, as these vessels are not directly in series with any measurement locations and have low average flow-rate. Figure 4.6 shows no visual differences between the pressure and flow-rate profiles at each measurement location. Table 4.5 shows the largest flow-rate error is occurring in the second femoral segments, with the right and left second femoral segments producing systolic errors of 0.52% and 0.515%, respectively. The largest pressure errors are occurring in the brachial arteries, with the right and left brachial arteries producing systolic errors of 0.636% and 0.642%, respectively. The errors produced when removing the first group of vessels are significantly lower than the 2% threshold imposed, and so the full group of vessels are removed from the arterial network used to create VPs. This reduction compacts four vessels with constant properties along their length and three Windkessel models, into a single Windkessel model.

The second group of vessels to be trialled for removal from the initial network consists of the common interosseous, posterior interosseous, and the second ulnar segments. This group of vessels forms a sub-network at the peripheries of the right and left arms, connected across one bifurcation and resulting in two terminal boundaries within each arm. This group are directly in series and close in spatial location to the brachial and radial arteries. Thus, it is expected that the removal of these vessels will induce larger errors within the brachial and radial arteries. The low average flow-rate in this group of vessels, however, will likely result in the error being significantly mitigated by the time it propagates to all other measurement locations. Figure 4.7 shows that, as is expected, there are large discrepancies between the pressure and flow-rate profiles in the brachial and radial arteries. No visual differences can be seen in the carotid and femoral arteries. Table 4.6 shows

4.4. ARTERIAL NETWORK TOPOLOGY

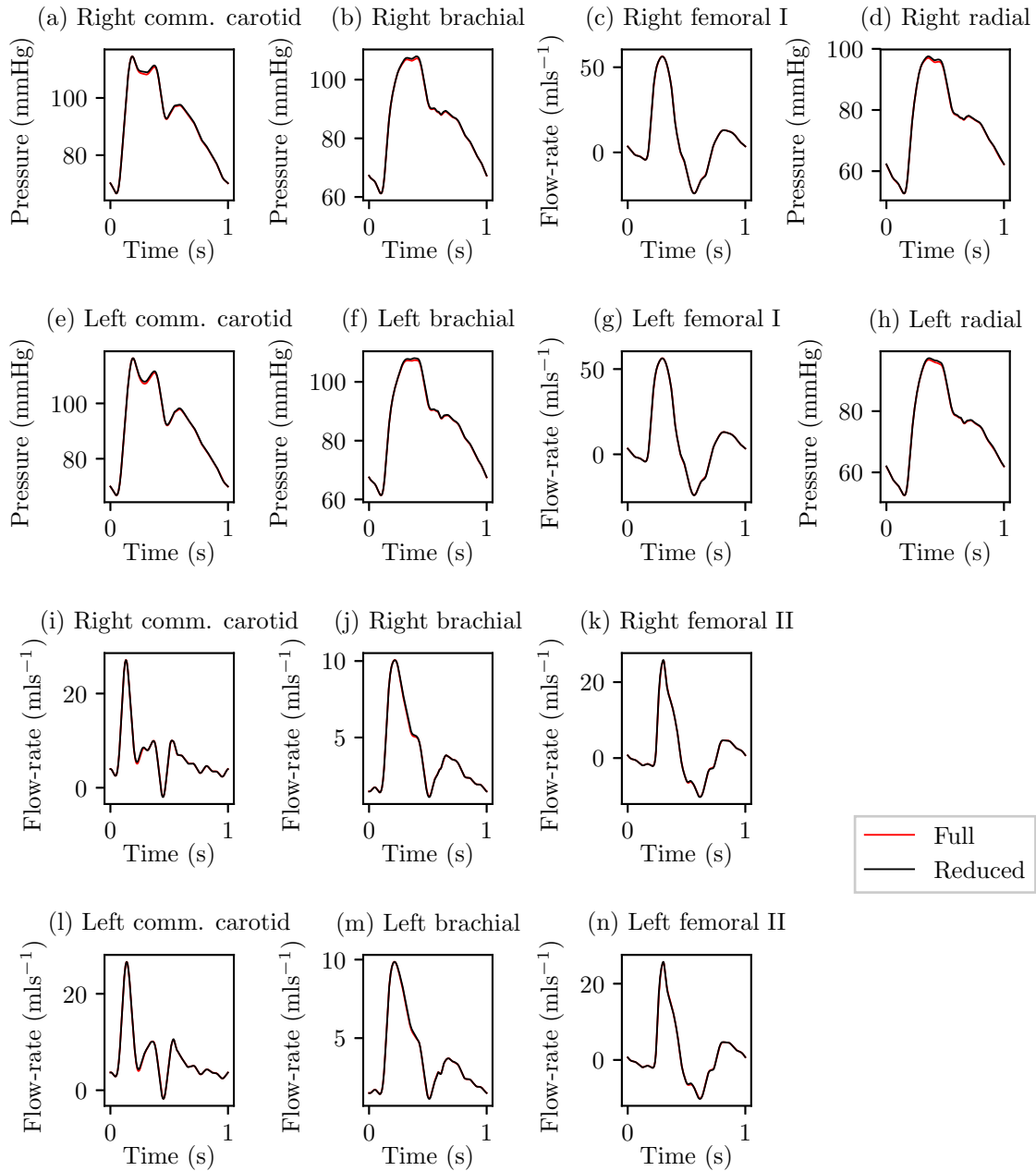


Figure 4.6: Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the first and second splenic segments; the left gastric; and the common hepatic arteries.

4.4. ARTERIAL NETWORK TOPOLOGY

the maximum pressure error is occurring in the radial arteries, with the right and left artery experiencing systolic errors of 4.847% and 3.981%, respectively. The maximum flow-rate error is occurring in the brachial arteries, with systolic errors of 16.158% and 16.031% in the right and left artery, respectively.

As the errors induced by the removal of the full second group of vessels exceeds the maximum allowable limit of 2%, the second ulnar segment and the common interosseous are re-introduced into the arterial network. The percentage discrepancies between the initial network and the reduced network produced by the removal of the posterior interosseous only are shown in Figure 4.8 and Table 4.7. Table 4.7 shows that, as is the case when removing the full group of vessels, the maximum pressure errors are occurring in the radial arteries and the maximum flow-rate errors are occurring in the brachial arteries. The re-introduction of the second ulnar segment and the common interosseous have reduced these errors, however. The maximum pressure errors are now equal to 0.678% in the right radial artery and 0.684% in the left radial artery. The maximum flow-rate errors are equal to 1.119% in the right brachial artery and 1.096% in the left brachial artery. As all errors are now below 2%, the posterior interosseous is removed from the arterial network of VPs. This reduction compacts one arterial vessel with constant properties and one Windkessel model into a single Windkessel model within each arm, removing a total of two arterial vessel segments.

The third group of vessels to be trialled for removal consists of the second popliteal segment, the anterior tibial, the tibiofibular trunk, and the posterior tibial. These vessels form a sub-network at the peripheries of the right and left legs, connected across one bifurcation and resulting in two terminal boundaries in each leg. As with the removal of peripheral vessels in the arms, it is expected that measurement locations directly in series and spatially nearby to this group of vessels will experience the greatest discrepancies when they are removed. Large discrepancies are, therefore, expected in the right and left femoral arteries. As the average flow-rate within the third group of vessels is greater than the average flow-rate in the peripheral arteries in the arms, the discrepancies experienced at measurement locations further from the reduction location are expected to be larger than those seen in Figure 4.7. Figure 4.9 shows the largest discrepancies are occurring in the femoral arteries, as is expected. Also as is expected, larger errors can be seen to be occurring at all other measurement locations, relative to those

4.4. ARTERIAL NETWORK TOPOLOGY

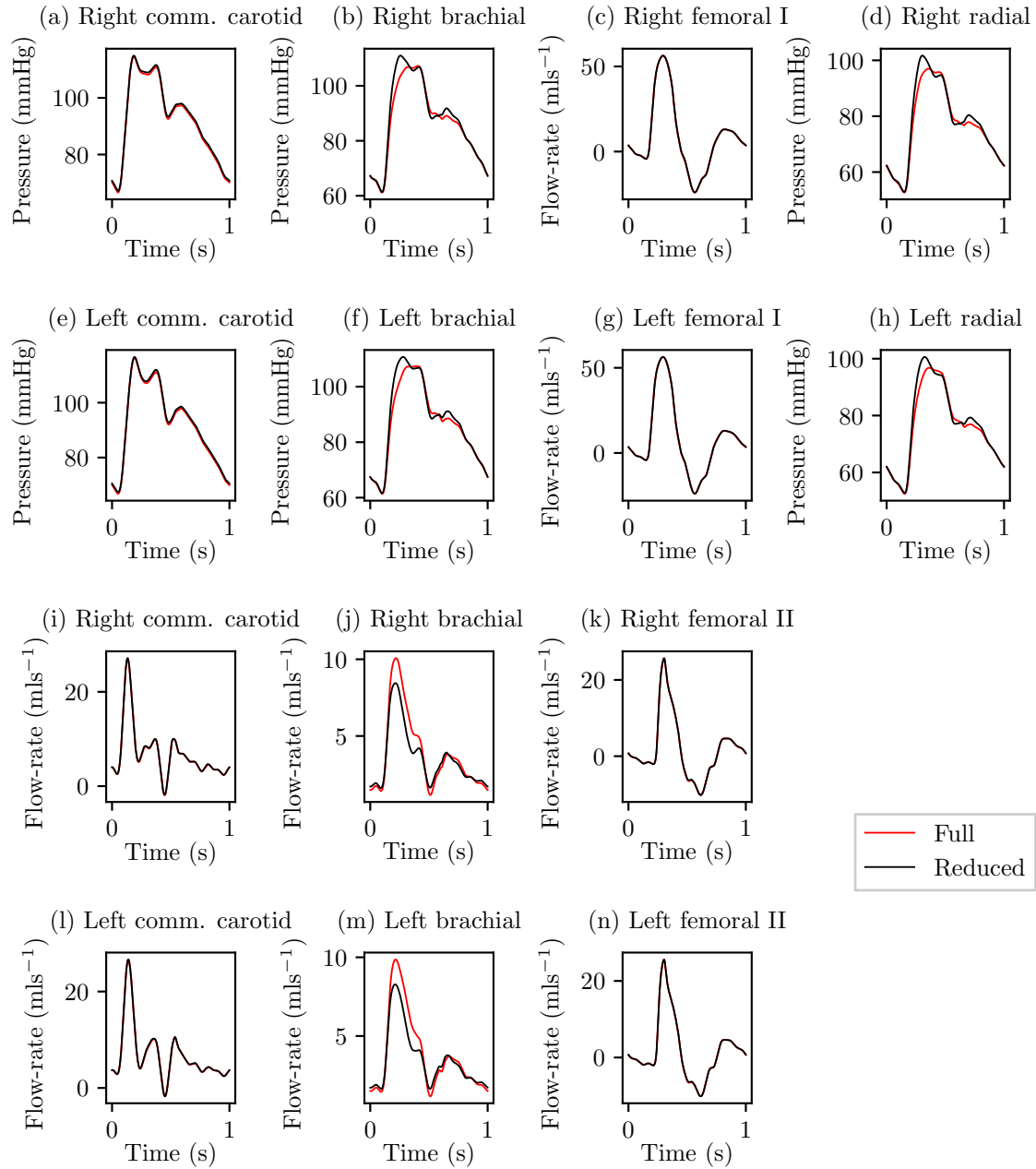


Figure 4.7: Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the common interosseous, the posterior interosseous, and the second ulnar segment.

4.4. ARTERIAL NETWORK TOPOLOGY

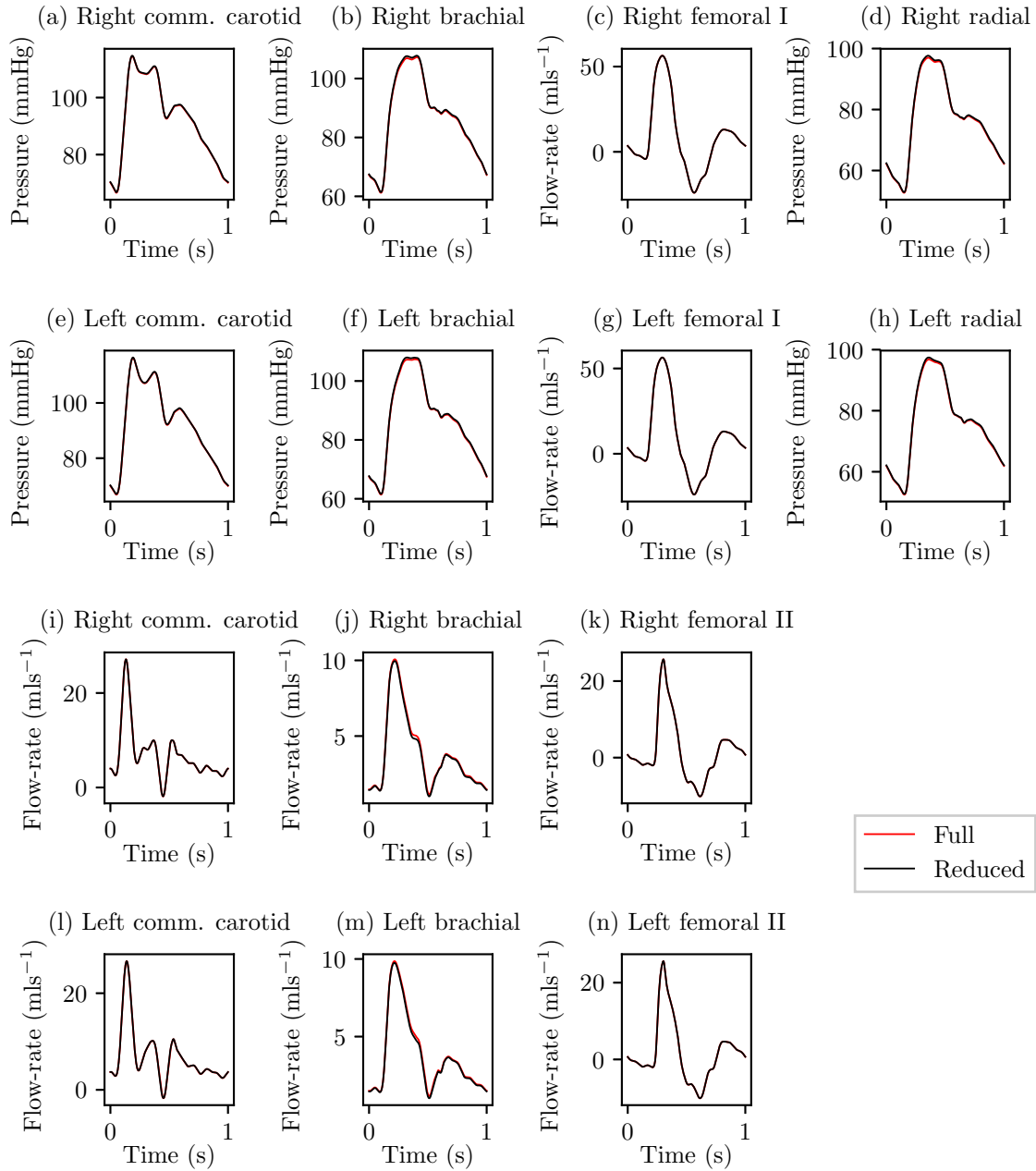


Figure 4.8: Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the the posterior interosseous.

4.4. ARTERIAL NETWORK TOPOLOGY

seen when removing either of the other two groups of vessels. Table 4.8 shows that the maximum flow-rate error is the average error in the second femoral segment. The average flow-rate error in the right and left second femoral segment is equal to 11.015% and 11.081%, respectively. The maximum pressure error is the average error in the right and left radial arteries, equal to 1.595% and 1.533%, respectively.

As the errors induced by the removal of the third group of vessels are greater than the allowable limit of 2% the second popliteal segment, the anterior tibial, and the tibiofibular trunk are re-introduced into the third proposed network. The pressure and flow-rate profiles corresponding to the reduced network produced by the removal of the posterior tibial only are shown in Figure 4.10. It can be seen from this figure that the re-introduction of the second popliteal segment, the anterior tibial, and the tibiofibular trunk have significantly reduced the discrepancies seen throughout the arterial network. Table 4.9 shows that the largest flow-rate errors are still occurring in the second femoral segments. The systolic error in the right and left second femoral segments are equal to 6.072% and 6.109%, respectively. The maximum pressure error remains in the radial arteries, with systolic errors in the right and left arteries of 0.776% and 0.750%, respectively. As the removal of even a single vessel within the peripheries of the legs results in flow-rate errors of more than 2%, it is not possible to remove any of these vessels from the arterial network of VPs.

Currently the affect of the removal of the first and second splenic segments; the left gastric; and the common hepatic and the removal of the right and left posterior interosseous have been examined independently. Before finalising the removal of these vessels from the arterial network of VPs, it is important to make sure that there is no significant amplifying affect on the errors induced when all vessels are removed in combination. The pressure and flow-rate profiles corresponding to the reduced arterial network produced by the removal of the first and second splenic segments; the left gastric; the common hepatic; and the posterior interosseous are shown in Figure 4.11. Figure 4.11 shows there are no significant errors induced when all 6 arterial vessels are removed in combination. Table 4.10 shows that, as is expected, greater errors are induced when both groups of vessels are removed in combination rather than isolation. These errors remain within the 2% allowable range, however. The maximum pressure error is the systolic error in the right and left radial arteries, equal to 1.236% and 1.231%, respectively. The maximum flow-

4.4. ARTERIAL NETWORK TOPOLOGY

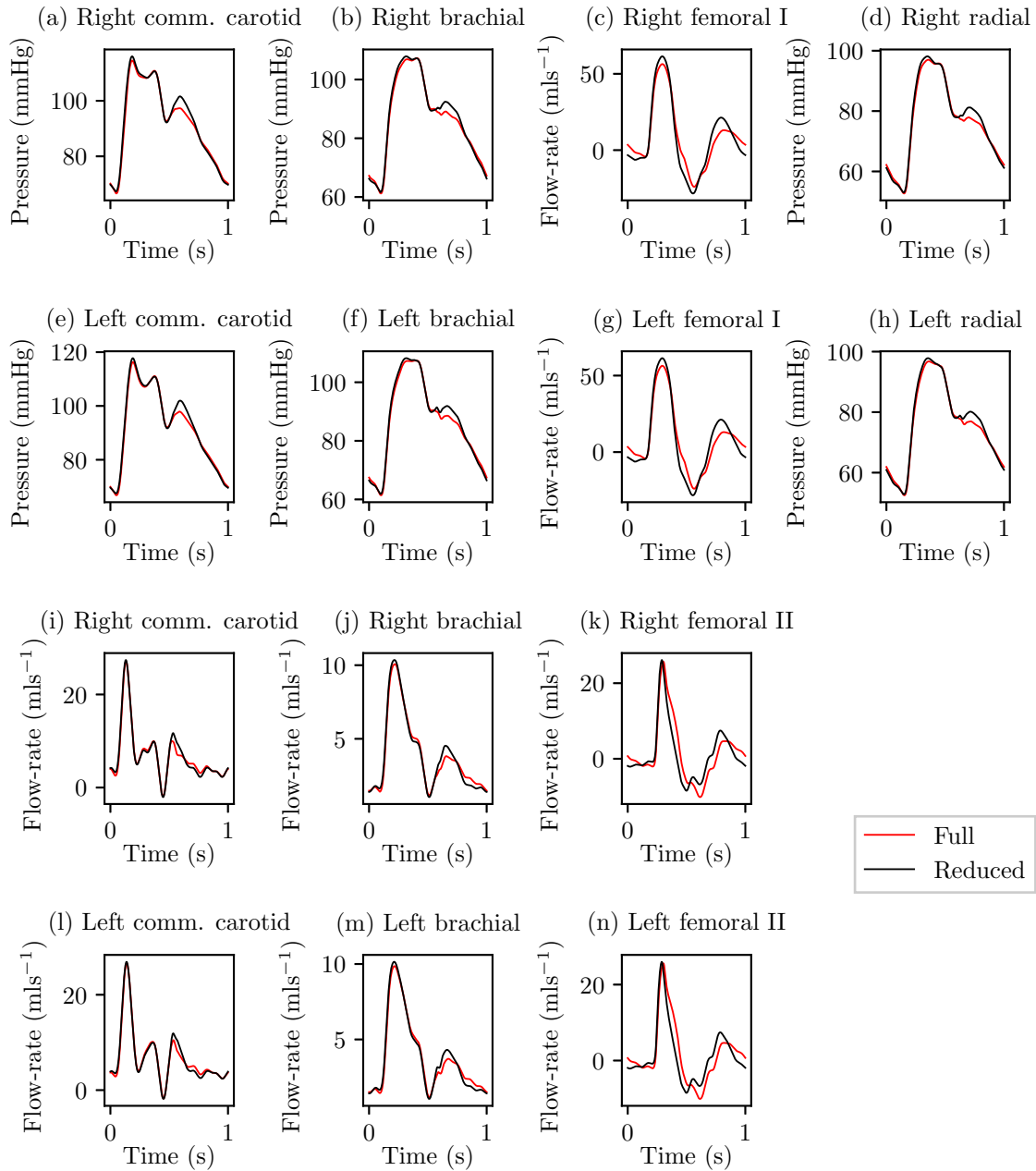


Figure 4.9: Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the the first and second popliteal segments; the anterior tibial; the tibiofibular trunk; and the posterior tibial.

4.4. ARTERIAL NETWORK TOPOLOGY

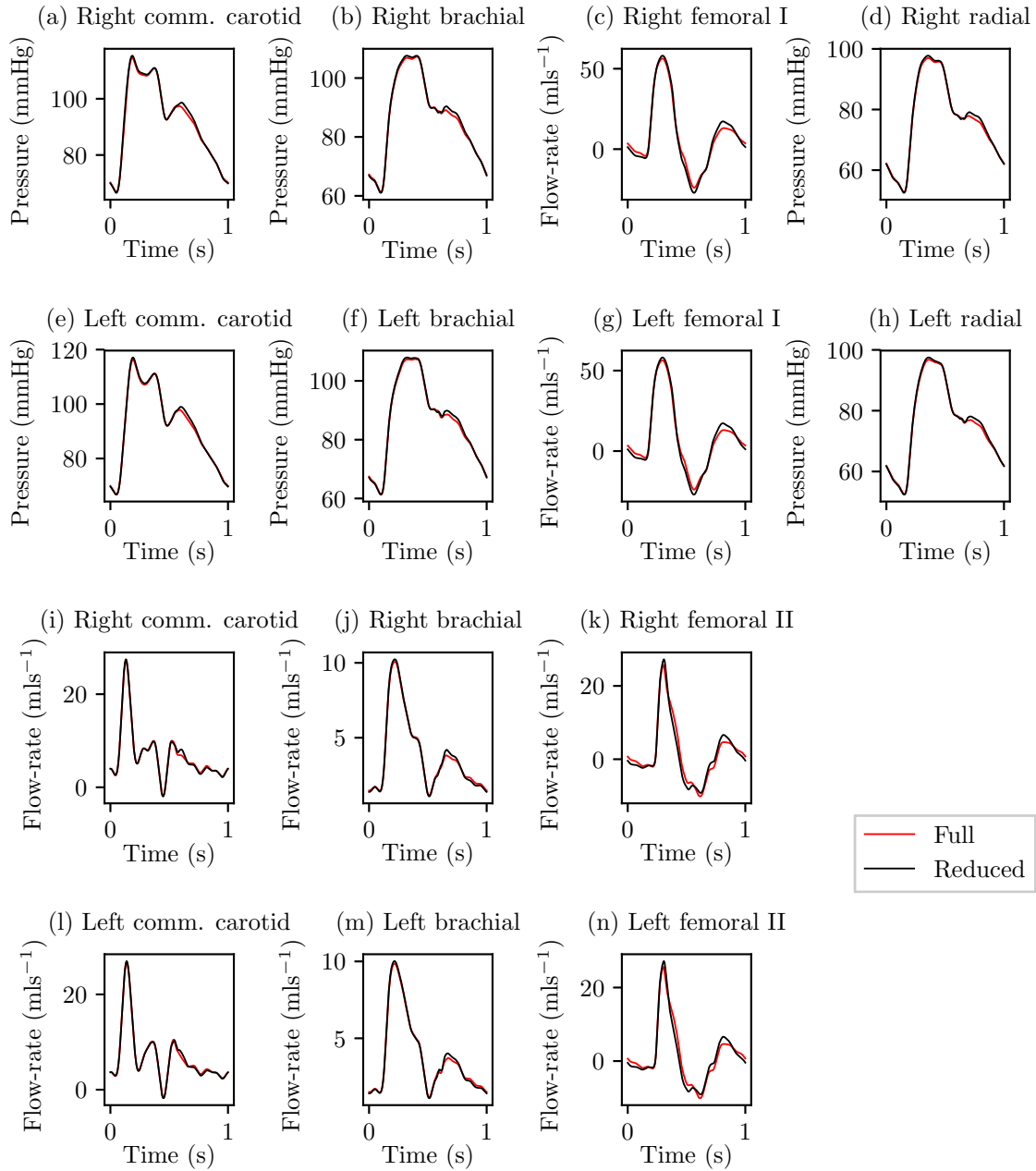


Figure 4.10: Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the posterior tibial.

rate error is occurring in the brachial arteries, with a systolic flow-rate error in the right and left brachial arteries of 1.223% and 1.131%, respectively.

It is found that the first and second splenic segments; left gastric; common hepatic; and the left and right posterior interosseous vessels can all be removed from the arterial network without introducing errors larger than 2%, relative to the full initial network. These six vessels are highlighted within in Figure 4.12. VPs are, therefore, created with these six vessels omitted from their arterial networks. All of the removed segments have constant reference radii and vessel wall mechanical properties along their lengths. The removal of these segments results in the need for two fewer Windkessel models. Thus, the removal of these six vessels reduces the dimensionality of VPs arterial networks by a total of 24—assuming the parameterisation used within the initial network. This network and its associated parameters—either directly taken from or computed based off of the initial network—forms the reference network for this study, and has a dimensionality of 389. A parameterisation of the arterial networks of VPs must be decided upon to both reduce this dimensionality further and enforce required physiological and geometrical behaviours, outlined next.

4.5 Parameterisation of the arterial network

A parsimonious representation is sought for the aforementioned reference (reduced) network so that its dimensionality can be further reduced. When deciding upon the parameterisation of the arterial network of VPs, two pairs of opposing characteristics must be considered and balanced:

- **Physiological precision versus dimensionality.** Requiring a high number of parameters to describe the arterial networks of VPs results in the creation of the new VPD being complex and likely computationally expensive. Excessive simplification of VPs, on the other hand, results in the pressure and flow-rate profiles produced being unrepresentative of the profiles seen throughout a real arterial network.
- **Variability versus control.** Strict relationships and behaviours can be imposed on the arterial networks of VPs, to ensure highly physiologically likely VPs are produced. This, however, comes at the cost of the variability of arterial

4.5. PARAMETERISATION OF THE ARTERIAL NETWORK

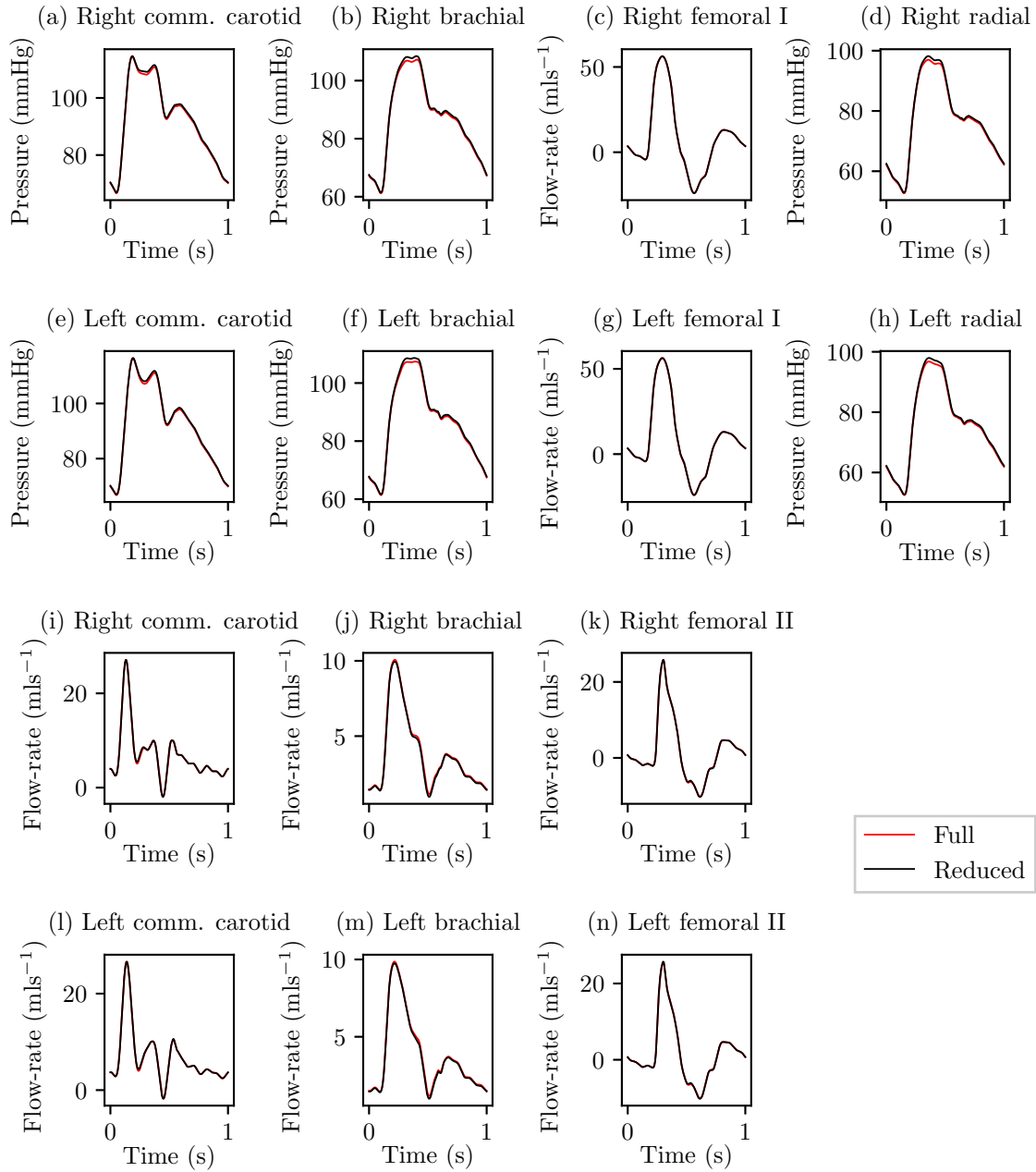


Figure 4.11: Comparison of pressure or flow-rate profiles at all measurement locations taken from the initial network and the reduced network produced by the removal of the first and second splenic segments, left gastric, common hepatic, and posterior interosseous. Percentage errors have been computed based on error metric presented in [58].

Location	Measurement	Average error (%)	Systolic error (%)	Diastolic error (%)
Right common carotid	Pressure	0.224	-0.017	0.009
	flow-rate	0.237	-0.074	-0.340
Left common carotid	Pressure	0.224	0.028	0.012
	flow-rate	0.232	-0.104	-0.258
Right brachial	Pressure	0.229	0.636	0.001
	flow-rate	0.285	-0.103	-0.474
Left brachial	Pressure	0.228	0.642	0.004
	flow-rate	0.275	-0.034	-0.368
Right radial	Pressure	0.249	0.554	-0.000
	flow-rate	0.249	0.546	0.003
Left radial	Pressure	0.256	-0.198	-0.181
	flow-rate	0.356	0.520	-0.483
Right femoral first segment	Pressure	0.254	-0.190	-0.181
Right femoral second segment	flow-rate	0.354	0.515	-0.483
Left femoral first segment	Pressure			
Left femoral second segment	flow-rate			

Table 4.5: Percentage discrepancy between non-invasive measurements taken from full arterial network and the reduced network produced by the removal of the first and second splenic segments; the left gastric; and the common hepatic arteries. Percentage errors have been computed based on error metric presented in [58].

Location	Measurement	Average error (%)	Systolic error (%)	Diastolic error (%)
Right common carotid	Pressure	0.708	0.333	0.834
	flow-rate	0.190	-0.016	0.147
Left common carotid	Pressure	0.708	0.327	0.839
	flow-rate	0.224	-0.053	0.06
Right brachial	Pressure	1.831	3.471	0.296
	flow-rate	5.275	-16.158	4.541
Left brachial	Pressure	1.833	3.127	0.334
	flow-rate	5.322	-16.031	4.713
Right radial	Pressure	1.911	4.847	0.395
	Pressure	1.911	3.981	0.419
Right femoral first segment	flow-rate	0.269	-0.073	0.064
Right femoral second segment	flow-rate	0.223	-0.122	-0.179
Left femoral first segment	flow-rate	0.268	-0.071	0.061
Left femoral second segment	flow-rate	0.223	-0.123	-0.179

Table 4.6: Percentage discrepancies between the non-invasive measurements taken from the full arterial network and the reduced network produced by the removal of the common interosseous, the posterior interosseous, and the second ulnar segment. Percentage errors have been computed using the error metrics presented in [58].

Location	Measurement	Average error (%)	Systolic error (%)	Diastolic error (%)
Right common carotid	Pressure	0.215	0.130	0.304
	flow-rate	0.055	-0.009	0.066
Left common carotid	Pressure	0.215	0.126	0.303
	flow-rate	0.055	-0.005	0.068
Right brachial	Pressure	0.417	0.458	0.375
	flow-rate	0.975	-1.119	-1.341
Left brachial	Pressure	0.418	0.568	0.374
	flow-rate	0.992	-1.096	-1.402
Right radial	Pressure	0.410	0.678	0.382
	Pressure	0.412	0.684	0.382
Left radial	Pressure	0.412	0.684	0.382
	flow-rate	0.0354	0.023	0.067
Right femoral first segment	flow-rate	0.023	0.0	0.018
Right femoral second segment	flow-rate	0.035	0.023	0.066
Left femoral first segment	flow-rate	0.023	0.023	0.066
Left femoral second segment	flow-rate	0.023	-0.002	0.0192

Table 4.7: Percentage discrepancies between the non-invasive measurements taken from the full arterial network and the reduced network produced by the removal of the posterior interosseous. Percentage errors have been computed based using the error metrics presented in [58].

Location	Measurement	Average error (%)	Systolic error (%)	Diastolic error (%)
Right common carotid	Pressure	1.245	1.295	1.117
	flow-rate	1.775	1.078	-0.581
Left common carotid	Pressure	1.226	1.24	1.054
	flow-rate	1.796	1.045	-0.405
Right brachial	Pressure	1.420	0.576	0.758
	flow-rate	2.704	2.802	-1.041
Left brachial	Pressure	1.369	0.779	0.744
	flow-rate	2.664	2.904	-1.119
Right radial	Pressure	1.595	1.165	0.622
	flow-rate	1.533	1.106	0.609
Right femoral first segment	flow-rate	9.477	8.963	-7.750
Right femoral second segment	flow-rate	11.015	1.764	6.553
Left femoral first segment	flow-rate	9.451	8.83	-7.598
Left femoral second segment	flow-rate	11.081	1.547	6.210

Table 4.8: Percentage discrepancy between the non-invasive measurements taken from the full arterial network and the reduced network produced by the removal of the second popliteal segment, the anterior tibial, the tibiofibular trunk, and the posterior tibial. Percentage errors have been computed using the error metrics presented in [58].

Location	Measurement	Average error (%)	Systolic error (%)	Diastolic error (%)
Right common carotid	Pressure	0.514	0.691	-0.101
	flow-rate	0.790	1.369	-0.271
Left common carotid	Pressure	0.505	0.687	-0.078
	flow-rate	0.774	1.301	-0.208
Right brachial	Pressure	0.582	0.312	-0.271
	flow-rate	1.098	1.643	-0.641
Left brachial	Pressure	0.568	0.523	-0.149
	flow-rate	1.085	1.611	-0.657
Right radial	Pressure	0.650	0.776	-0.321
	Pressure	0.632	0.750	-0.216
Right femoral first segment	flow-rate	3.837	3.068	-5.544
Right femoral second segment	flow-rate	4.762	6.072	3.718
Left femoral first segment	flow-rate	3.834	3.077	-5.520
Left femoral second segment	flow-rate	4.794	6.109	3.950

Table 4.9: Percentage discrepancy between the non-invasive measurements taken from the full arterial network and the reduced network produced by the removal of posterior tibial. Percentage errors have been computed using the error metrics presented in [58].

Location	Measurement	Average error (%)	Systolic error (%)	Diastolic error (%)
Right common carotid	Pressure	0.4380	0.114	0.316
	flow-rate	0.250	-0.085	-0.271
Left common carotid	Pressure	0.438	0.157	0.318
	flow-rate	0.246	-0.111	-0.186
Right brachial	Pressure	0.643	1.108	0.378
	flow-rate	0.875	-1.223	-1.833
Left brachial	Pressure	0.644	1.179	0.380
	flow-rate	0.889	-1.131	-1.788
Right radial	Pressure	0.654	1.236	0.383
	Pressure	0.657	1.231	0.386
Right femoral first segment	flow-rate	0.256	-0.174	-0.114
Right femoral second segment	flow-rate	0.365	0.515	-0.463
Left femoral first segment	flow-rate	0.254	-0.166	-0.115
Left femoral second segment	flow-rate	0.363	0.514	-0.463

Table 4.10: Percentage discrepancy between the non-invasive measurements taken from the full arterial network and the reduced network produced by the removal of the first and second splenic segments; left gastric; common hepatic; and posterior interosseous. Percentage errors have been computed based using the error metrics presented in [58].

4.5. PARAMETERISATION OF THE ARTERIAL NETWORK

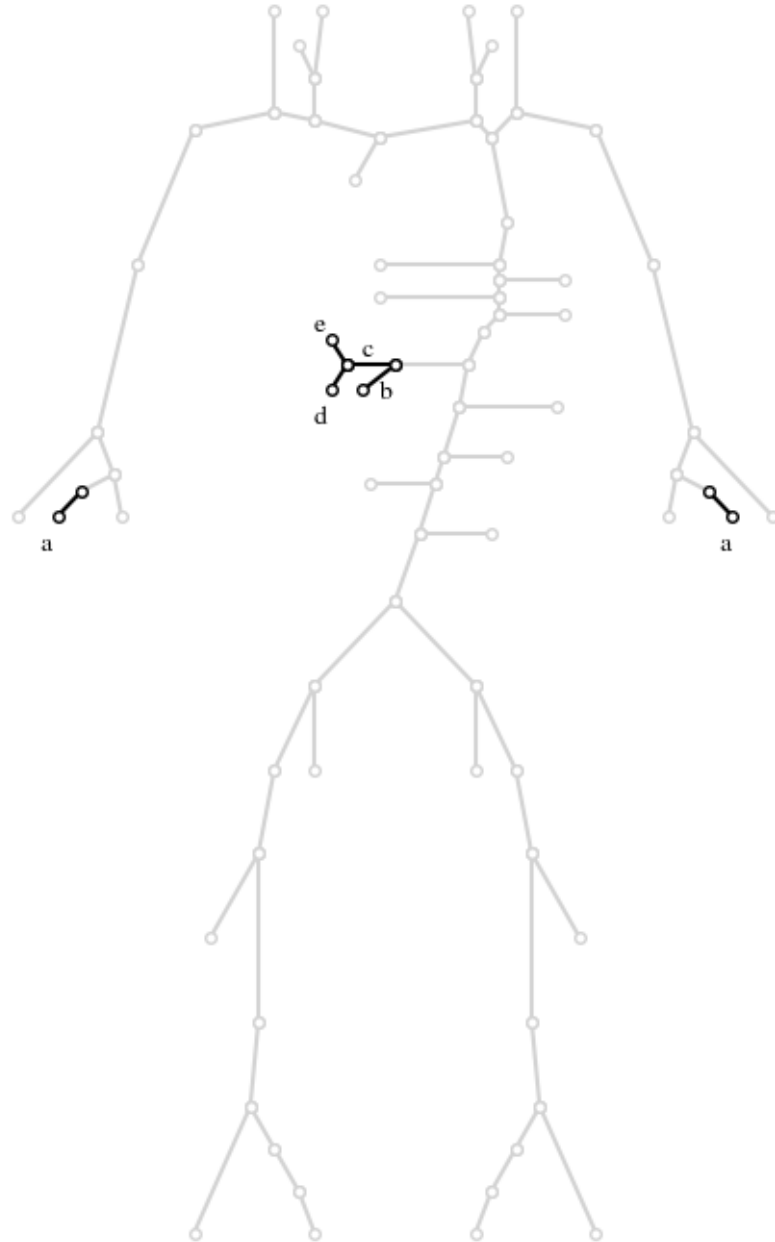


Figure 4.12: The location of the vessels that have been found to be able to be removed from the initial network without inducing errors of greater than 2%. In the above figure 'a', 'b', 'c', 'd', and 'e' indicate the second ulnar segments, the common hepatic artery, the first splenic segment, the left gastric artery, and the second splenic segment respectively.

networks, and so consequently pressure and flow-rate profiles, seen across the final VPD. More variability can be produced across the VPD by apply a non-restrictive parameterisation to the arterial network. This is likely to result in an increase in the frequency of the occurrence of physiologically unlikely VPs.

The vessels in the reference network can be split into the following three categories, which are also indicated in Figure 4.13:

- **Category 1:** 33 vessel segments with varying β and r_0 , and continuous variation with respect to the prior and subsequent vessels.
- **Category 2:** six vessel segments with varying β and r_0 , and discontinuous variation with respect to the prior and subsequent vessels.
- **Category 3:** 32 vessels where β and r_0 are constant along their lengths.

When solving the physics based model of pulse wave propagation, vessel segments are split into a series of elements. The radii and vessel wall mechanical properties of the nodes at the proximal and distal ends of each element are computed based on a given profile of the vessel properties. The SDC implementation being used to solve the physics based model of pulse-wave propagation assumes that the properties of elements vary linearly along their lengths. It can, therefore, be seen that vessel properties are subject to two profiles; an intra-vessel profile dictating the properties at each nodal position, and an intra-elemental profile dictating the profile of properties between these nodes. The intra-elemental profile is forced to be linear by the SDC implementation being used, while the intra-vessel profile is decided upon through the parameterisation of the arterial network chosen. Given a fine enough spatial discretisation, *i.e.* a small enough element length, the successive intra-elemental property profiles should accurately approximate the desired intra-vessel profile. For Categories 1 and 2 with varying intra-vessel property profiles, the reference network assumes a linear variation [20]. Thus, any property $g(x)$ (reference radius r_0 or β) varies along length x as:

$$g(x) = g_0 - (g_0 - g_L) \frac{x}{L}, \quad (4.4)$$

where g_0 , and g_L represents the properties at the proximal end $x = 0$ and the distal end $x = L$, respectively. A re-parameterisation of these properties to reduce the

total dimensionality is presented next. For all the aforementioned three categories of vessels, a unified parameterisation is proposed, which also results in reduction in the dimensionality associated with the description of Category 1 vessels. This parameterisation includes description of the mechanical properties (Section 4.5.1), geometric properties (Section 4.5.2) and boundary conditions (Section 4.5.3).

4.5.1 Mechanical properties.

Exponential property profiles

As opposed to describing the property variations linearly, see Equation (4.4), the key idea is to use exponential variations. The re-parameterisation of mechanical properties β for the three categories is as follows:

Category 1: These vessels have continuous variation in properties between successive vessels. Thus, successive vessels can be lumped together into chains and a single parameterisation adopted. The 33 vessel segments (Category 1) that meet this description can be lumped into the following chains:

- **The aortic chain.** This chain of vessel segments includes the first to forth aortic arch segments; the first to sixth thoracic aorta segments; and the first to fifth abdominal aorta segments.
- **The right and left arm chains.** These chains of vessels include the first and second subclavian segments; the axillary artery; and the brachial artery.
- **The right and left leg chains.** These chains of vessels include the external iliac; the first and second femoral segments; and the first and second popliteal segments.

The β profiles for an entire vessel chain can now be described using a single function as opposed to separate functions (and hence separate parameters) for each individual vessel.

When vessel segments are lumped into chains, as is being done within this study, linear intra-chain property profiles are problematic. Proximal vessels within the network are expected to have steeper property profile gradients, while distal vessels are expected to have shallower property profile gradients. Describing the property profiles of vessel chains through a linear variation of the properties at the proximal

end of the chain results in a constant gradient along the entire length of the chain. This consequently results in too little decay of vessel properties in segments at the proximal end of the chain, and too great decay of vessel properties in segments at the distal end of the chain. An appropriate choice for intra-chain variation is an exponential decay of the properties at the proximal end of the chain:

$$\beta(x) = \beta_0 \exp(-\Omega_v x), \quad (4.5)$$

where $\beta(x)$ and β_0 represents the mechanical properties of the chain at the spatial position x and the proximal end of the chain, respectively; and $\Omega_v > 0$ represent a decay parameter.

Category 2: These vessels do not form part of any chain but show an intra-vessel property variation. For consistency and uniformity, their variation is also described by the exponential re-parameterisation of Equation (4.5).

Category 3: These vessels are also not part of a chain and have constant intra-vessel properties. Again, for uniformity, the re-parameterisation of Equation (4.5) is used, with the added explicit condition that $\Omega_v = 0$.

Hierarchical assignment of vessel properties

The assignment of β properties of the entire network, as described by Equation (4.5), is presented here. Proximal-distal coherence dictates that the β properties at the distal end of a vessel segment must be greater than or equal to the corresponding property at the proximal end of any subsequent daughter segments [17]. To ensure this a hierarchical procedure is adopted.

The β properties of the arterial network are initialised by explicitly assigning a value at the inlet of the first aortic arch segment. The property profile of the aortic chain, and so consequently the vessel wall mechanical property at the proximal and distal ends of each segment within the chain, is computed using Equation (4.5). The properties at the proximal end of any vessel segments branching from the aorta are computed by applying a scaling term (introduced to ensure proximal-distal coherence) to the corresponding property at the distal end of the parent aortic segment:

$$\beta_0^d = \beta_{L_p}^p \Omega_s, \quad (4.6)$$

4.5. PARAMETERISATION OF THE ARTERIAL NETWORK

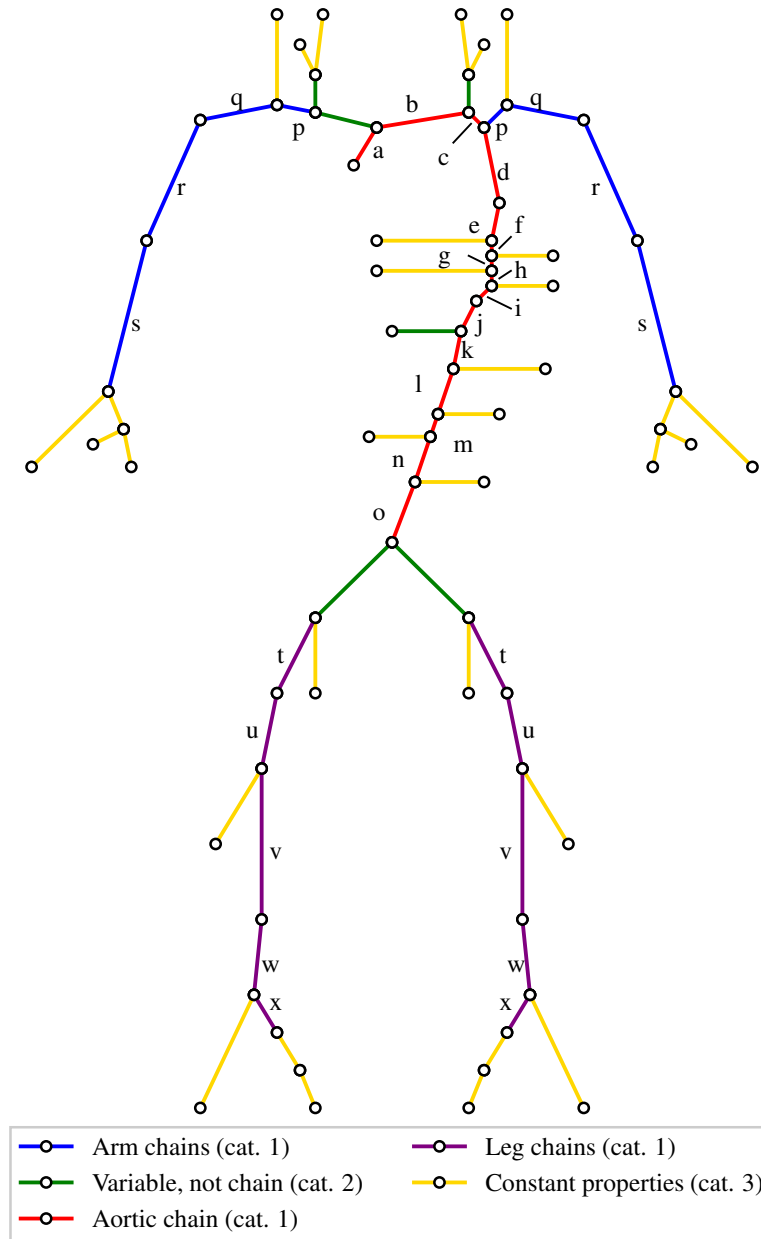


Figure 4.13: The location of the vessels within each of the five chains. The aortic chain consists of: the first to fourth aortic arch segments denoted by “a” through to “d”, the first to sixth thoracic aorta segments denoted by “e” through to “j”, and the first to fifth abdominal aorta segments denoted by “k” through to “o”. The right and left arm chains consist of: the first and second subclavian segments denoted by “p” and “q” respectively, the axillary artery denoted by “r”, and the brachial artery denoted by “s”. The right and left leg chains consist of: the external iliac denoted by “t”, the first and second femoral segments denoted by “u” and “v” respectively, and the first and second popliteal segments denoted by “w” and “x” respectively.

4.5. PARAMETERISATION OF THE ARTERIAL NETWORK

where β_0^d represents the property at the proximal end of the daughter segment, $\beta_{L_p}^p$ represents the property at the distal end of the parent segment, and Ω_s represents a daughter-parent scaling term, where:

$$0 < \Omega_s \leq 1. \quad (4.7)$$

The intra-vessel or intra-chain property profiles of all daughter branches bifurcating from the aorta are computed using Equation (4.5). This process is sequentially repeated through each generation of the arterial network until all terminal boundaries are reached.

The β property profile of the aortic chain is fully described by Equation (4.5), as β_0 is overtly stated to initialise the network. To fully understand the parameterisation of vessel segment property profiles beyond the aortic chain Equations (4.5) and (4.6) must be synthesised, *i.e.*:

$$\beta^d(x_d) = \beta_{L_p}^p \Omega_s \exp(-\Omega_v x_d), \quad (4.8)$$

where $\beta^d(x_d)$ represents the β property of the daughter vessel at position x_d . The number of parameters required to describe the intra-chain property profiles and the intra-vessel property profiles of vessels with varying properties along their length, beyond the aortic chain, can be seen to remain as two. The property at the distal end of the parent vessel, denoted by $\beta_{L_p}^p$, is described by the parameters of the parent intra-chain or intra-vessel property profile.

There are two arterial vessel segments with a daughter-parent ratio of one, each with a right and left instance. These vessels are the second ulnar segment, and the tibiofibular trunk. The daughter-parent ratio for these four vessels is subject to the explicit condition $\Omega_s = 0$, removing the need for four dimensions. The β properties of all arterial vessel segments, therefore, requires provision of 50 parameters.

Figure 4.14 shows an example of the assignment of vessel properties across two generations of the arterial network. From this figure, the formation of a Bayesian network can be seen [37]. The parameters of an intra-vessel or intra-chain property profile are independent of the parameters of the parent intra-vessel or intra-chain property profile. If the intra-vessel or intra-chain property profile of a vessel of interest is known, however, a conditional dependency is formed between

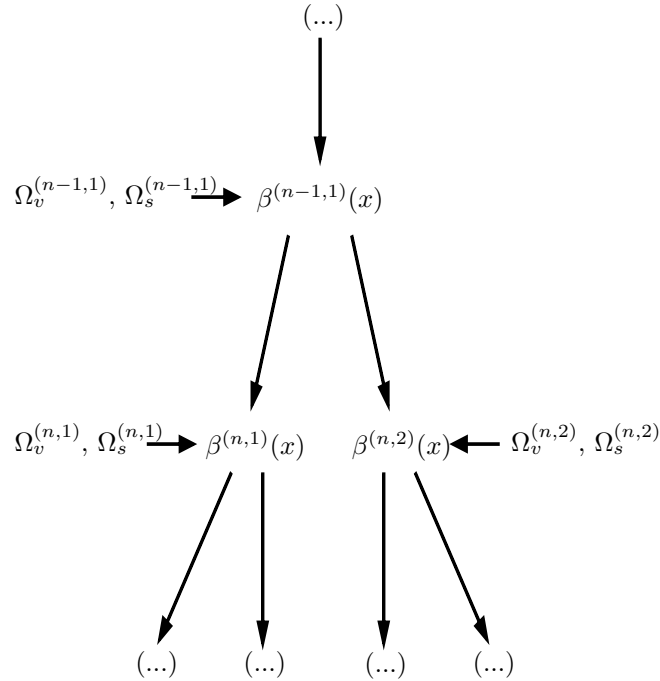


Figure 4.14: Illustration of hierarchical procedure used to assign intra-vessel and intra-chain property profiles to segments within the arterial network. $\Omega_v^{(n,m)}$ and $\Omega_s^{(n,m)}$ represent the decay and scaling parameters of the m^{th} branch at the n^{th} generation respectively; and $\beta^{(n,m)}(x)$ represents the vessel wall mechanical property or radius of the m^{th} branch at the n^{th} bifurcation generation at location x

the parameters of the profile and the parameters of property profile of the parent.

4.5.2 Geometric properties

Here, the re-parameterisation of geometric properties—reference radius and length—is presented.

Arterial vessel reference radius

Since the variation of reference radius (r_0) is subject to the same requirements as those for the variation of β , the variation of r_0 is described in exactly the same manner as β , as presented in 4.5.1. Lumping arterial vessels into chains, where appropriate, and applying exponential functions to describe the properties β and

r_0 of all the vessels reduced the number of dimensions from 389 in the reference network to 269.

Arterial vessel length

Each arterial vessel requires specification of its length. To individually assign a length to every vessel requires 71 parameters, accounting for a large proportion of the 269 remaining dimensions. It is thus proposed to reduce the dimensionality of the network by applying a single scaling term to the lengths of all vessels. It is empirically found that behaviours, patterns, and overall variability observed in the pressure and flow-rate profiles when allowing maximum freedom to the length of vessels, *i.e.* assigning independent lengths to each arterial vessel, are not lost when applying a singular vessel length scaling term to all the vessels. This analysis justifies the use of a single scaling term and can be found in Appendix G. With this assumption, the dimensionality is reduced from 269 to 199.

4.5.3 Boundary conditions

The boundary conditions consist of the inlet flow-rate and the terminal lumped parameters. These are parameterised as follows:

Inlet flow-rate

The volumetric inlet flow-rate to the network—at the free end of vessel ‘a’ in Figure 6.1—is described using a Fourier series (FS), as previously employed in the PoC study. It is found in Chapter 3 that a time domain inlet flow-rate profile can be described to a high level of precision using a FS truncated at the 5th order. Thus, the time domain inlet flow-rate profiles can be described by:

$$\mathbf{Q}_{\text{inlet}} = \{a_0 = 0, b_0, a_1, b_1, \dots, a_5, b_5\}, \quad (4.9)$$

and requires specification of 11 coefficients.

Terminal lumped models

There are 29 terminal boundaries in the network. Each of these terminal boundaries is coupled to a Windkessel model. Each Windkessel model requires three parameters; two resistances, and a compliance. It is chosen that the parameters of all Windkessel models will be treated as *a priori* independent—*i.e.* no direct inter- or intra-Windkessel model relationships are stated. These Windkessel models, therefore, require specification of a total of 87 parameters.

4.5.4 Final network summary

The final parsimonious representation consists of:

- **33 category 1 vessel segments**, for which the intra-vessel property profiles are described through the use of one of five vessel chains. Each chain of vessels requires two exponential functions—one describing the r_0 intra-chain profile, and one describing the β intra-chain profile. Each exponential function requires two parameters. In the case of the aortic chain these parameters are an initialising term and an exponential decay term. For the other four vessel chains the two required parameters are a daughter-parent ratio term and an exponential decay term. These 33 vessel segments, therefore, require a total of 20 dimensions to describe their properties.
- **Six category 2 vessel segments**, for which the intra-vessel property profiles are described by their own individual exponential function. Each of these segments requires two profiles described by two parameters. These six vessel segments require a total of 24 dimensions to describe their properties.
- **32 category 3 vessel segments**. 28 of these vessels have properties that differ from that of their parent vessel. The properties of these 28 vessels are, therefore, described using two parameters—a daughter-parent scaling term for their r_0 and β property, respectively. The remaining four vessel segments with constant properties along their length have r_0 and β properties equal to that of their parent. These four vessels require no parameters to be described, as their daughter-parent ratio is set to one. The 32 vessel segments with constant

properties along their lengths, therefore, require a total of 56 dimensions to describe their properties.

- **The inlet flow-rate**, described using a FS truncated at the 5th order (see Section 4.5.3).
- **29 Windkessel models**, requiring 87 parameters (see Section 4.5.3).
- **The lengths of all vessels** in the network are computed based on a single scaling term (see Section 4.5.2).

The above description has 199 parameters. These represent 22 decay parameters Ω_v , 78 daughter-parent scaling parameters Ω_s , 87 Windkessel parameters at the outlets, 11 FS coefficients describing the inlet flow-rate, and 1 length scaling term. These 199 parameters represent a reduction of approximately 48% with respect to the original description with 389 parameters. The final step in parameterisation is to describe all these 199 parameters relative to the reference network, and is described next.

4.5.5 Scaling with respect to the reference network

Instead of describing the network by directly stating values of the 199 parameters, they are specified relative to the reference network:

$$\mathbf{\Omega} = \mathbf{\Phi}^T \mathbf{\Omega}_{\text{ref}}, \quad (4.10)$$

where $\mathbf{\Phi} = [\Phi_1, \Phi_2 \cdots \Phi_{199}]^T$ represents the scaling vector in relation to the reference network, and $\mathbf{\Omega}_{\text{ref}}$ represents the vector of reference network parameters. The reference network parameters for the vessel wall mechanical property and radius exponential profiles (see Sections 4.5.1 and 4.5.2) are computed by curving fitting each to the reference (reduced) network presented in Section 4.4.3. The pressure and flow-rate profiles corresponding to the reference network presented in Section 4.4.3 and those produced when using the reference network parameters (employing the aforementioned parsimonious parameterisation) are not compared, as the purpose of the reference network parameters is not to perfectly recreate the ADAN network but instead act as a known physiologically highly likely example of the parsimonious parameterisation. By applying scaling terms to a reference set of parameters, easy

comparison can be made to a known reference patient. Random sampling from the distribution of Φ will result in the VPD. This is described in the next section.

4.6 Statistical modelling

In Sections 4.4 and 4.5 a reduced version of the initial network, consisting of 71 arterial vessel segments, and the parameterisation of this network, which is fully described by 199 parameter scaling terms, is outlined. Random realisations of these 199 arterial network parameter scaling terms are sampled to create the VPD. When creating the VPD it is important to ensure that both the arterial network parameter scaling terms assigned to VPs and the resulting pressure and flow-rate profiles capture the range and distribution measured in a real population. The information used to impose physiological realism on the VPD can be split into two general categories:

- **Literature reported measurements:** The literature reported measurements used to impose realism on the VPD are primarily based on pressure and flow-rate profiles. This is due to the fact that pressure and flow-rate measurements can be taken at certain locations within the arterial network relatively inexpensively and non-invasively, and so are frequently reported in literature. Parameters required to describe VPs arterial networks, on the other hand, are generally either non-physical, and so can not be directly measured, or are invasive to obtain. As a consequence of this there are significantly fewer cases of reported one-dimensional arterial network parameters. The distributions of the literature based measurements being incorporated into the VPD, to ensure physiological realism, is described as:

$$\mathcal{N}(\boldsymbol{\mu}_{\text{lit}}, \boldsymbol{\Sigma}_{\text{lit}}), \quad (4.11)$$

where $\boldsymbol{\mu}_{\text{lit}}$ is the vector of the mean measurements, and $\boldsymbol{\Sigma}_{\text{lit}}$ represents the measurement covariance matrix. Each literature based measurement is assumed to be independent and so $\boldsymbol{\Sigma}_{\text{lit}}$ is diagonal.

- **Known physiological restrictions and geometrical constraints:** The known physiological restrictions and geometrical constraints are used to impose

bounds on the supports of the distributions of arterial network parameter scaling terms.

To create the VPD, random samples of Φ from their joint probability distribution are required. This distribution should satisfy the two aforementioned pieces of information.

The pulse wave propagation model, denoted as \mathcal{M} , takes the parameters Φ as inputs and outputs the pressure and flow-rates, collectively denoted as the vector Λ , at all the locations:

$$\Lambda = \mathcal{M}(\Phi), \quad (4.12)$$

The specific measurements of the model—corresponding to the literature reported measurements—are denoted by τ , and are essentially transformations of Λ corrupted by measurement noise \mathcal{E} :

$$\tau = \mathcal{H}(\Lambda) + \mathcal{E}, \quad (4.13)$$

where \mathcal{H} represents the observation operator. Commonly, the observation operator \mathcal{H} represents an identity transform of selected components of Λ . The measurement error \mathcal{E} is typically assumed to be a zero-mean multivariate normal:

$$\mathcal{E} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\text{error}}), \quad (4.14)$$

where $\mathbf{0}$ is a zero-vector and Σ_{error} represents the error covariance matrix (diagonal when errors are independent).

Classical inversion problem

One method to create a distribution of arterial network parameter scaling terms is to treat the distribution of measurements taken from literature as targets that must be recreated across the final VPD, *i.e.*:

$$\tau \sim \mathcal{N}(\mu_{\text{lit}}, \Sigma_{\text{lit}}). \quad (4.15)$$

To find the distribution of arterial network parameter scaling terms required to recreate the literature based measurements across the VPD, the forward model described by Equations (4.12) and (4.13) is assumed to be fully deterministic, *i.e.* Σ_{error} is a zero-matrix. The required distribution of arterial network parameter

scaling terms is found through the inversion of Equations (4.12):

$$\Phi = \mathcal{M}^{-1}(\Lambda), \quad (4.16)$$

where $\mathcal{M}^{-1}(\Lambda)$ represents the inverted backward model mapping the distribution of model outputs to the distribution of arterial network parameter scaling terms. Two problems with this method for determining a distribution from which VPs must be sampled are:

- It is difficult to incorporate the known physiological restrictions and geometrical constraints.
- The inverse model \mathcal{M}^{-1} is difficult to be obtained. Even if a numerical approach was taken, the solution of the inverse problem is expensive.

Bayesian formulation

As opposed to considering the distribution of measurements taken from literature to be target distributions, they are instead treated as observations (measurements) of τ from Equations (4.12) and (4.13). A Bayesian formulation is adopted to find the posterior distribution of the parameters Φ , subject to the affect of the observations τ . This is suitable because the geometric and physiological constraints, in the form of a prior, can be combined with the literature reported measurements, in the form of a likelihood, to result in a posterior distribution that considers both the pieces of available information. The Bayes' theorem allows such a combination naturally:

$$\underbrace{p(\Phi | \tau)}_{\text{posterior}} = \frac{\overbrace{p(\tau | \Phi)}^{\text{likelihood}} \overbrace{p(\Phi)}^{\text{prior}}}{\underbrace{p(\tau)}_{\text{evidence}}}. \quad (4.17)$$

The prior distributions $p(\Phi)$ are described in section 4.6.1 and are constructed based on the known physiological restrictions and geometrical constraints. Strong prior beliefs are not held about the distributions of arterial network parameter scaling terms, and so these prior distributions are weakly informative or uninformative. Their primary purpose is to impose appropriate bounds on the supports of the distributions based on geometric and physiological constraints.

For the computation of the likelihood, in the context of the statistical model described by Equations (4.12)–(4.17), it is assumed that $\boldsymbol{\tau}$ is measured to be equal to $\boldsymbol{\tau} = \boldsymbol{\mu}_{\text{lit}}$ and the error covariance is $\boldsymbol{\Sigma}_{\text{error}} = \boldsymbol{\Sigma}_{\text{lit}}$. The likelihood for individual measurements is described in Section 4.6.2.

The evidence term within Bayes' theorem is independent of the distribution of arterial network parameter scaling terms. This means that the evidence does not affect the shape of the posterior distribution, and instead purely acts as a normalising constant, ensuring the integral of the posterior is equal to one. It is, therefore, possible to evaluate the posterior distribution, up to the limit of the normalising constant, without the need for computation of the evidence term.

4.6.1 Prior distributions

Known physiological restrictions and geometrical constraints are imposed on the posterior distribution of arterial network parameter scaling terms through the use of the prior distribution. When choosing appropriate prior distributions for each parameter it is important to consider the affect they have on not only the individual parameter, but also on the overall posterior distribution from which VPs are sampled. Prior distributions can be split into three categories, dependent on the strength of specific beliefs held about the parameter they are being applied to:

- **Informative prior distributions** are applied when strong specific beliefs are held about a parameter. An informative prior distribution has low variance, and so a high probability corresponding to its mean value. Due to the low variance, an informative prior significantly restricts the values a parameter can take, as any deviation from the mean results in a significant penalisation to the probability of that parameter. The high mean probability associated with an informative prior has a significant affect on the shape of the posterior distribution.
- **Weakly informative prior distributions** are applied when looser or partial beliefs are held about a parameter. A weakly informative prior has higher variance than an informative prior, constraining a parameter to a high probability region rather than a particular value. Weakly informative priors generally have lower mean probability than informative priors, and so have

less affect on the shape of the posterior distribution.

- **Uninformative priors** are the final common form of prior distribution. The name “uninformative” is misleading as these prior distributions do provide information about parameters. The information provided, however, is generally objective—for example the parameter must be within an allowable range. Uninformative priors are often based on the principle of indifference [143], assigning a constant probability to all values within a parameters support. Uninformative priors, therefore, provide little or no regularisation to a parameter within its support, and subsequently have little influence on the shape of the posterior distribution.

Since scarce information is available about the variations of the parameters, the prior distributions are primarily constructed based on the known physiological and geometrical bounds that should be observed. Thus, weakly informative or uninformative priors are applied to all arterial network parameter scaling terms. This allows the shape of the posterior distribution to be driven by the likelihood term within Bayes’ theorem, while the prior distribution acts purely as a regularisation term. One of three general distributions are used as the prior distribution of all parameter scaling terms, dependent on the support of that particular parameter:

1. **Bounded parameters:** A uniform prior distribution is chosen for all the scaling parameters which are supported on a bounded interval. The daughter-parent ratio (see Equations (4.6)) is the only example of such a parameter. For the i^{th} pair of vessels with reference parent-daughter ratio $\Omega_{i,\text{ref}}$, it can be seen that when the bounds imposed on the daughter-parent ratio in the parameter space (see Equation (4.7)) are mapped to the parameter scaling space (see Equation (4.10)) the corresponding scaling term Φ_i is bounded between 0 and $1/\Omega_{i,\text{ref}}$. The prior probability density function (PDF) is thus:

$$p(\Phi_i) = \begin{cases} \Omega_{i,\text{ref}} & \text{if } \Phi_i \in (0, 1/\Omega_{i,\text{ref}}] \\ 0 & \text{otherwise} \end{cases}, \quad (4.18)$$

and shown in the left plot of Figure 4.15.

2. **Semi-bounded parameters:** A log-normal prior distribution is chosen for

all the scaling parameters with semi-infinite support. These parameters with a lower bound of zero and no upper bound are:

- initial values of radii r_0 and vessel wall mechanical properties β at the inlet of the aorta (see Sections 4.5.1, and 4.5.2);
- the decay terms used within all exponential property profiles (see Sections 4.5.1, and 4.5.2);
- the length of arterial vessels (see Section 4.5.2); and
- the terminal boundary Windkessel model parameters (see Section 4.5.3).

The prior PDF of the i^{th} such parameter is thus described by the following log-normal distribution:

$$p(\Phi_i) = \frac{1}{\Phi_i \sigma_i \sqrt{2\pi}} \exp\left(-\frac{(\ln(\Phi_i) - \mu_i)^2}{2\sigma_i^2}\right), \quad (4.19)$$

where μ_i and σ_i represent the mean and standard deviation of the underlying Normal distribution, *i.e.* $\ln(\Phi_i) \sim \mathcal{N}(\mu_i, \sigma_i^2)$. This log-normal distribution is created with large variance, resulting in a weakly informative prior. For all the scaling parameters μ_i is set to 0.5 and σ_i^2 is set to 0.8, respectively. This PDF is shown in the middle plot of Figure 4.15.

3. **Unbounded parameters:** A normal prior distribution is chosen for all the scaling parameters with infinite support. The only such parameters are the inlet flow-rate FS coefficients (see section 4.5.3). The prior PDF of the i^{th} FS coefficient with mean μ_i and standard deviation σ_i is thus described as:

$$p(\Phi_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\Phi_i - \mu_i}{\sigma_i}\right)^2\right). \quad (4.20)$$

While the FS scaling parameters do have infinite support, they are expected to be within physiological range and hence within several multiplications of the reference FS coefficients. Thus, $\mu_i = 1$ and a standard deviation $\sigma_i = 2$ is set, resulting in a weakly informative prior.

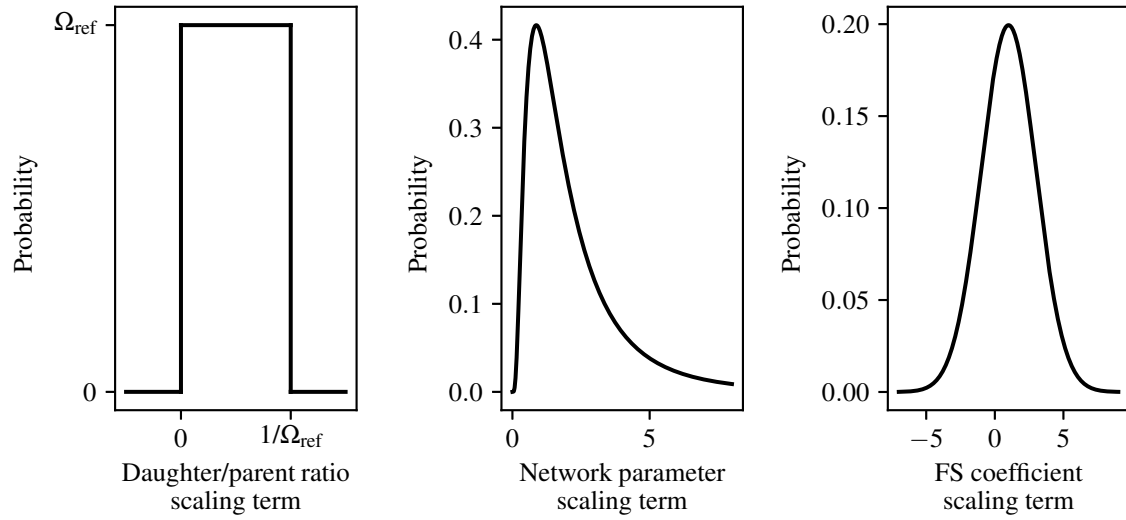


Figure 4.15: The three type of prior distributions used for all arterial network scaling parameters.

It is assumed that all the priors are independent of each other, thus a singular joint prior PDF is:

$$p(\Phi) = \prod_{i=1}^N p(\Phi_i), \quad (4.21)$$

where $N = 199$ is the total number of scaling parameters.

4.6.2 Likelihood

The likelihood term within Bayes' theorem represents the likelihood of observing the measurements μ_{lit} from the model described by Equations (4.12), and (4.13). To construct a function for the likelihood term, examples of observed measurements taken from the model described by Equations (4.12), and (4.13) must be obtained from literature. The literature based measurements being incorporated into the VPD are split into three components. These three components are:

- **Scalar pressure and flow-rate measurements.**
- **A time varying inlet flow-rate and cardiac output measurement.**
- **A vessel length scaling term measurement.**

Each of these components is explained next.

Scalar pressure and flow-rate measurements

Scalar pressure (diastolic and systolic) and flow-rate (average) measurements at locations (radial artery, ascending aorta, common carotid, and femoral artery) reported in literature are incorporated into the likelihood. Their statistics and sources are shown in Table 4.11. When multiple measurements are available for the same quantity, they are pooled together into a single mean and variance [178]. The pooling method used accounts for not only the mean and variance of each individual measurements, but also the number of patients from which each measurement has been computed. Due to the interconnected nature of the arterial network, if physiologically realistic scalar pressure and flow-rate measurements are incorporated at multiple discrete locations throughout the arterial network, it is reasonable to assume that physiologically realistic distributions are occurring continuously throughout the network across the resulting VPD. A break down of the demographics of patients included within these literature based measurements are as follows:

- Two radial artery pressure measurements, one for systolic and one for diastolic pressure, are created based on 71 patients taken from two studies. The first study [152] contains 45 males and 6 females, with age ranges of 48 to 77 years old, all of whom have been recorded to have coronary artery disease. The second study used to create the radial artery pressure measurement [38] contains 16 males and 4 females, with age range of 36 to 78 years old, with a range of cardiac diseases.
- Systolic and diastolic ascending aorta pressure measurements are created based on 69 patients taken from two studies. The first study [152] is the same as the first study used for the radial artery measurement. The second study [138] contains 18 patients, with age range of 19 to 54 years old, with no cardiovascular disease detectable through hemodynamic measurements.
- Systolic and diastolic common carotid artery pressure measurements are created based on 134 patients taken from 3 studies. The first study [199]

contains 4 males and 2 females with heart disease, and a mean age of 45 years old. The second study [11] contains 9 healthy subjects with age range between 24 and 34 years old. The third study [80] contained 118 patients, 54 of whom are female and 65 males, with age range between 9 and 75 years old.

- An average femoral artery flow-rate measurement is created based on 63 patients from two studies. The first study [128] contains 12 healthy patients, 8 male and 4 female, with an average age of 55 years old. The second study [121] contains 51 patients with no evidence of peripheral vascular disease. 26 of these patients are males and 25 females, with age range of 18 to 75 years old.

For each measurement τ_i (either directly measured or pooled together) to be $\mu_{i,\text{dis}}$ with standard deviation $\sigma_{i,\text{dis}}$, the likelihood is:

$$p(\tau_i = \mu_{i,\text{dis}} | \Phi) = \frac{1}{\sigma_{i,\text{dis}} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\mu_{i,\text{dis}} - \mathcal{H}_i(\mathcal{M}(\Phi))}{\sigma_{i,\text{dis}}}\right)^2\right), \quad (4.22)$$

where \mathcal{H}_i represents the observation operator (see Section 4.6) that extracts the τ_i component of the model output. Denoting all such discrete measurements collectively with the vector $\boldsymbol{\tau}_{\text{dis}}$, the combined likelihood with the assumption of independence is:

$$p(\boldsymbol{\tau}_{\text{dis}} = \boldsymbol{\mu}_{\text{dis}} | \Phi) = \prod_i^N p(\tau_i = \mu_{i,\text{dis}} | \Phi). \quad (4.23)$$

Time varying inlet flow-rate and cardiac output measurement.

While the incorporation of the above pressure and flow-rate measurements updates the posterior distribution with regards to expected ranges, it provides no information about the expected shape of pressure and flow-rate profile. Some time-varying information about the behaviour of pressure and flow-rates over the cardiac cycle is necessary to obtain a physiologically realistic posterior. Statistics on any such time-varying behaviour is not reported in literature to the authors' knowledge. Thus, a time-varying pseudo measurement is constructed by combining the reference (see [20]) time varying inlet flow-rate $Q_{\text{ref}}(t)$ and a measurement of average cardiac output [92, 35]. It must be ensured that this pseudo measurement accurately captures not

Measurement location	Measurement type	No. of patients	Recorded measurement	Known error	Sources
Radial artery	Diastolic pressure (mmHg)	71	65	12.84	[152], [38]
	Systolic pressure (mmHg)		123	22.75	
Ascending aorta	Diastolic pressure (mmHg)	69	65	7.14	[152], [138]
	Systolic pressure (mmHg)		103	8.24	
Common carotid artery	Diastolic pressure (mmHg)	134	75.58	6.01	[199], [11] [80]
	Systolic pressure (mmHg)		122.78	12.01	
Femoral artery	Average flow-rate (<i>ml/sec</i>)	63	5.84	2.11	[128] [121]

Table 4.11: The measurement of discrete pressure and flow-rate taken from literature. When multiple source measurements are available for the same quantity, they are pooled together into a single mean μ_{dis} and variance σ_{dis} [178].

only the mean and standard deviation of the inlet flow-rate profile at any time point within the cardiac period, but also the covariance between any two time points. An appropriate method to create such a pseudo measurement is a Gaussian Random Field (GRF, also referred to as a Gaussian Process) [167, 91]. Using a GRF allows for a description of the mean (see Equation (4.24)), standard deviation (Equation (4.25)), and covariance (Equation (4.26)) to be chosen, and then these components combined into a time-varying distribution. The pseudo measurement $Q_{\text{meas}}(t)$ is described by a GRF with the following mean $\mu_{\text{meas}}(t)$ and standard deviation $\sigma_{\text{meas}}(t)$:

$$\mu_{\text{meas}}(t) = Q_{\text{ref}}(t), \quad (4.24)$$

$$\sigma_{\text{meas}}(t) = \left(1 + \frac{Q_{\text{ref}}(t) - \min(Q_{\text{ref}})}{\max(Q_{\text{ref}}) - \min(Q_{\text{ref}})}\right) \mathcal{F}, \quad (4.25)$$

where \mathcal{F} is a positive scaling parameter. The scaling of $\sigma_{\text{meas}}(t)$ is to ensure that i) the standard deviation is positive, and ii) the variance is proportional to the magnitude of the flow rate (*i.e.* the percentage variance with respect to the magnitude remains fixed). This behaviour is shown in Figure 4.16. $Q_{\text{ref}}(t)$ varies between approximately 550mls^{-1} and 0mls^{-1} . Applying a constant standard deviation (*i.e.* $\sigma_{\text{meas}}(t)$ is fixed) based on the maximum measured flow-rate results in time points corresponding to low flow-rates having very large variance proportional to their measured value, and vice versa. The value of \mathcal{F} is empirically tuned after the GRF has been described.

The covariance between the pseudo measurements at any two times t_i and t_j is described through a periodic kernel [167]:

$$\text{cov}(Q_{\text{meas}}(t_i), Q_{\text{meas}}(t_j)) = \sigma_{\text{meas}}(t_1) \sigma_{\text{meas}}(t_2) \exp\left(\frac{-2 \sin^2(\pi|t_i - t_j|/T)}{(vT)^2}\right), \quad (4.26)$$

where T represents the cardiac period, and v represents the ratio of the correlation length to the cardiac period. The correlation length dictates the distance within which two points have a strong affect on one another. Instead of directly stating the correlation length, it is decided to computed it based on the cardiac period. In this study the cardiac period of all VPs is being kept constant. Describing the correlation length in this way, however, ensures that if the cardiac period of VPs were

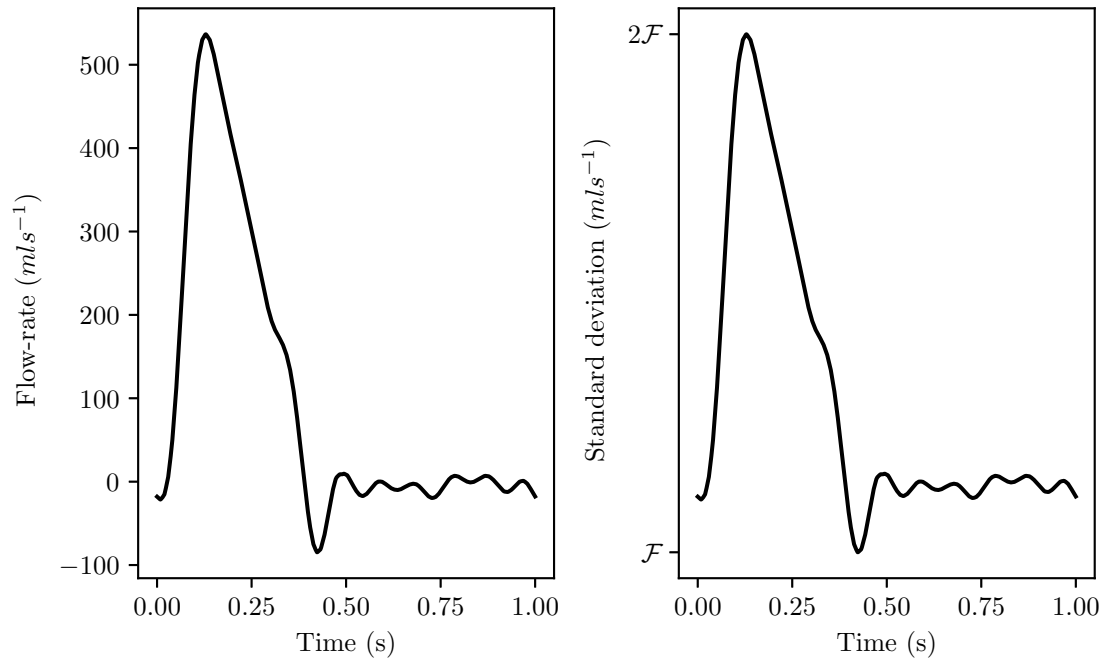


Figure 4.16: On the left the measured flow-rate profile taken from the initial network, being incorporated into the posterior distribution, is shown. On the right the corresponding standard deviation profile, computed using Equation (4.25), is shown in terms of the parameter \mathcal{F} .

to be randomised in future studies two discrete points at a fixed proportion of the period apart will have the same influence on each other, irrespective of the cardiac period. The affect of different correlation lengths, as a proportion of the cardiac period, on the correlation between time points is shown in Figure 4.17. The purpose of the pseudo measurement being incorporated into the posterior distribution is to impose control over the shapes of pressure and flow-rate profiles being produced. If the correlation length used to create the GRF is too low, adequate control will not be imposed on the realisations of inlet flow-rate profiles that can be produced. A high correlation between time points throughout the entire cardiac period, however, will result in a heavy penalisation to the likelihood of the measured inlet flow-rate profile occurring if it is not a direct translations of the observed profile extracted from the model. A correlation length to cardiac period ratio of $\nu = \frac{1}{2}$ is chosen. This ratio can be seen to maintain a high correlation for a reasonable proportion of the cardiac period while still allowing the correlation to drop to 0, balancing the

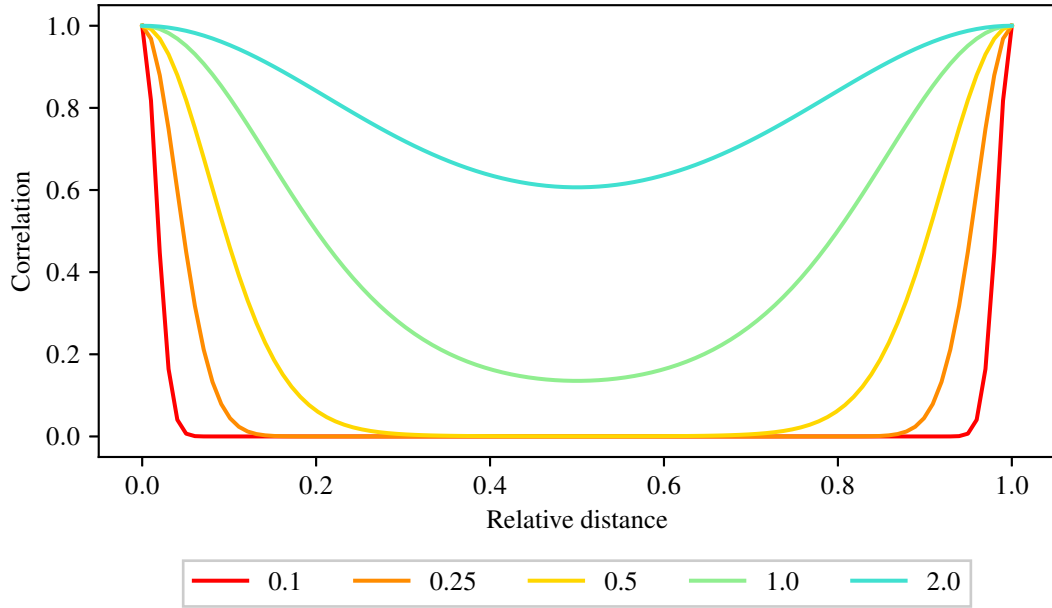


Figure 4.17: The affect of correlation length as a proportion of the cardiac period on the correlation between time points. The x-axis represents the relative distance between two time points as a proportion of the cardiac period, and the y-axis represents the correlation between these time points. Each line shows the correlation profile when a different correlation length to cardiac period ratio is used.

two aforementioned behaviours.

Using the measurement of average cardiac output and the associated standard deviation taken from [92] and [35] the value assigned to the scaling term within Equation (4.25), denoted by \mathcal{F} , is tuned empirically. The pooled measurement of average cardiac output and the standard deviation taken from literature is equal to $98ml/sec$ and $35.55ml/sec$, respectively. A series of unique GRFs are created using standard deviation profiles computed using Equation (4.25) and different values of \mathcal{F} between $50mls^{-1}$ and $200mls^{-1}$. 100,000 realisations of inlet flow-rate profiles are drawn from each GRF. The average cardiac output of each of these realisations is computed, and the mean of the average cardiac output and the associated standard deviation across each GRF is empirically found. The absolute percentage difference between the empirical mean and the associated standard deviation of the average cardiac output of each GRF, and the literature based measurement is computed. This analysis is shown in Figure 4.18.

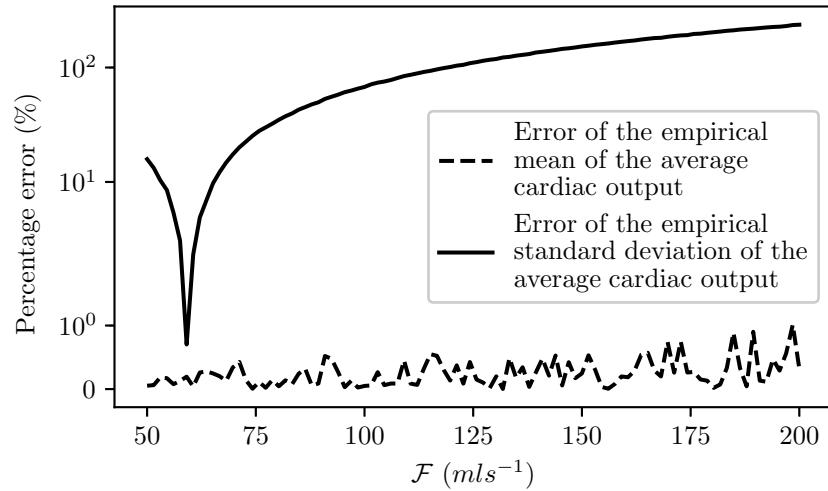


Figure 4.18: The percentage difference between the empirical mean of the average cardiac output and the associated standard deviation of 100,000 realisations drawn from GRFs with different standard deviation scaling terms, and the measured average cardiac output and associated standard deviation taken from literature [92][35].

The discrepancy between the empirical mean of the average cardiac output of each GRF and the literature based measurement should not change when different values of \mathcal{F} are used. Any oscillatory behaviour seen in the percentage discrepancy of the mean average cardiac output will therefore be introduced due to statistical errors. Figure 4.18 shows that the amplitude of the oscillations of the mean average cardiac output discrepancy are negligible, with a difference between maximum and minimum of 1.01%. Figure 4.18 shows that setting $\mathcal{F} = 59.09$ produces a clear and significant minimum total discrepancy. Using this value of \mathcal{F} produces a mean average cardiac output of $97.80\text{ml}/\text{sec}$, compared to the literature based measurement of $98\text{ml}/\text{sec}$, and a standard deviation of $34.84\text{ml}/\text{sec}$, compared to the literature based standard deviation of $35.55\text{ml}/\text{sec}$.

To be able to incorporate the affect of observing the pseudo measurement $Q_{\text{meas}}(t)$ into the posterior distribution, the GRF created above is discretised into a series of time points. As with the correlation length ratio used within Equation (4.26), the number of evaluation points used must be carefully considered to balance control over the shape of flow-rate profiles produced with freedom to produce variability across the final VPD. Evaluating the likelihood of the time domain inlet

flow-rate measurement occurring at a high number of discrete points results in more control being imposed over the shape of realisations of inlet flow-rate profiles assigned to VPs, and so less variability across the final VPD. Due to the greater range of both magnitudes and gradients seen within the first half of the cardiac period, it is decided that two thirds of the evaluation points are in the first half of the cardiac period. The two thirds of evaluation points assigned to the first half of the cardiac period and the one third assigned to the second half are evenly spread across their respective half's. The resulting positioning of evaluation points when discretising the time domain inlet flow-rate profile into six, nine, and twelve points are shown in Figure 4.19. From the left two plots, six and nine evaluation points can be seen to provide too little information about the shape of the time domain inlet flow-rate profile. When using six evaluation points, the shape of the time domain inlet flow-rate profile is not evaluated between the start of the cardiac period and the peak systolic flow-rate. There is also no evaluation of the profile around the minimum flow-rate position. When using nine evaluation points there appears to be better monitoring of the minimum flow-rate value, however there is still a lack of evaluation points between the start of the cardiac period and the peak flow-rate. Using twelve evaluation points appears to alleviate both of these problems. The likelihood of the time domain inlet flow-rate profile occurring is, therefore, evaluated at twelve discrete evaluation points.

Denoting these twelve evaluation time points as $\{t_1, t_2, \dots, t_{12}\}$, the vector of measurements at these times with $\boldsymbol{\tau}_{\text{inflow}}$, the mean GRF vector as $\boldsymbol{\mu}_{\text{inflow}} = [\mu_{\text{meas}}(t_1), \mu_{\text{meas}}(t_2), \dots, \mu_{\text{meas}}(t_{12})]$ computed through Equation (4.24), and the vector of inlet flow-rates produced by the network parameters as $\boldsymbol{Q}_{\text{inflow}} = [Q_{\text{inflow}}(t_1), Q_{\text{inflow}}(t_2), \dots, Q_{\text{inflow}}(t_{12})]$, the likelihood for flow-rate can be written as:

$$p(\boldsymbol{\tau}_{\text{inflow}} = \boldsymbol{\mu}_{\text{inflow}} | \boldsymbol{\Phi}) = (2\pi)^{-\frac{k}{2}} \det(\boldsymbol{\Sigma})^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{\xi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}\right), \quad (4.27)$$

where:

$$\boldsymbol{\xi} = \boldsymbol{\mu}_{\text{inflow}} - \boldsymbol{Q}_{\text{inflow}}, \quad (4.28)$$

and $\boldsymbol{\Sigma}$ is the GRF covariance matrix whose i^{th} row and j^{th} column element $\Sigma_{i,j} = \text{cov}(Q_{\text{meas}}(t_i), Q_{\text{meas}}(t_j))$ can be computed from Equation (4.26); and k is the number of time points at which the GRF is evaluated ($k = 12$).

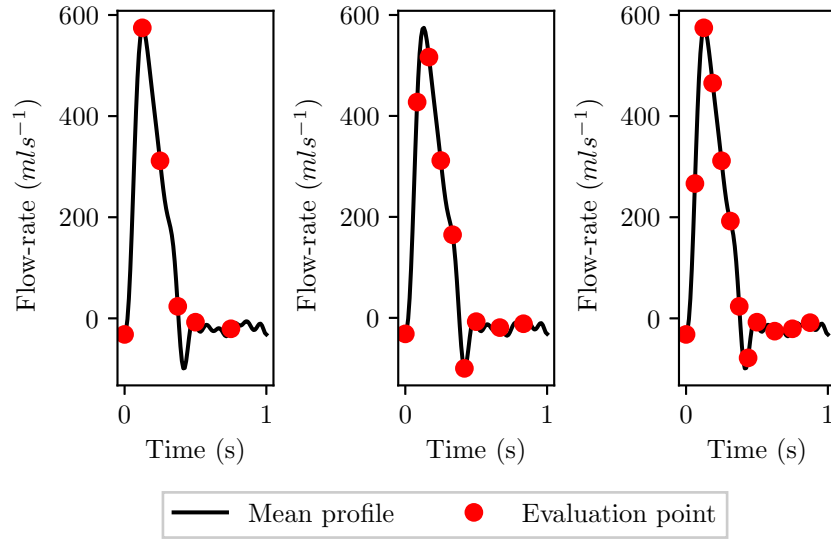


Figure 4.19: The discrete time points at which the likelihood of the time domain inlet flow-rate profiles is evaluated is shown on the mean flow-rate profile when different numbers of evaluation points are used. The left, middle, and right plots show the locations of evaluation points when using 6, 9, and 12 points respectively.

Vessel length scaling term measurement

As mentioned in Section 4.5.2, the length of the vessels are parameterised by a single scaling term relative to the reference network. Since statistics of any direct vessel length measurement are not available, it is assumed that the lengths of arterial vessels are directly proportional to the height of a subject. A study of 25,945 twins from eight countries (Australia, Denmark, Finland, Italy, Netherlands, Norway, Sweden, and the UK) found the mean and standard deviation of the height of the full cohort to be 172.0cm and 9.308cm respectively [191]. Since the reference arterial network has a patient of height 170 cm, the measurement data corresponds to a mean of $\mu_{\text{len}} = 1.0118$ and the standard deviation of $\sigma_{\text{len}} = 0.0548$ for the vessel length scaling term. Denoting this measurement as τ_{len} , the likelihood for the scaling parameters Φ_{len} is thus

$$p(\tau_{\text{len}} = \mu_{\text{len}} | \Phi) = \frac{1}{\sigma_{\text{len}} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\mu_{\text{len}} - \Phi_{\text{len}}}{\sigma_{\text{len}}}\right)^2\right), \quad (4.29)$$

Assuming all the measurements to be independent, the combined likelihood in

Equation (4.17) is:

$$p(\boldsymbol{\tau}|\boldsymbol{\Phi}) = p(\boldsymbol{\tau}_{\text{dis}} = \boldsymbol{\mu}_{\text{dis}}|\boldsymbol{\Phi}) \times p(\boldsymbol{\tau}_{\text{inflow}} = \boldsymbol{\mu}_{\text{inflow}}|\boldsymbol{\Phi}) \times p(\tau_{\text{len}} = \mu_{\text{len}}|\boldsymbol{\Phi}), \quad (4.30)$$

where the RHS terms can be computed from Equations (4.23), (4.27), and (4.29).

In summary, the prior distributions for the 199 network parameters are specified in Section 4.6.1. These priors will be modified through the likelihood—specified by 7 scalar measurements of pressure and flow-rate, 1 time-varying flow-rate profile evaluated at 12 time points in the cardiac cycle, and a measurement of vessel lengths—to yield the posterior distribution of the parameters through Equation (4.17). Random sampling from this posterior will result in the virtual patient database. The sampling procedure is presented next.

4.7 Sampling from the posterior distribution

With the prior and the likelihood specified, the posterior distribution is given by Equation (4.17), and can be evaluated at any given $\boldsymbol{\Phi}$ up to a normalising constant (the evidence term in the denominator of the equation). Sampling from this analytically intractable posterior is achieved through the Markov-chain Monte Carlo (MCMC) method [69, 71] (see Chapter 2), which is a widely used method to sample from unnormalised distributions.

Given an initialising position within a distribution of interest, an MCMC method draws sequential dependent samples from that distribution. An iteration within a Markov chain is required to be only dependent on the previous iteration of the chain. Therefore, given an adequate length chain, an MCMC method is able to accurately sample from a distribution when direct sampling is infeasible. It is chosen that 75,000 VPs are to be sampled from the posterior distribution.

The MCMC algorithm is sequential—each subsequent sample depends on the previous sample and hence the chain grows only one sample at a time. Generating a long chain this way thus leads to very high computational run times, since computation of the likelihood requires running a pulse wave propagation simulation at each step. It is shown in Section 4.3 that the physic based model of pulse wave propagation takes approximately 25 seconds to solve for a healthy VP (using the optimum mesh found for healthy patients in Section 4.3, an initial pressure of 75

mmHg, and running for eight cardiac cycles to allow for enough time to converge to periodic solutions). It is likely that for some VPs convergence will be reached in significantly fewer than eight cardiac cycles, however it is chosen to use eight to ensure more than enough time for all VPs to converge. The process of VP creation could be optimised (*i.e.* a reduction in the computational expense) in future studies by automatically adapting the mesh used to solve VPs and the number of cardiac cycles run to better suit individual VPs. It is known that 75,000 VPs are to be sampled from the posterior distribution. It is empirically found that using a standard deviation of 0.0375—equivalent to 3.75% of the reference values—within the proposal distribution (see Equation (2.11)) for the scaling terms applied to the inlet flow-rate FS coefficients, and 0.025—equivalent to 2.5% of the reference values—within the proposal distribution of all other scaling terms results in an acceptance rate of approximately 0.2 (see Chapter 2 for details). This analysis is based on MCMC chains of short length, and so the acceptance rate of the final chain is likely to differ from this estimate. Assuming that the acceptance rate of the chain is equal to the previously estimated value of 0.2, however, the required wall clock time to complete the MCMC chain is computed to be 109 days.

To achieve some level of parallelisation pre-fetching is employed (see Chapter 2 for details). For this study a decision tree depth $\eta = 4$ is used. Assuming an acceptance rate of 0.2 and no increase in time per decision tree as the depth increase, this results in the creation of 75,000 VPs taking 28 days. Before this new VPD is used to train and test machine learning classifiers, the *a posteriori* behaviour of the new VPD itself must be analysed, presented next.

4.8 Results and discussion

The VPD is created by generating samples through the MCMC algorithm. The results for the MCMC method and the VPD are presented and discussed here.

4.8.1 Burn-in analysis and chain diagnostics

Formally there are no restrictions on the starting position of a Metropolis-Hastings MCMC chain, other than it must be within the support of all parameters. It is common, however, for the initial iterations of an MCMC chain to be considered

as the “burn-in” period, during which the chain converges on an equilibrium distribution. This burn-in period is therefore discarded from the final sampled posterior distribution. Intuitively this practice of discarding the initial portion of a MCMC chain does not make sense. A defining feature of a Markov chain is its memorylessness, and so the samples drawn from the posterior distribution should be independent of the initialising position. Ideally an MCMC chain is run for infinite length, in which case a very high number of samples are drawn from the entire distribution. In this case the initialising position of the chain has no influence on the final distribution from which samples have been drawn. It is not possible, however, to run a MCMC chain of infinite length, and instead the chain must be truncated. This results in the distribution from which samples are drawn being an estimate of the real posterior distribution.

While an MCMC chain is within a region of high posterior density it will move around in a random manner. Eventually the chain will move beyond the region of high density, into a region of low density. Once the chain is within a region of low density it can be thought of as behaving pseudo-randomly. While successive steps in the chain will have no clear pattern, it is expected that while in a region of low density there will be a net movement towards regions of higher density. As the chain once again drifts towards regions of high density, the net “pull” on the chain is alleviated, allowing it to behave in a manner more similar to its desired random walk. This pattern of behaviour is repeatedly completed, building a clearer picture of the posterior distribution each time it is. If a MCMC chain is initialised within a region of low posterior density, the chain will begin within a “migratory” period. Initialising an MCMC chain of finite length within a region of low density results in an inflation in the number of samples drawn from that region comparative to the number of samples drawn from regions of high density. The magnitude of this inflation is dependent on two factors i) how long the chain initially takes to reach a region of high density, and ii) the total length of the chain. If the chain is run for a very short period, so that only a small number of samples are drawn from the region of post migration high density, this inflation is great. As the length of the chain increases, and so more samples are drawn from the entire distribution, the affect of this initial migration is mitigated. In a similar manner if the MCMC chain moves slowly through the distribution, and so takes a long time to reach the region of high density, the affect of this inflation on the estimated posterior distribution is greater.

If the initial migratory period of the chain is reduced, the affect of this period on the estimate of the posterior distribution is also reduced. Aside from simply running the MCMC chain for more iterations, removing a burn-in period from the samples drawn is one of the simplest ways of increasing the accuracy of the estimate of the posterior distribution. A burn-in period is used with the intention of removing the initial migratory period, so that the chain starts within a region of high density. This method is not exact, and it is often impossible to ensure all parameters are starting within a high density region. The MCMC chain is initialised at the reference network, which is known to be a reasonable realisation of an arterial network, and thus should be in a region of high density.

To estimate the required initial burn-in period, trace plots for every parameter at each iteration of the MCMC chain are plotted. An example of a trace plot is shown for the scaling terms applied to the length of VPs arterial vessels in Figure 4.20, and all other trace plots are shown in Appendix H. To aid visualisation, these plots are thinned out by plotting every 100th iteration of the chain. These trace plots are visually assessed to determine burn-in. Essentially, if there is an initial period where a clear migration of the chain is seen away from the sample used to start the chain, then the burn-in is selected to be until this initial migration is complete. If no initial migratory period is seen, and so the initial position is suspected to be within a region of high density, a precautionary initial burn-in period is still applied. Aside from determining the required initial burn-in period, the trace plots of the values assigned to each parameter at each step of the chain allows the movement of the chain around the distribution, referred to as the “mixing” of the chain, to be examined.

Generally figures in Appendix H show that most parameters are initialised in a region of high density and hence no net migration is observed—the samples oscillate around a central mean value. Some interesting behaviours in the trace plots, including features that deviate from such a general/desired behaviour, are:

- The scaling terms applied to the initial β at the inlet of the aortic chain are seen to centre around 5. The maximum scaling term applied is approximately 7.5.
- The scaling terms applied to the reference decay term of the left leg chain β profile initially oscillates around a value of approximately 1 for 25,000

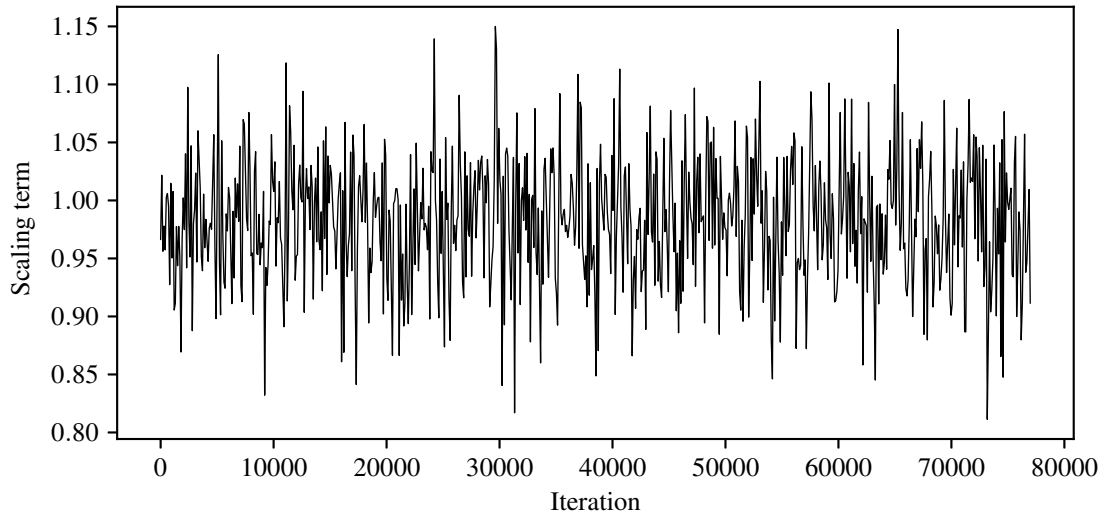


Figure 4.20: MCMC trace plot of the scaling term applied to the length of the arterial vessels at every 100th iteration. In this case, not net migration of the chain is observed.

iterations, before migrating and oscillating around a value of approximately 7 (see Figure H.1). This may suggest that this distribution is multi-modal.

- The scaling terms applied to the reference decay term of the brachiocephalic trunk β profile oscillates around a value of 1 for the first 40,000 iterations, before migrating and oscillating around a value of 9 (see Figure H.2). The scaling terms applied to the reference decay term of the brachiocephalic trunk r_0 profile shows no complimentary behaviour and remains within the region of 0–4 throughout. This may again suggest a multi-modal distribution.
- A spike is seen in the scaling terms applied to the compliance of the right external carotid Windkessel model at approximately 55,000 iterations (see Figure H.16). This behaviour is not seen within the scaling terms applied to the compliance of the left external carotid Windkessel model which remains within the region of 0–3 throughout.

Even though most of the parameters do not show a clear migration, an initial burn-in period of 10,000 iterations is chosen as a precautionary measure to minimise any affect the initial sample position may have. Once this burn-in period has been

removed, the VPD contains 65,000 VPs. Along with the burn-in period, all VPs with negative average flowrate in any arterial vessel are removed from the VPD. These VPs are removed as it is physiological unlikely for a patient to have negative average flowrate in any vessel. Of the 65,000 post burn-in VPs, 12,857 are removed due to the presence of negative average flowrate, reducing the VPD to 52,143

For the assessment of convergence diagnostics, the measure of integrated auto-correlation time (IACT) [155, 41, 176] for the parameters and the measurements (pressure and flow-rate) is employed. The IACT represents the average temporal distance within which two samples are correlated—*i.e.* two samples at a temporal distance greater than the IACT are independent of one another. In theory samples draw by an MCMC chain are only dependent on the previous iteration of the chain, thus the IACT can be thought of as the computational inefficiency of the chain. The IACT is computed by first finding the autocorrelation (AC)—*i.e.* the correlation of the MCMC chain with its future self, as a function of the lag length (the temporal distance between which the correlation is evaluated)—across all possible lag lengths. The AC of the j^{th} arterial network parameter scaling term (Φ_j) at a lag length of ‘lag’—described as $\text{Autocorr}_{(j,\text{lag})}$ —is equal to:

$$\text{Autocorr}_{(j,\text{lag})} = \frac{\sum_{i=1}^{m-\text{lag}} (\Phi_{(j,i)} - \bar{\Phi}_j)(\Phi_{(j,i+\text{lag})} - \bar{\Phi}_j)}{\sum_{i=1}^{m-\text{lag}} (\Phi_{(j,i)} - \bar{\Phi}_j)^2}, \quad (4.31)$$

where m represents the total number of VPs sampled ($m = 65,000$ as this analysis is performed post burn-in however pre-removal of negative average flow-rate VPs), $\Phi_{(j,i)}$ represents the j^{th} arterial network parameter scaling term for the i^{th} VP, and $\bar{\Phi}_j$ represents the mean value of the j^{th} arterial network parameter scaling term across the m VPs. For the parameter representing the scaling term for the vessel lengths, and for the measurement of diastolic pressure in the ascending aorta, plots of autocorrelation versus lag length [155, 41, 176] are shown in Figure 4.21. From these AC values, the IACT of the j^{th} arterial network parameter scaling term (IACT_j) is equal to:

$$\text{IACT}_j = 1 + 2 \sum_{\text{lag}=1}^{m-2} \text{Autocorr}_{(j,\text{lag})}. \quad (4.32)$$

The maximum lag length is equal to $m - 2$ as this is the distance between the first and last sample. A lower IACT is desirable as this suggested less inefficiency in the

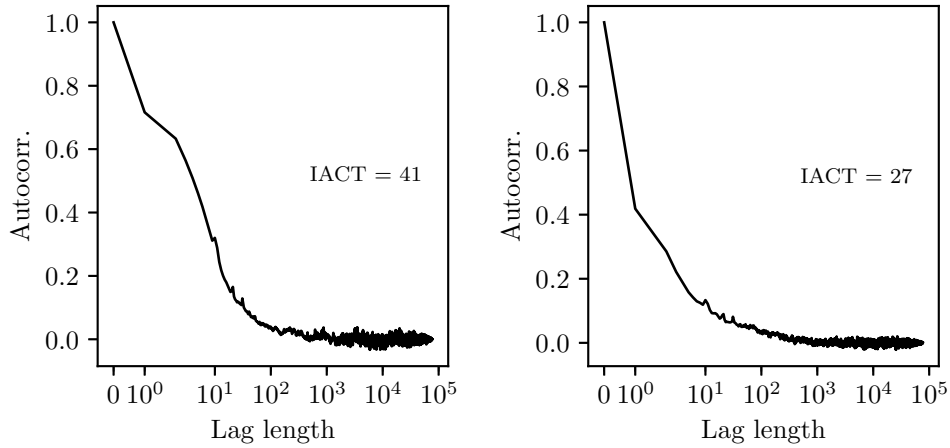


Figure 4.21: Autocorrelation as a function of lag-length in the MCMC chain for the vessel length scaling parameter (left), and for the diastolic pressure in the ascending aorta (right). IACT denotes the integrated autocorrelation time [41, 155, 176].

MCMC chain.

It is found that most of the parameters and the measurements show an IACT of less than 50, comparable to those reported in previous studies utilising 1D flow models[155]. For some parameters, even higher IACT (> 200) are observed. While this does not invalidate the VPD, it does imply that the chain efficiency can be improved by either (i) reducing the dimensionality of the problem further; (ii) further tuning individual step-sizes in Equation (2.11); or (iii) by utilising more efficient MCMC samplers; see, for example, the work by Paun and Husmeier[153]. One reason for higher IACTs for some parameters may be the correlation between the parameters in the posterior. Across the 19,701 pair combinations, a histogram of the correlation coefficients is shown in Figure 4.22, showing that most pairs show low to moderate correlations $|r| < 0.5$, with few combinations showing more extreme correlations. Since the dimensionality is high, even moderate correlations can make the navigation of chain harder because of the curse of dimensionality[14].

4.8.2 Posterior vs. literature reported measurements

The joint posterior distribution of the 199 parameters is difficult to visualise. Even if only pairwise slices of the joint distribution are considered, the number of pairs (19,701) is too high to be shown in this thesis. To be consistent

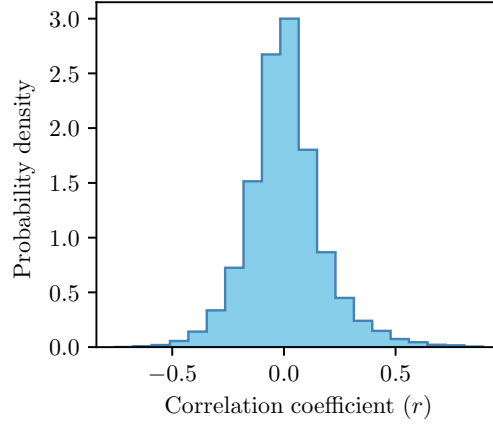


Figure 4.22: Histogram of correlation coefficient (r) between the 19,701 pairs of parameters in the posterior.

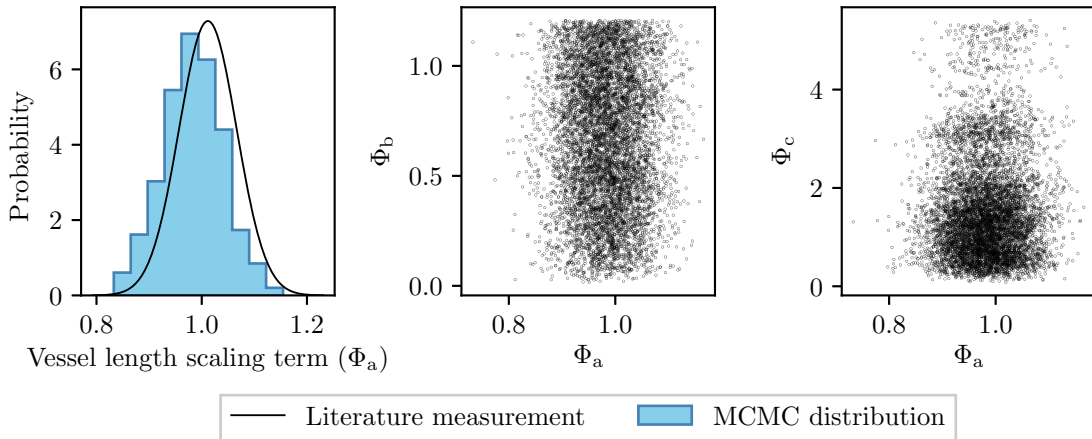


Figure 4.23: Posterior marginal distribution for the vessel length scaling term, denoted as Φ_a (left); joint distribution between Φ_a and the right arm chain’s scaling factor for the β daughter/parent ratio denoted as Φ_b (mid); and joint distribution between Φ_a and the brachiocephalic trunk’s scaling factor for β decay denoted as Φ_c (right). While Φ_a and Φ_b show a largely uniform joint distribution, Φ_a and Φ_c show a potentially multi-modal distribution with local peaks at $\Phi_c \approx 1.0$ and $\Phi_c \approx 3.2$.

with the trace-plot presented in the previous section for the vessel length scaling parameter (described here as Φ_a), its posterior distribution is shown in Figure 4.23 (left). The remaining two plots (centre and right) show pairwise slices of the joint distribution of this parameter with two other parameters—the right arm chain’s scaling factor for the β daughter/parent ratio (described as Φ_b) and the brachiocephalic trunk’s scaling factor for β decay (described as Φ_c). Since the goal of this work is not to make inference about the parameters but to produce a VPD consisting of realistic measurements (pressure and flow-rate), the posterior distribution of the measurements within the VPD is important to be considered. As described in Section 4.6, the posterior distribution is a combination of the prior distribution and the likelihood term within Bayes’ theorem. Since the chosen prior distributions are weakly informative or uninformative (see Section 4.6.1), the distributions of measurements within the VPD should be close to the literature based measurements incorporated through the likelihood. These are not necessarily identical though, as the prior distribution corrects the posterior to account for the geometrical and physiological constraints. Furthermore, since the incorporated measurements are taken from several different sources, the posterior distribution resolves inconsistencies between such measurements, the physics of pulse-wave propagation, and the constraints imposed by the prior. For example, if the prior distribution corresponds to lower valued outputs compared to the literature measurements, the posterior distribution will underestimate the literature measurements, and vice versa. A comparison of the VPD pressure and flow-rate distributions to the literature based measurements (see Table 4.11 and Section 4.6.2) are shown in Figure 4.24 and Table 4.12. A similar comparison for the vessel length scaling term (see Section 4.6.2) is shown in Figure 4.23 (left). Finally, the statistics of the time varying inlet flow-rate profiles in the VPD are compared to the GRF (see Section 4.6.2) in Figure 4.26.

Generally, a good agreement between the scalar pressure and flow-rate measurements are seen in Figure 4.24. This agreement enforces confidence in the overall approach. However, for the average flow-rates in the left and right femoral arteries, larger than expected differences are observed. This is likely due to the following reasons:

- A large inconsistency between the femoral flow-rate measurements and the

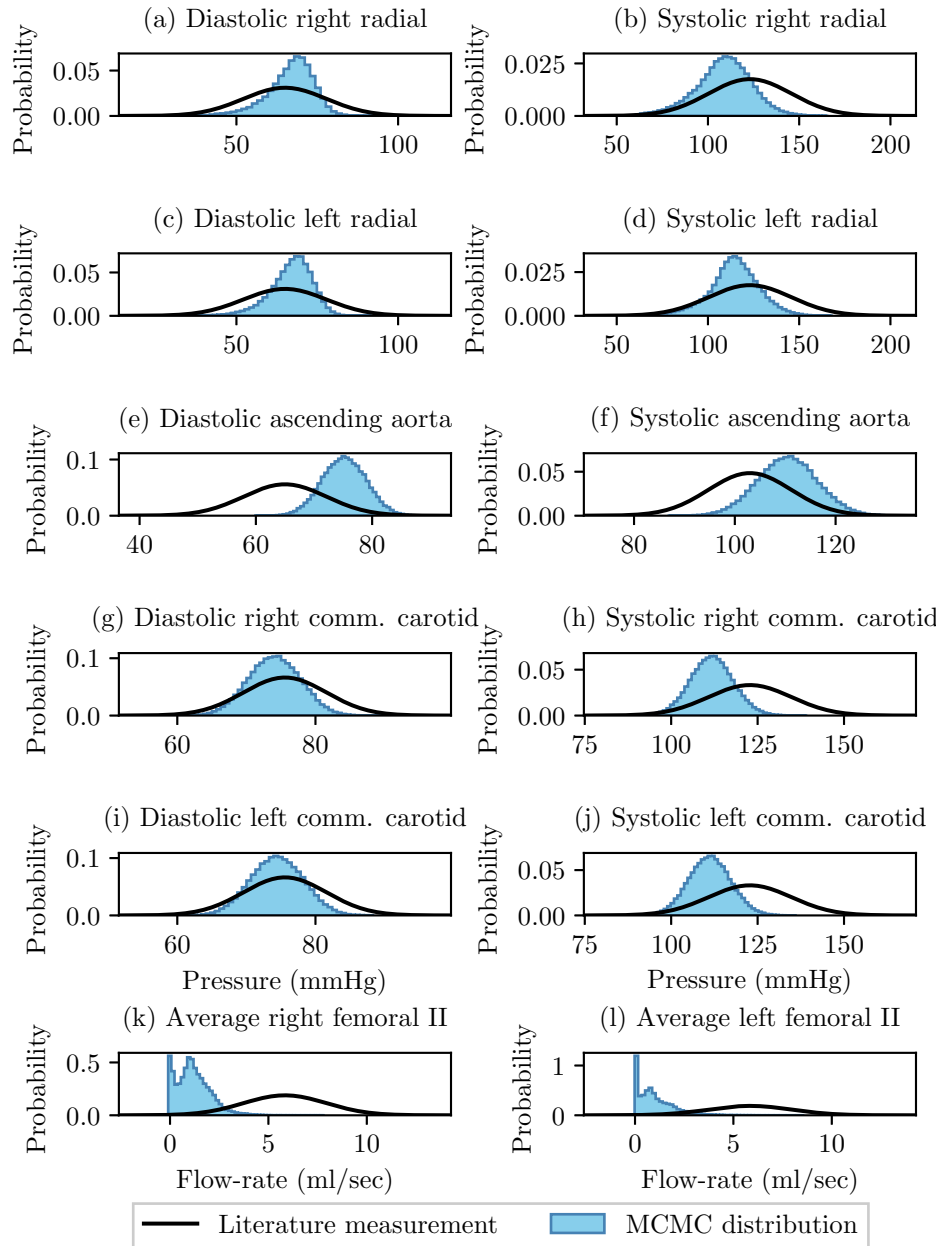


Figure 4.24: Histograms of the MCMC distributions of the pressure and flow-rate measurements at all measurement locations. The literature based measurements and associated error distribution at each location are overlaid in black. Diastolic and systolic pressure in the right radial artery are shown in (a) and (b), respectively; the diastolic and systolic pressure in the left radial artery are shown in (c) and (d), respectively; the diastolic and systolic pressure in the ascending aorta are shown in (e) and (f), respectively; the diastolic and systolic pressure in the right common carotid artery are shown in (g) and (h), respectively; the diastolic and systolic pressure in the left common carotid artery are shown in (i) and (j), respectively; and the average flow-rate in the second segments of the right and left femoral artery are shown in (k) and (l), respectively.

Measurement location	Measurement type	Source	Side	Measured mean (μ_{dis})	Measured std. (σ_{dis})
Radial artery	Diastolic pressure (mmHg)	Literature	Both	65	12.84
		Posterior	Right	67.91	6.76
			Left	67.69	6.29
	Systolic pressure (mmHg)	Literature	Both	123	22.75
		Posterior	Right	108.07	15.48
			Left	114.60	13.66
Ascending aorta	Diastolic pressure (mmHg)	Literature	-	65	7.14
		Posterior	-	75.52	3.69
	Systolic pressure (mmHg)	Literature	-	103	8.24
		Posterior	-	110.70	6.03
Common carotid artery	Diastolic pressure (mmHg)	Literature	Both	75.58	6.01
		Posterior	Right	74.43	3.72
			Left	74.80	3.75
	Systolic pressure (mmHg)	Literature	Both	122.78	12.01
		Posterior	Right	111.91	6.25
			Left	111.45	6.19
Femoral artery (left and right)	Average flow-rate (ml/sec)	Literature	Both	5.84	2.11
		Posterior	Right	1.30	0.77
			Left	1.08	0.81

Table 4.12: The measurement of the discrete pressure and flow-rate taken from literature and the corresponding posterior distributions. See Table 4.11 for details and sources of the literature measurements.

other pressure measurements. Note that they are taken from different sources and hence not from the same population. The percentage difference between the average flow-rate in the right and left femoral arteries taken from the reference network and the literature based measurements are equal to -73.11% and -74.14% respectively, while all other pressure and flow-rate measurements exhibit a percentage difference from the reference network of less than 30%. Furthermore, there may be an inconsistency between the cardiac output measurement used for generating the pseudo-measurement for the time varying inlet flow-rate measurement (see Section 4.6.2).

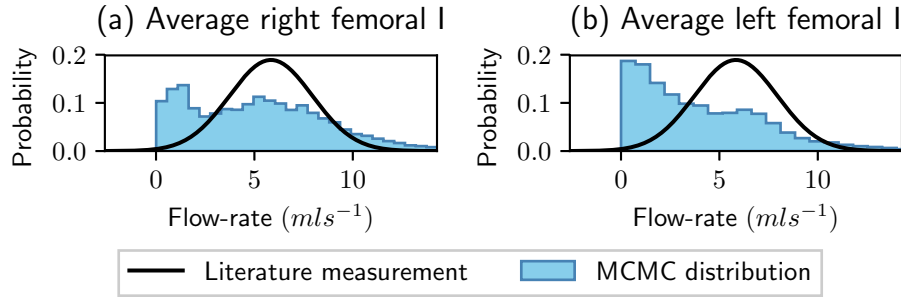


Figure 4.25: Histograms of the MCMC distributions of the average flow-rate in the first segment of the right femoral artery (a), and the first segment of the left femoral artery (b). The literature based measurement of average flow-rate in the femoral arteries is overlaid in black.

- The presence of more measurements for pressure as opposed to flow-rates. Since the likelihoods for all the measurements are weighted equally, it may be possible that the chain is influenced weakly by the few flow-rate measurements. Such an issue can be resolved in future studies by assigning variable weightings to the measurements.
- In the network the femoral arteries are split into the first (I) and second (II) segments, shown in Figure 4.13 by ‘u’ and ‘v’, respectively. The precise location at which the literature based femoral flow-rate measurement has been taken is unknown. In this study, it is assumed that this measurement was acquired at the centre of the second segment ‘v’, as this vessel has a reference length of 31.92cm accounting for 90.86% of the reference length of the two femoral segments. Since, the first segment ‘u’ bifurcates into ‘v’ and another segment *profunda femoris* (Figure 4.13), the flow-rate in ‘v’ is smaller than that in ‘u’. Thus, it is possible that the measured flow-rate was in ‘u’. A comparison of VPD flow-rate in the first femoral segment against the measurement is shown in Figure 4.25, showing the agreement is significantly better.

Figure 4.23 shows a good agreement between the vessel length scaling term in the VPD and the literature based measurement. The mean and standard deviation of this term in the VPD are 0.9819 and 0.0551, respectively, which compare well to those reported in literature with values of 1.0118 and 0.0548, respectively.

Figure 4.26 shows a good agreement between the statistics of the time varying

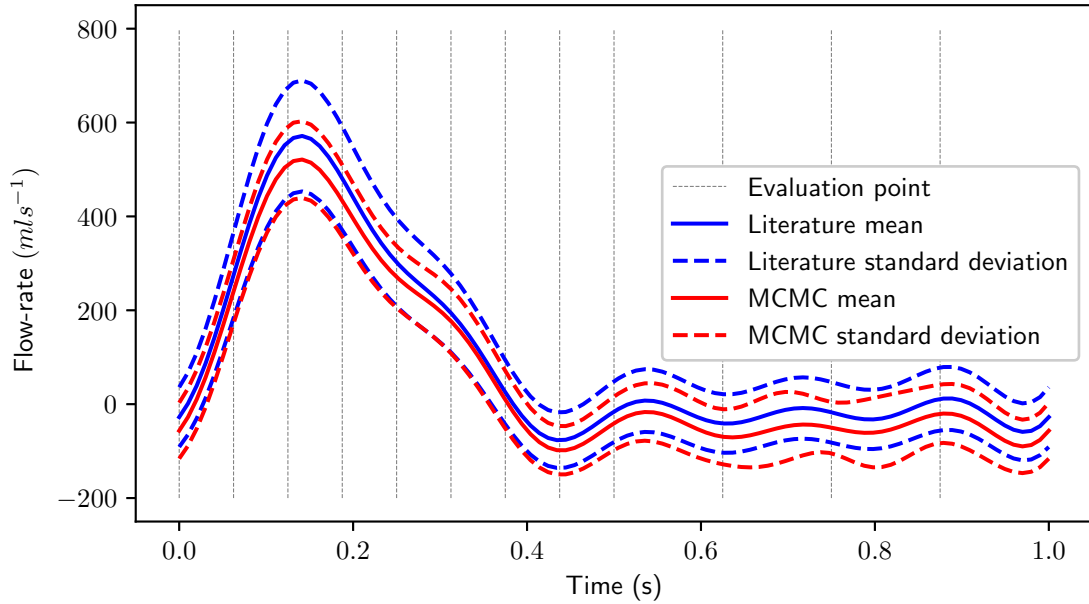


Figure 4.26: Comparison of the empirical distribution of the time varying inlet flow-rate profiles between the MCMC samples and the pseudo measurement created through a GRF. All standard deviation curves in dashed lines depict mean ± 1 standard deviation. See Section 4.6.2 for details on the GRF and the literature reported measurements used to create it.

inlet flow-rate profile in the VPD and the pseudo measurement constructed through the GRF (Section 4.6.2). The mean profile closely follows that of the GRF throughout the cardiac cycle, with maximum difference of approximately 10% at peak systole. Figure 4.26 also shows \pm one standard profiles for both the VPD flow-rate and the GRF pseudo measurement. Throughout the cardiac cycle, a high degree of agreement is seen, with largest errors of approximately 10% magnitudes, again, at peak systole. The VPD mean inlet flow-rate profile having a lower magnitude than the literature based pseudo measurement maybe partially responsible for the mean average femoral flow-rates being lower than the measurement taken from literature. The difference between the VPD distribution of inlet flow-rate profiles and the literature based measurement does not appear to be severe enough to be entirely responsible for the differences in the distribution of average femoral flow-rate, however.

Correlation profiles—similar to that shown in Figure 4.17—are produced for the

posterior inlet flow-rate profiles at each discrete evaluation point. For each profile, a different evaluation point is treated as the initial position, and the correlation between the flow-rate at this point and all other discrete evaluation points is computed across the VPD. The profiles of correlation between the flow-rate at each initial position and all other evaluation points are plotted against the temporal distance between the points as a proportion of the cardiac period. These profiles are shown within Figure 4.27. Ideally all twelve correlation profiles—one starting at each discrete evaluation point—should closely match the correlation profile used when creating the GRF. The individual correlation profiles within Figure 4.27 have not been identified as they all follow very similar shapes, and so the overall patterns and behaviours are more important than the behaviour of the profile corresponding to each discrete evaluation point. Overall, it appears as if all correlation profiles show relatively good agreement with the desired shape. While it first appears as if there is significant negative correlation in the posterior profiles, looking at the y-axis of Figure 4.27 it can be seen that the maximum magnitude of negative correlation is approximately 0.5. The correlation between most discrete evaluation points beyond the correlation length, for which the correlation should be equal to 0, are greater than -0.2.

4.8.3 Evaluation of individual VPs

While analysis of the distribution of measurements is very useful for understanding the overall patterns and behaviours of the VPD, it is also important to look at what pressure and flow-rate profile shapes are occurring through out the arterial network of individual VPs. To determine if sufficient restraint has been imposed on the range of arterial networks that can be produced, the pressure and flow-rate profiles of individual VPs must be examined. Random samples from the VPD are assessed to gain further insights on the VPs and the behaviour of pressure and flow-rate profiles. Pressure profiles are examined at the ascending aorta; right and left radial arteries; and right and left common carotid arteries. Flow-rate profiles are examined at the right and left second femoral segments. Along with the distributions, these examination locations should be sufficient to understand the behaviour of pressure and flow-rate profiles throughout the arterial networks of VPs, while limiting the number of profiles being visualised to a feasible number. 15 VPs are randomly drawn

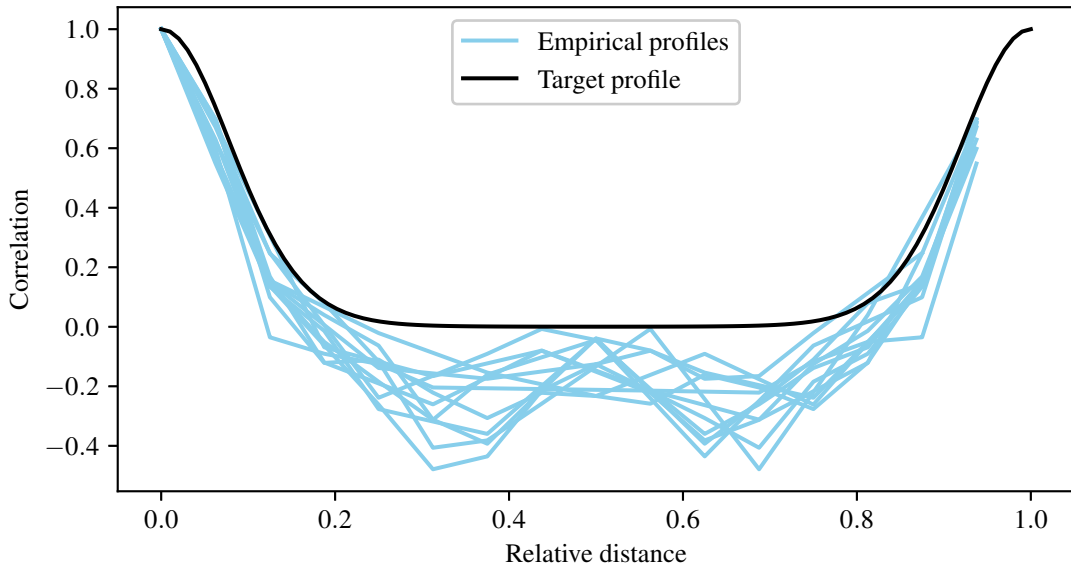


Figure 4.27: Each blue line shows the profile of correlation between a different starting position and all other evaluation points. The black line shows the correlation profile used within the pseudo measurement GRF, outlined in Section 4.6.2.

from the VPD (excluding the burn-in and VPs with negative average flow-rate), and the pressure and flow-rate profiles associated with each are shown at all examination locations in Appendix I. Five VPs of interest are extracted from Appendix I and shown in Figure 4.28.

Figures in Appendix I show that pressure profiles in proximal vessels, *i.e.* the ascending aorta and two common iliacs, show more consistency and similarity to the reference profiles compared to the pressure and flow-rate profiles in distal vessels, *i.e.* the radial artery pressure and the femoral artery flow-rate. This suggests that the pseudo measurement of the time varying inlet flow-rate profile is sufficient to impose control over the shape of pressure and flow-rate profiles in the proximal vessels. As the spatial distance from the inlet of the aorta increases, an increase in the variability of pressure and flow-rate profiles is observed. Thus, access to more measurements, distributed across the network in both space and time, will result in an even more realistic database as the likelihood at several such locations will guide the posterior.

Figure 4.28 shows a selection of undesirable behaviour in the VPD profiles.

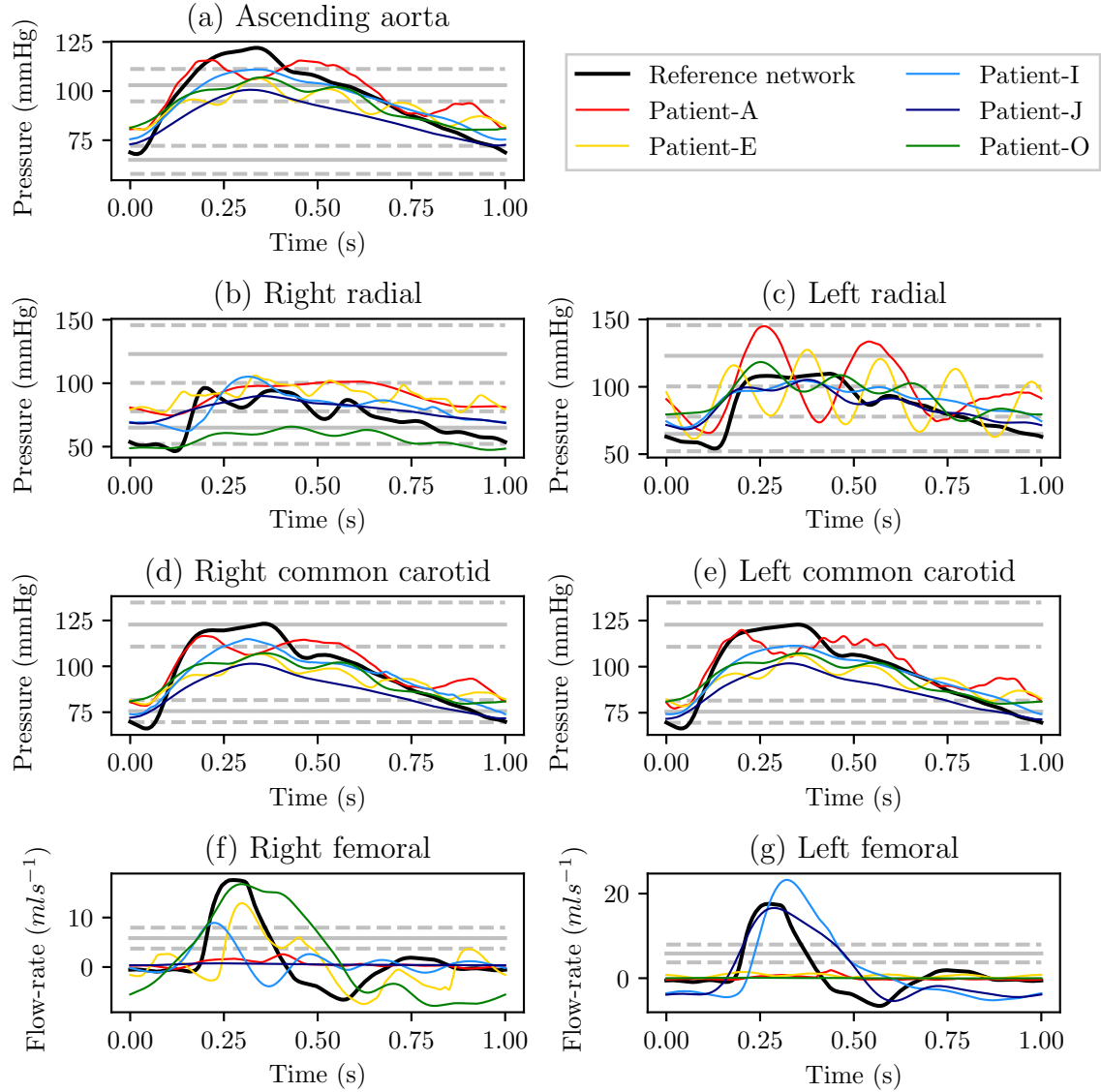


Figure 4.28: In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error (see Table 4.11) are shown by the solid and dashed grey lines, respectively.

Highly oscillatory pressure profiles are observed in the left radial artery, shown in Figure 4.28(c), of Patient-A and Patient-E. In the case of Patient-E, oscillatory behaviour is seen within all the pressure and flow-rate profiles. This suggests that the cause of oscillations is systemic, and possibly the shape of the inlet flow-rate profile prescribed. While this behaviour is not very common in the database, if needed its occurrence can be further reduced by imposing stronger correlations between inlet flow-rates at two time points through the parameter v in Equation (4.26) in the GRF. In the case of VP Patient-A there is no significant oscillatory behaviour in any other profile, but the pressure waveform in the right radial artery is featureless with no clear systolic and diastolic points. This apparent over- and under-damping of the pressure profiles in the radial arteries may suggest an imbalance in the compliances or resistances of the right and left arms. This hypothesis of left-right imbalance is further supported by Figures 4.28(f) and (g) when the femoral flow-rates are observed for Patient-I: the flow-rate in the right femoral artery shows a high-mean high-oscillation behaviour while in the left femoral artery it is low-mean with significantly lower oscillations. Similar behaviour is seen within Patient-J and Patient-O. In future studies, a symmetry metric that balances, while still allowing for some variability, the left and right side parameters for symmetric left and right side vessels may be implemented.

4.8.4 Analysis of resistance and compliance ratios of opposite extremities

In the previous subsection it is seen that there appears to be inadequate control imposed on the properties of opposite extremities. To examine the affect of the compliances and resistances of opposite extremities on the pressure and flow-rate profiles produced, the resistances and compliances of the lower extremities are computed (see Section 4.4.3 and Figure 4.5) up to the second bifurcation in the leg vessels, *i.e* up to the end of vessels labelled ‘u’ in Figure 4.13. To assess the left-right imbalance, the ratio of femoral pulse (maximum flow-rate minus the minimum flow-rate) on the left and right sides is considered:

$$\text{RFFP} = \frac{\max(Q_{\text{RSF}}) - \min(Q_{\text{RSF}})}{\max(Q_{\text{LSF}}) - \min(Q_{\text{LSF}})}, \quad (4.33)$$

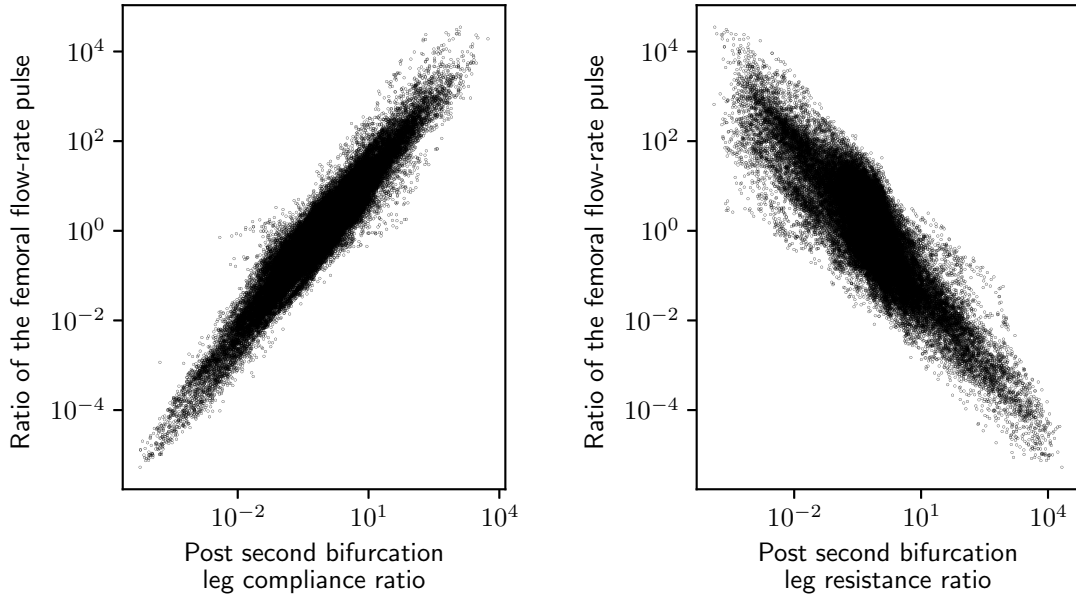


Figure 4.29: The relationship between the ratio of femoral flow-rate pulse to the compliance and resistance ratios.

where RFFP represents the ratio of the femoral flow-rate pulse; $\max(Q_{\text{RSF}})$ and $\min(Q_{\text{RSF}})$ represent the maximum and minimum flow-rate in the right second femoral segment; and $\max(Q_{\text{LSF}})$ and $\min(Q_{\text{LSF}})$ represent the maximum and minimum flow-rate in the left second femoral segment. If, as hypothesised, the observed over- and under-damping of flow-rate profiles in opposite extremities is due to an imbalance in the properties of the each extremity, then there should be a clear relationship between either the ratio of resistances or the ratio of compliances to the ratio of the femoral flow-rate pulse. The relationship between the resistance ratio and RFFP, and the compliance ratio and RFFP is shown in Figure 4.29. Other variables and combinations (such as the product of resistance and compliance) are tested, however it is found that the resistance and compliance show the greatest correlation to the RFFP. The left-right imbalance is apparent in Figure 4.29, which shows that the ratio of left-to-right femoral pulses varies from 10^{-4} and 10^4 across the VPD.

While Figure 4.29 shows a high correlation between the RFFP and both the resistance and compliance ratios, it is found that the femoral flow-rate pulse ratio shows highest correlation against the ratio of reduced network compliances. While

the compliance ratio is largest, the physical cause of this positive correlation is unclear. Intuitively it is expected that a highly compliant vessel will significantly flatten the flow-rate profile, resulting in a low flow-rate pulse. One possible hypothesis is the inverse relationship between the resistance and compliance of an arterial vessel. The compliance of an arterial vessel increases as the reference radius of the vessel increases, whereas the resistance of the vessel decreases as the reference radius increases.

The strong negative and positive correlations between the ratios of resistances and compliances to the ratios of femoral flow-rate pulses, respectively, suggests that an imbalance in the properties of opposite extremities is the cause of the over- and under-damping seen in VPs flow-rate profiles, as hypothesised. How this insufficient control over realisations can be mitigated within the already created VPD, and future VPDs can be improved from this *a posteriori* finding, are outlined in the next section.

4.9 Conclusions

A physiologically realistic virtual patient database is presented for the human arterial network. A methodology to create virtual patients guided by prior beliefs, geometrical/physiological constraints, and literature reported measurements is presented. Starting from a reference network describing the arterial network, the methodology includes: i) network reduction without compromising relevant behaviour; ii) re-parameterisation to reduce dimensionality; iii) incorporation of geometrical and physiological constraints in the form of a prior; iv) incorporation of literature reported clinical measurements in the form of the likelihood; v) combination of the prior and likelihood to generate the posterior; and vi) sampling from the posterior with MCMC to create the VPD. This generic methodology, given a mathematical description of a biological system, can be adopted to create virtual patients for any biological system while accounting for all available information.

There appears to be an imbalance in the resistances and compliances of opposite extremities in certain VPs, that results in over- and under-damped pressure and flow-rate profiles. To correct for this a post simulation filter is applied in the next chapter. The underlying physical cause of the undesirable pressure and flow-rate

profiles seen, *i.e.* the imbalance of properties in opposite extremities, is understood and so the VPD can be filtered based on this. If the VPD were to be remade the *a posteriori* finding of the need for control of the ratios of properties in opposite extremities could be directly built into the parameterisation of the arterial network, or the posterior distribution from which VPs are sampled. A possible future adaptation to the VPD—to remove the need for post simulation filters—could be to introduce correlations into the prior distributions assigned to the parameters describing opposite extremities.

Overall, the high degree of agreement between the literature reported measurements and the posterior distributions in the VPD enforce confidence in the approach taken. While some computational expense has been wasted, creating unrealistic VPs that are then filtered out of the VPD, this can i) be easily corrected if the VPD were to be remade ii) does not detract from the viability of the remaining VPs. The VPD created here can now be exploited to train and test ML classifiers in the subsequent chapters of this thesis.

Chapter 5

Pre-processing and creation of literature advised unhealthy VPDs

5.1 Introduction

In the previously completed chapter (Chapter 4) a low dimensional parameterisation is chosen to describe the arterial networks of VPs, a physiologically realistic distribution is constructed for these parameters by incorporating literature based measurements into prior distributions, and random realisations of arterial networks are sampled from this distribution. The pressure and flow-rate profiles associated with random realisations of arterial networks are computed using the physics based model of pulse wave propagation. Ideally this new VPD could directly be used to create unhealthy VPs, and then subsequently train and test ML classifiers. It is found through analysis of the *a posteriori* behaviour of the VPD, however, that certain VPs exhibit physiologically highly unlikely behaviours. It is desirable to remove these VPs from the VPD. Before the previously created VPD is used to train and test ML classifiers it is, therefore, important to apply a post simulation filter. The filtered healthy VPD can then be used to create unhealthy VPs. This chapter outlines the pre-processing carried out on the VPD created in Chapter 4, to create a data set that can be used to train and test ML classifiers.

This chapter begins by explaining how undesirable VPs are removed from the previously created VPD. It is chosen to apply a physics advised filter to the new VPD, as opposed to a hard filter as used within the PoC study (Chapter 3). To construct this physics advised filter the underlying physical cause of the physiologically highly unlikely behaviours is examined. Next, unhealthy VPs are created. As outlined in Chapter 1 the four most common forms of arterial disease are carotid artery stenosis (CAS), subclavian artery stenosis (SAS), peripheral artery disease (PAD), and abdominal aortic aneurysm (AAA). The prevalence of SAS and PAD have been recorded to vary between 1.9% and 18.83% within different demographics [66, 186], while CAS has been recorded to affect 3.8% of men and

2.7% of women [129]. The prevalence of AAA is estimated to be 4.8% [122]. The second section of this chapter outlines how VPs are created containing each of these four forms of disease. To create unhealthy VPs the likely locations and severity of disease present within stenosed and aneurysm VPs is outlined; the parameterisation of diseased vessels is presented; and the distribution of disease parameters is chosen and explained.

This chapter culminates in the analysis of the *a posteriori* behaviour of both the filtered healthy, and the unhealthy VPDs. Analysis of pressure and flow-rate profiles taken from VPs randomly sampled from the filtered healthy VPD is carried out to evaluate the ability of the physics advised filter to remove undesirable VPs. Finally the empirical distributions of pressure and flow-rate measurements across the healthy and unhealthy VPDs are analysed.

5.2 Application of physics advised filters to the healthy VPD

Through analysis of the *a posteriori* behaviour of the physiologically realistic VPD in Chapter 4, it is seen that certain VPs exhibit excessively over- and under-damped pressure and flow-rate profiles within opposite extremities. It is desirable to remove these VPs from the VPD, as they can be considered to be physiologically highly unlikely. In the PoC (Chapter 3) a hard filter, *i.e.* the direct imposition of bounds on pressure profiles, is applied. This hard filter is undesirable for several reasons:

- No consideration is given to the underlying physical cause of haemodynamic behaviours.
- No restrictions are imposed on the shape of pressure profiles, as long as the magnitudes remain within the allowable range.
- This filter is not based on any geometrical, mechanical, or physiologically restrictions. It instead employs haemodynamic expectations.

A more appropriate choice of post simulation filter to apply is a physics advised filter. In Chapter 4 the probable physical cause of the excessively over- and under-damped pressure and flow-rate profiles in opposite extremities is identified as significant

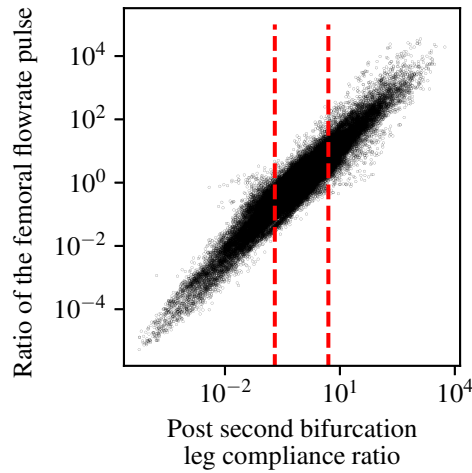


Figure 5.1: The limits imposed on the ratios of the post second bifurcation lower extremity compliances in opposite lower extremities are plotted over the ratios of compliances of all post burn-in VPs sampled from the posterior distribution with positive average flow-rate in all vessels.

asymmetry in the resistances and compliances. By imposing limits on the ratios of the resistances or compliances in opposite extremities, a filter can be constructed that alleviates the problems associated with the previous hard filter.

In the previous chapter the left-right lower extremity imbalance is assessed by reducing the network up to the second bifurcation in the leg vessels. To assess the left-right imbalance, the ratio of femoral pulse (maximum flow-rate minus the minimum flow-rate) on the left and right sides is considered. It is found that this ratio shows highest correlation against the ratio of reduced network compliances on the left and right sides. A plot of the femoral pulse ratio against the compliance ratio is shown in Figure 5.1. To limit the left-right imbalance, a filter on the left-to-right compliance ratio between 0.2–5 is introduced. These limits are shown as vertical lines in Figure 5.1 and constrains the femoral ratio between 10^{-2} and 10^2 across the data set. With this filter, approximately 45% (23,275 patients) of the patients are discarded, leaving with 28,868 physiologically realistic patients in the VPD.

5.3 Creation of unhealthy VPDs

A filtered physiologically realistic VPD containing healthy subjects is created in Chapter 4 and Section 5.2, and forms the starting point for diseased VP creation. The arterial network contains 71 vessel segments and is shown in Figure 5.2, along with the locations where disease occurs in high prevalence (see Chapter 4 for details). The filtered healthy VPD contains 28,868 VPs and is referred as VPD_H . Disease is introduced into these healthy arterial networks as described next.

5.3.1 Disease forms

The four most common forms of arterial disease are carotid artery stenosis (CAS), subclavian artery stenosis (SAS), peripheral arterial disease (PAD, a form of stenosis), and abdominal aortic aneurysm (AAA) [53, 115, 1, 39, 122]. Their prevalence is restricted to the following vessels (see Chapter 4 for details) and shown in Figure 5.2:

- **CAS** is assumed to only affect the common carotid arteries. For simplification and consistency of notation these vessels are referred to as the **carotid artery chains** (CA_x).
- **SAS** is assumed to affect the first and second subclavian segments. These two chains of vessels (one on the right and left side) are referred to as the **subclavian artery chains** (SA_x).
- **PAD** is assumed to affect the common iliacs; external iliacs; first and second femoral segments; and the first popliteal segments. These chains are referred to as the **peripheral artery chains** (PA_x).
- **AAA** is assumed to affect the first to fourth abdominal aorta segment. This chain of vessels is referred to as the **abdominal aortic chain** (AA_x).

It is assumed that each diseased VP has only one of the four forms of arterial disease. Four complementary databases corresponding to VPD_H are constructed, each pertaining to one form of arterial disease. To create the diseased VPD corresponding to CAS, referred to as VPD_{CAS} , for every subject in VPD_H , disease is

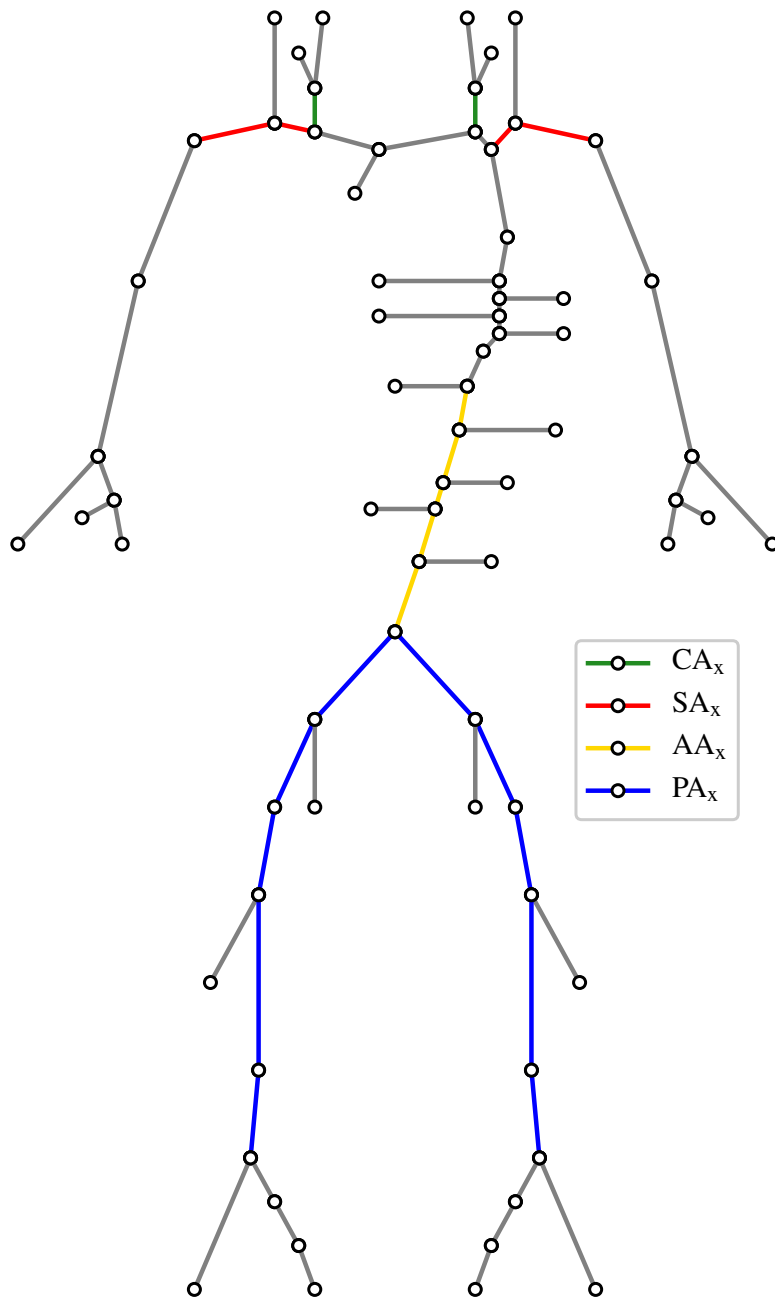


Figure 5.2: The connectivity of the arterial network. The location of the four forms of disease (see Chapter 4) are highlighted.

introduced in CA_x (*i.e.* the left or right carotid artery). This is achieved by taking the arterial network of a subject from VPD_H , artificially introducing a stenosis in CA_x , and then re-running the pulse wave propagation model to compute the pressure and flow-rate waveforms. Thus, VPD_{CAS} contains 28,868 VPs with CAS. Similarly, the databases corresponding to SAS, PAD, and AAA are created, and referred to as VPD_{SAS} , VPD_{PAD} , and VPD_{AAA} , respectively. The disease severities, locations, and shapes are varied randomly across these databases as described next.

5.3.2 Severities of diseases

It must be ensured that the created unhealthy VPDs encompass not only the full range of highly likely disease locations, but also the likely range of disease severities. When deciding upon the range of disease present within the unhealthy VPDs the balance between variability and control must be carefully considered. The magnitude of any biomarkers introduced into pressure and flow-rate profiles, and so likely ease of detection, is likely to be proportional to the severity of disease present—*i.e.* more significant changes to the area of a vessel are likely to produce more significant changes to the pressure and flow-rate profiles of blood.

Stenosis severities

The severity of stenoses (percentage reduction in area) is varied between 50% and 95%. The lower 50% limit is set for the stenoses to be haemodynamically significant [1, 197] and the upper limit of 95% reflects near total occlusion. Stenosed VPs are not created with a complete occlusion, although this behaviour can physiologically occur [53], as VPs with a complete blockage can obviously and easily be identified by a lack of downstream flow-rate.

Aneurysm severities

Determination of the range of severities of disease to include in VPD_{AAA} is more difficult than in the case of stenosed VPDs, as there are no fixed geometrical restrictions on the maximum increase in vessel area. It is stated in [59] that only 2% of AAA with diameter less than 4cm rupture, and so these aneurysms are unlikely to be operated on, while 25%–41% of AAAs with diameter greater than 5cm rupture

5.3. CREATION OF UNHEALTHY VPDS

within 5 years. The current AAA screen program offered by the United Kingdom's National Health Service splits patients into four categories, dependent on the severity of the AAA present [48]:

- If the aorta of a patient is found to have a diameter less than 3cm, the patient is discharged with no future scanning carried out.
- If the aorta of a patient has a diameter between 3cm and 4.4cm the patient is invited back for a follow up scan in 1 years time.
- If the aorta of a patient has a diameter between 4.5cm and 5.4cm the patient is invited back for a follow up scan in 3 months time.
- If the diameter of a patient's aorta is found to be greater than 5.4cm the patient is invited to undergo surgery.

A further studying breaking down patients into categories, based on the severity of any AAAs present, is [44]. In [44] patients with AAA are split into three categories:

- Patients with a AAA with diameter less than 2.6cm are discharged.
- Patients with a AAA with diameter between 2.6cm–5.4cm are monitored.
- Patients with a AAA with diameter greater than 5.4cm are treated.

Based on these studies an allowable range of AAA severities of 4cm–6cm diameters is chosen. Assuming that arterial vessels are cylindrical, this severity range results in a vessel cross sectional area range of 12.56cm^2 – 28.27cm^2 . The diameter of the abdominal aorta in the reference network in Chapter 4 has a range of between 1.5cm–1.18cm, resulting in a cross sectional area range of between 1.76cm^2 – 1.09cm^2 . The minimum severity of AAAs to be included within unhealthy VPs is computed by finding the required increase in area to produce the minimum clinically significant cross sectional area, *i.e.* 12.56cm^2 , from the maximum area of the reference arterial network, *i.e.* 1.76cm^2 . The maximum severity of AAA to be included within unhealthy VPs is computed by finding the required increase in area to produce the maximum clinically significant cross sectional area, *i.e.* 28.27cm^2 , from the minimum area of the reference arterial network, *i.e.* 1.09cm^2 . The corresponding AAA severities are set to vary between 713% ($12.56/1.76$) and 2,593% ($28.27/1.09$).

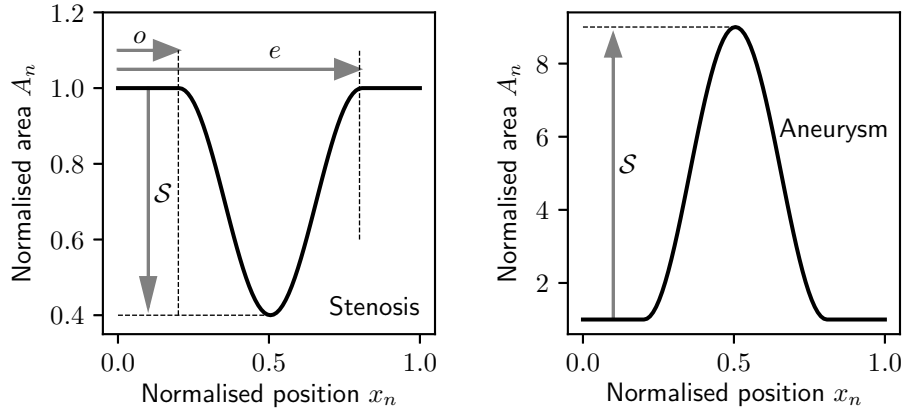


Figure 5.3: An example of a stenosis of severity 0.6 and aneurysm of severity 8.0 are shown. These disease profiles are created with a start location of 0.2 and an end location of 0.8.

5.3.3 Parameterisation of diseased vessels

With the above ranges, parameterisation of area increase/reduction proposed in the PoC is adopted (see Chapter 3 for details), see Figure 5.3. For a chain of diseased vessels (CA_x , SA_x , PA_x , or AA_x), the normalised area A_n as a function of the normalised x-coordinate, x_n , is represented as:

$$A_n = \begin{cases} \left(1 \mp \frac{\mathcal{S}}{2}\right) \pm \frac{\mathcal{S}}{2} \cos\left(\frac{2(x_n - o)\pi}{e - o}\right) & \text{for } o \leq x_n \leq e \\ 1 & \text{otherwise} \end{cases} \quad (5.1)$$

where \mathcal{S} represents the severity, o represents the normalised starting location of the disease in the vessel chain, e represents the normalised end location, A_n is normalised with respect to the healthy version of the vessel in VPD_H , and \pm creates an aneurysm or stenosis, respectively. In CA_x , SA_x , and PA_x , the left and right side vessels are chosen with equal probability.

The disease severity \mathcal{S} , start location o , and end location e are assigned uniform distributions based on physical considerations. To sample values for these parameters, a fourth parameter, the reference location of the disease (represented by r) is introduced. This is included to impose a minimum length of 10% of the chain length on the disease profiles. Thus, the parameters for disease are sampled

sequentially from uniform distributions within the following bounds:

$$\text{Bounds: } \left\{ \begin{array}{l} 0.2 \leq r_e \leq 0.8, \\ 0.1 \leq o \leq r_e - 0.05, \\ r_e + 0.05 \leq e \leq 0.9, \\ \left\{ \begin{array}{ll} 0.5 \leq \mathcal{S} \leq 0.95 & \text{stenosis} \\ 7.13 \leq \mathcal{S} \leq 25.93 & \text{aneurysm.} \end{array} \right. \end{array} \right. \quad (5.2)$$

Based on the above parameterisation, examples of healthy and diseased SA_x , PA_x , and AA_x area profiles are shown in the left and right columns of Figure 5.4, respectively.

5.4 Results and discussion

5.4.1 Evaluation of pressure and flow-rate profiles measured within individual VPs.

Similar to the analysis carried within Chapter 4, 15 VPs are randomly sampled from VPD_H (containing 28,868 usable VPs). The pressure and flow-rate profiles associated with these randomly sampled VPs are plotted, to allow for evaluation of the behaviour of individual VPs in VPD_H . Pressure profiles are examined at the ascending aorta; right and left radial arteries; and right and left common carotid arteries. Flow-rate profiles are examined at the right and left second femoral segment. The pressure and flow-rate profiles associated with the 15 VPs randomly sampled from VPD_H are shown within Figures 5.5, 5.6, and 5.7.

Looking at the three figures listed above, it is seen that there now appears to be more consistency in the pressure profiles within the radial arteries (shown in plots (b) and (c) of the three figures), than was seen in the unfiltered VPD. While some oscillatory behaviour is seen in the radial pressure profiles (Patient-E, Patient-F, and Patient-O) the amplitude of these oscillations is much lower than previously observed (see Figure 4.28 for the original profiles).

Figures 5.5, 5.6, and 5.7, show that the some under- and over-damped femoral flow-rate profiles are still occurring. Previously, however, the occurrence of under- and over-damped profiles was asymmetrical. There appears to be a higher degree of

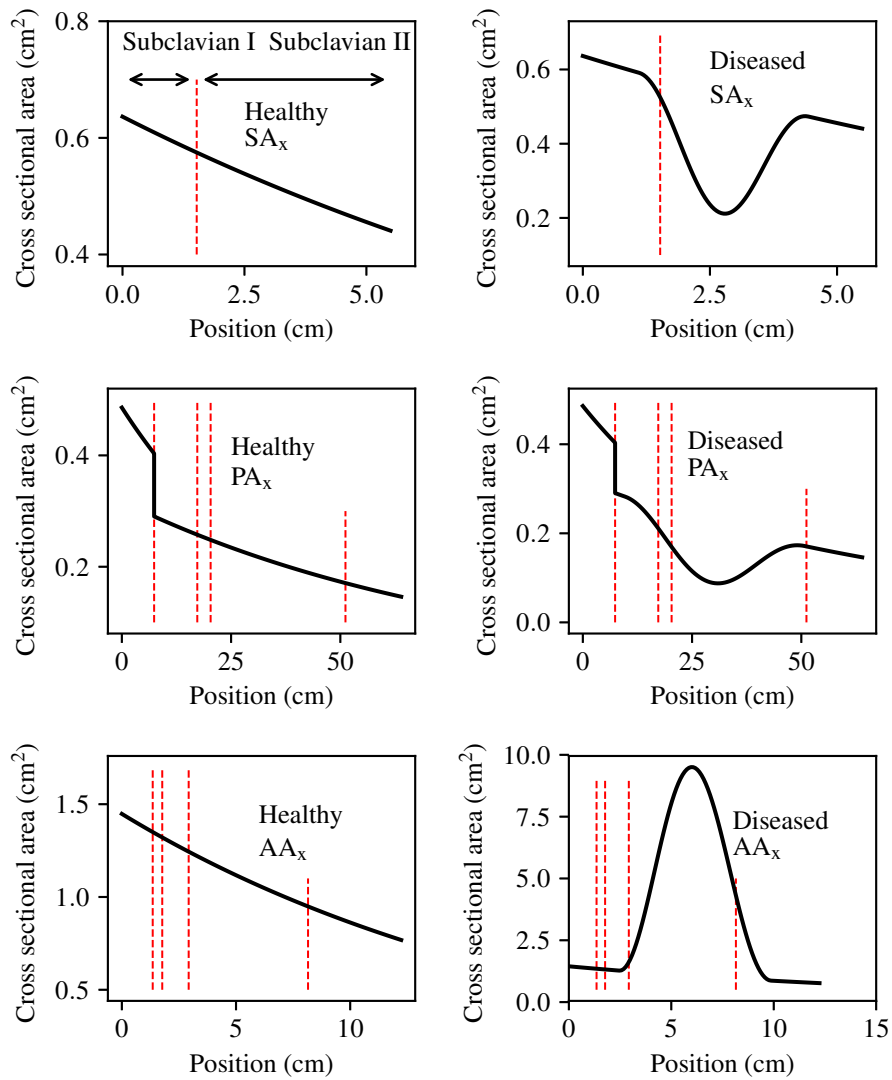


Figure 5.4: Examples of healthy and diseased SA_x, PA_x, and AA_x area profiles. The geometrical boundaries between vessel segments that form the chains are indicated by red dashed lines. See Section 5.3.1 for details of the vessels in each chain.

symmetry between the flow-rate profiles in opposite extremities in VPD_H . Under-damped flow-rate profiles are seen to be occurring in both the right and left lower extremities in VPs Patient-B (shown in Figure 5.5), Patient-L (shown in Figure 5.7), and Patient-N (shown in Figure 5.7). Over-damped flow-rate profiles are seen within both the right and left lower extremities of VPs Patient-A (shown in Figure 5.5), Patient-H (shown in Figure 5.6), and Patient-O (shown in Figure 5.7). Overall, it appears as if the physics advised filter has successfully served its purpose of removing VPs that exhibit excessive asymmetrical profiles in opposite extremities.

5.4.2 Analysis of scalar pressure and flow-rate distributions

In Chapter 4 histograms of the distributions of systolic and diastolic pressure in the radial arteries, ascending aorta, and common iliac arteries; and average flow-rate in the second femoral segments are shown. These histograms are now recreated first for the filtered healthy VPD (VPD_H), and then for each of the five VPDs (VPD_H , VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , and VPD_{AAA}) created within this chapter. The distributions of pressure and flow-rate measurements across VPD_H (overlaid with the literature based measurements built into the likelihood term in Section 4.6.2) are shown in Figure 5.8. The histograms of the distributions of measurements across all five of the VPDs are shown in Figure 5.9.

Comparing the distributions of measurements across the filtered (Figure 5.8) and unfiltered (Figure 4.24) healthy VPDs, it is seen that there is very little visual difference between the distributions of pressure measurements across each of the two databases. The most obvious difference between the two is seen in the distribution of average femoral flow-rate. There is a significant reduction in the proportion of VPs belonging to the lowest flow-rate bin (with a lower band of 0ml/sec) in the filtered VPD relative to the unfiltered VPD. This suggests that the post simulation filter is able to remove a large proportion of the VPs with very low (near zero) flow-rate in either of the two lower extremities.

The differences between the distributions of measurements taken from each of the five VPDs (as shown in Figure 5.9) are expected to be minor. The indicative biomarkers of arterial disease captured within pressure and flow-rate profiles are expected to consist of micro inter- and intra-measurement details, and so there is not expected to be significant differences to these discrete pressure and flow-rate

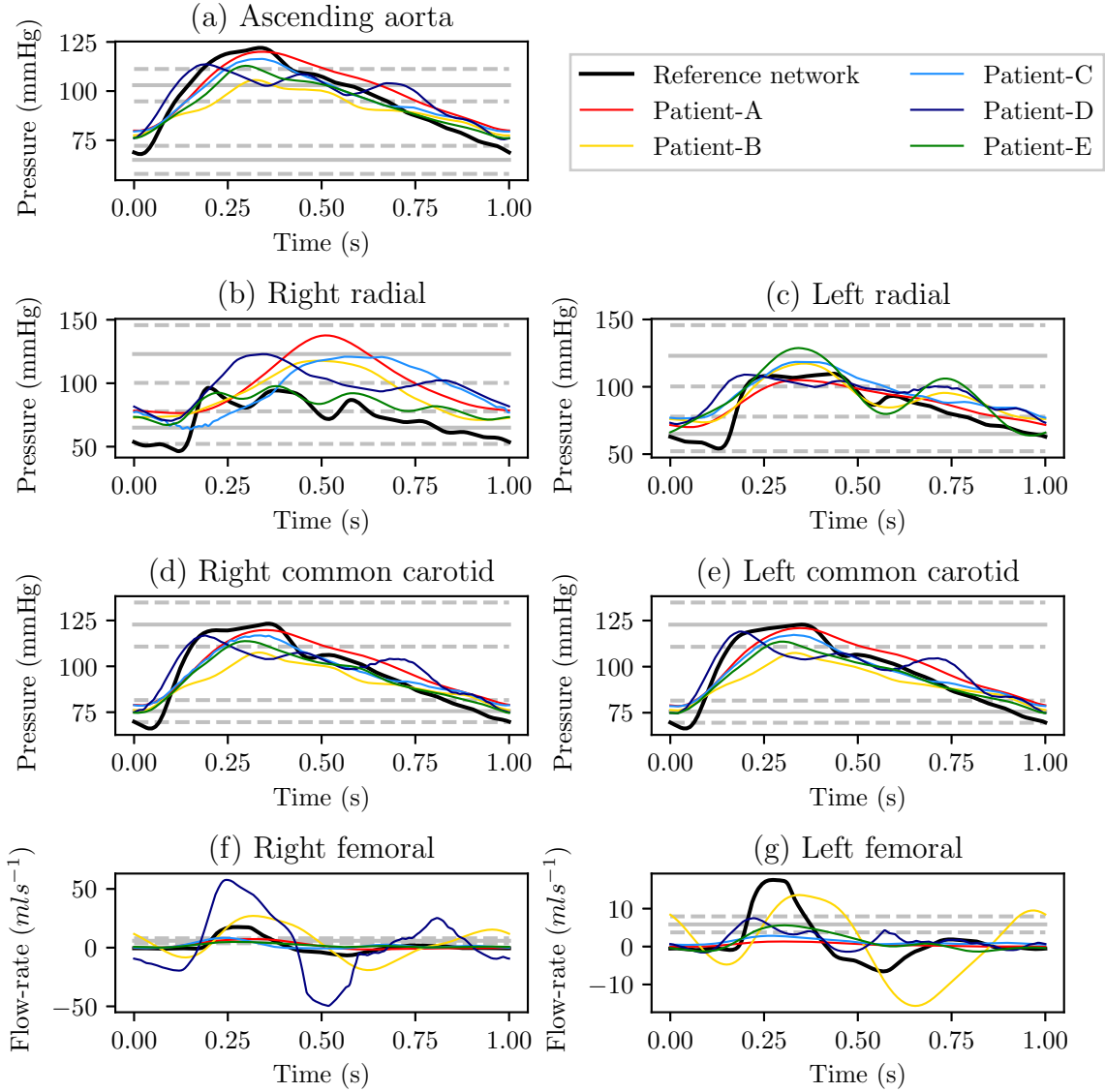


Figure 5.5: The subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively (see Chapter 4 for details).

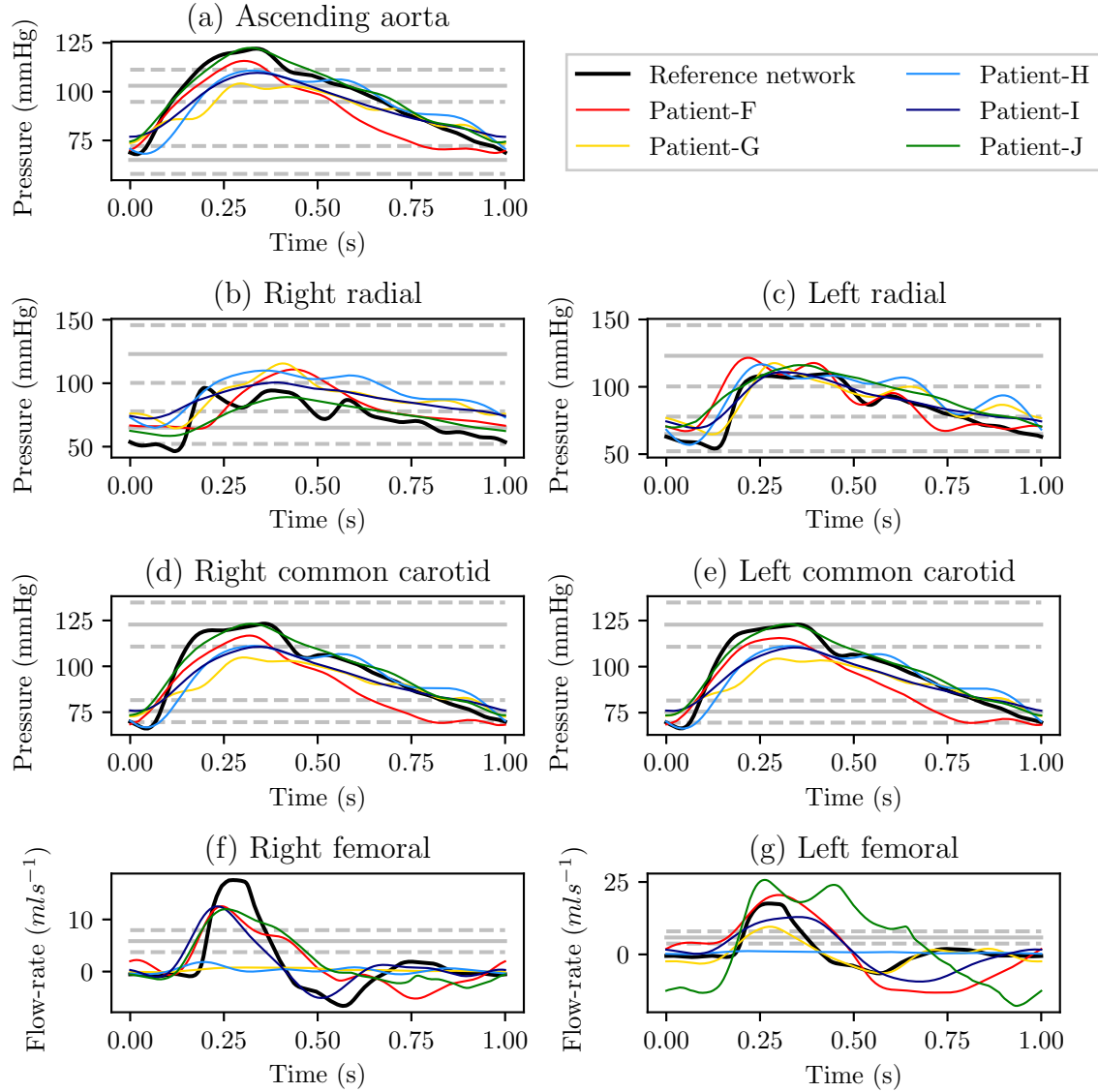


Figure 5.6: The subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively (see Chapter 4 for details).

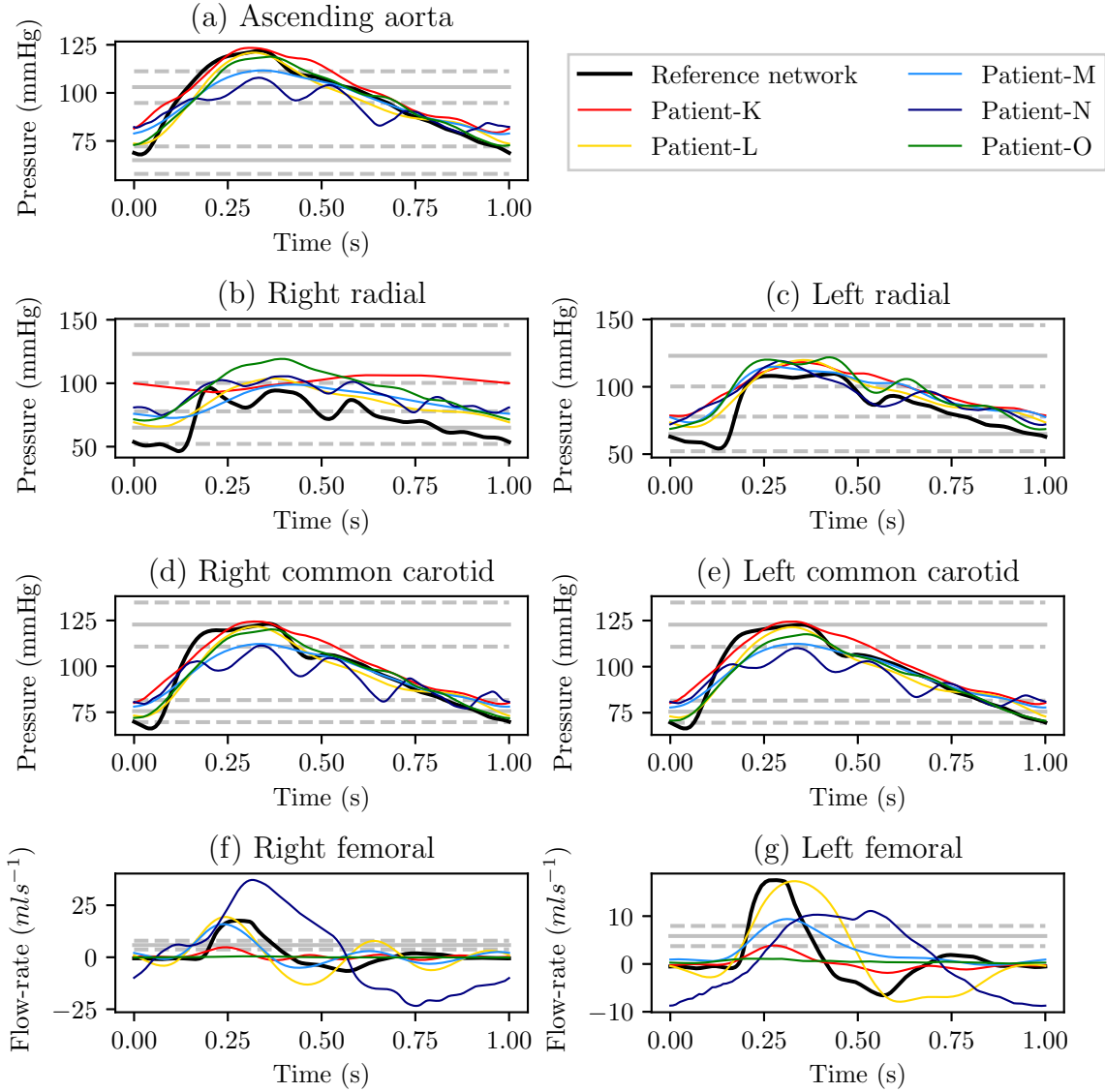


Figure 5.7: The subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively (see Chapter 4 for details).

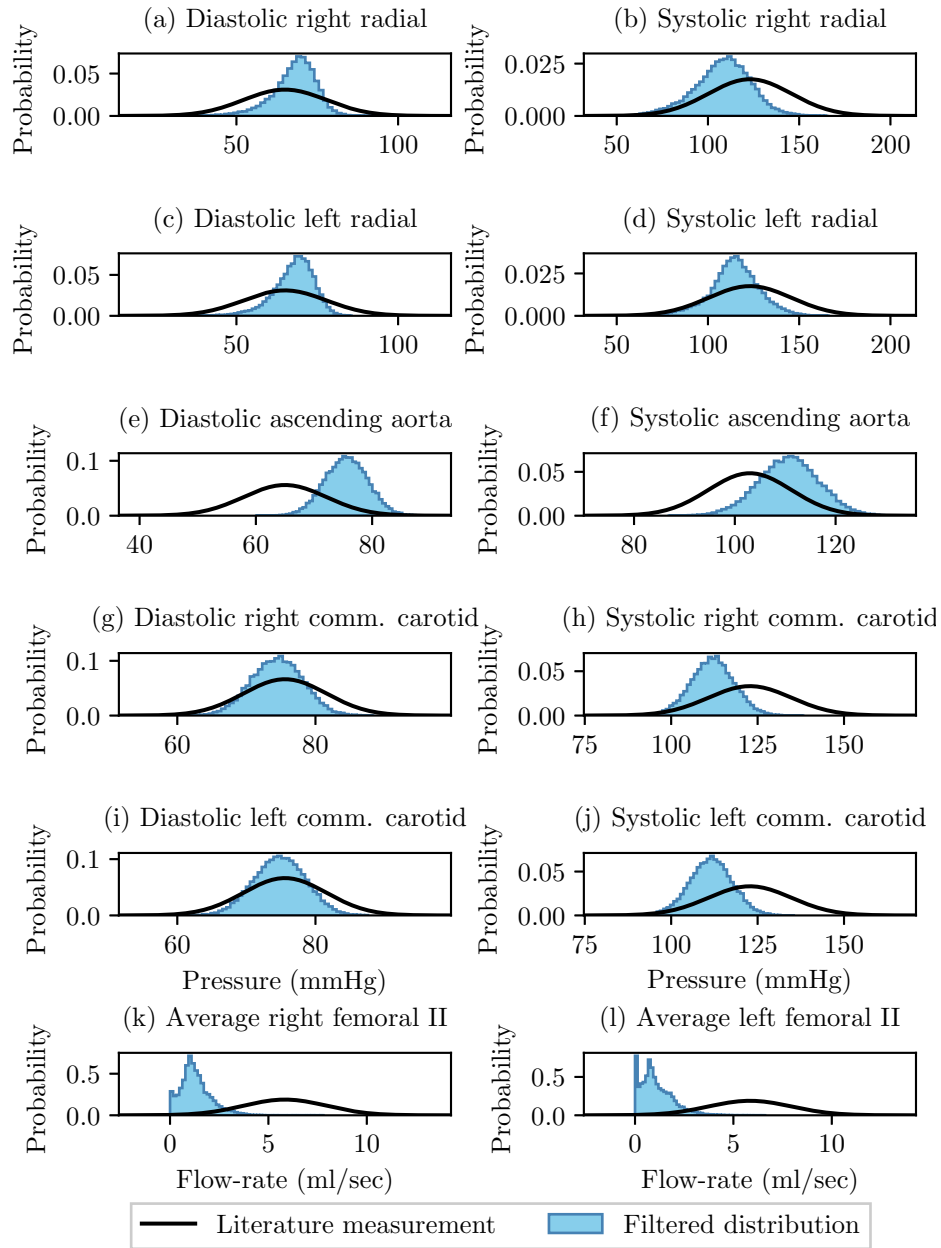


Figure 5.8: Histograms of the filtered distributions of the pressure and flow-rate measurements at all measurement locations. The literature based measurements and associated error distribution at each location are overlaid in black. Diastolic and systolic pressure in the right radial artery are shown in (a) and (b), respectively; the diastolic and systolic pressure in the left radial artery are shown in (c) and (d), respectively; the diastolic and systolic pressure in the ascending aorta are shown in (e) and (f), respectively; the diastolic and systolic pressure in the right common carotid artery are shown in (g) and (h), respectively; the diastolic and systolic pressure in the left common carotid artery are shown in (i) and (j), respectively; and the average flow-rate in the second segments of the right and left femoral artery are shown in (k) and (l), respectively.

measurements.

Figure 5.9 shows that there does not appear to be any significant differences between the distributions of measurements taken from VPD_H , VPD_{CAS} , VPD_{SAS} , and VPD_{PAD} . While there are some minor differences between the distributions measured in each of these four VPDs, there does not appear to be any consistent or significant patterns present. This suggests that, as is *a priori* expected, the haemodynamic affects of arterial disease can not be sufficient captured by scalar pressure and flow-rate measurements.

Greater distinction can, however, be made between the distribution of measurements taken from VPD_{AAA} . Figure 5.9 shows a consistent increase in diastolic pressure—shown in plots (a), (c), (e), (g), and (i)—and a decrease in the systolic pressure—shown in plots (b), (d), (f), (h), and (j)—in VPD_{AAA} , relative to VPD_H . This suggests that the pressure profiles of VPs in VPD_{AAA} are being flattened. This finding suggests strong potential for the use of pressure and flow-rate profiles to detect AAAs, as physical changes are producing consistent haemodynamic differences.

5.5 Conclusions

In this chapter pre-processing has been carried out on the physiologically realistic VPD previously created in Chapter 4 to allow for the training and testing of ML classifiers. A physics advised filter has been applied to the raw VPD to remove VPs exhibiting physiologically unlikely behaviours. This filter removes VPs with highly asymmetric compliances in the lower extremities. Before filtering, the VPD contained 52,143 (75,000-10,000-12,857) VPs. The physics advised filter removes 23,275 VPs, and the remaining 28,868 VPs are referred to as VPD_H .

Four complimentary VPDs are then created for four of the most common forms of arterial disease (CAS, SAS, PAD, and AAA). Each of the four unhealthy VPDs contains 28,868 VPs that share identical underlying arterial networks to the corresponding parent healthy VP, however with the inclusion of disease in the appropriate location. The pressure and flow-rate profiles associated with these unhealthy VPs are computed using the previously employed physics based model of pulse-wave propagation.

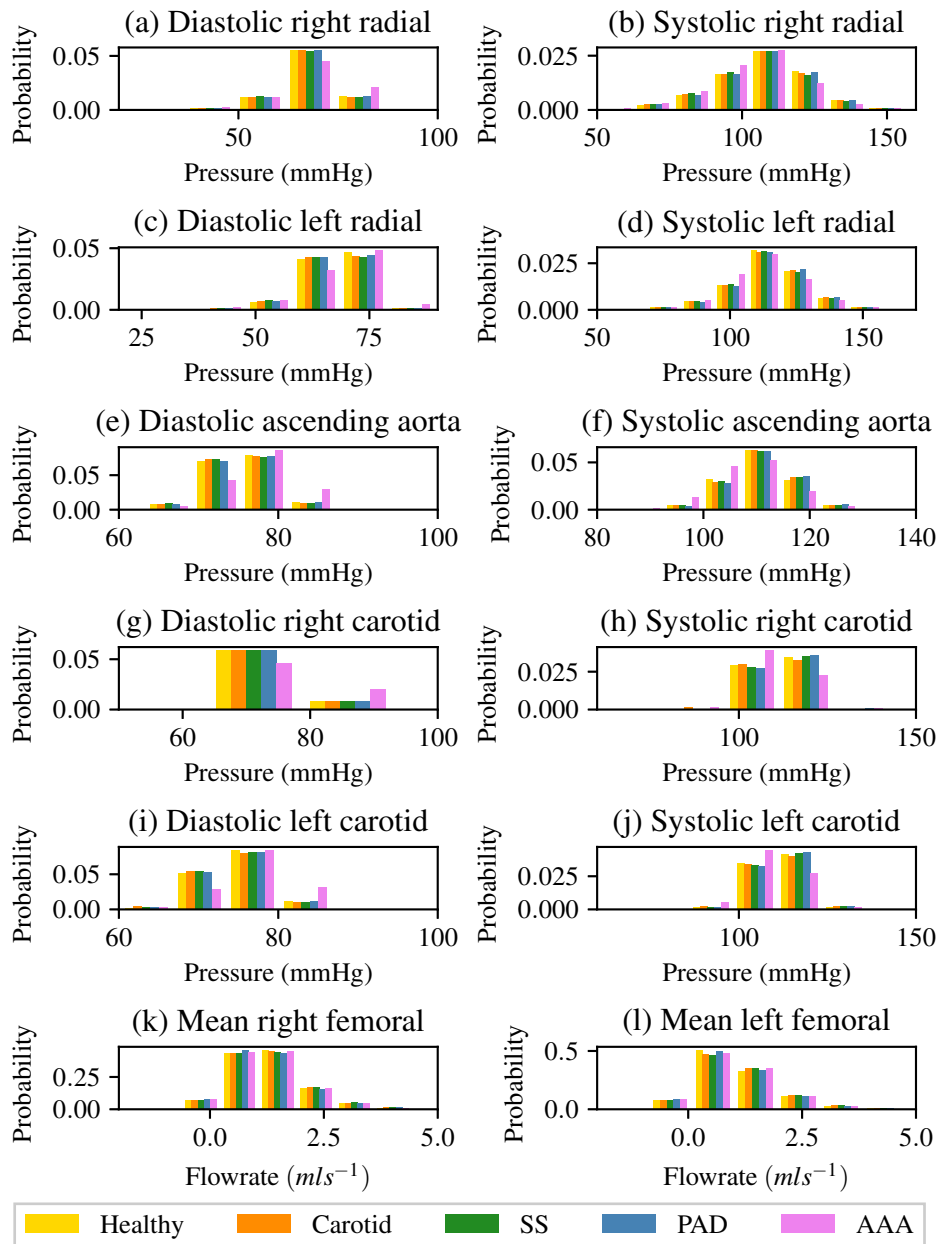


Figure 5.9: The histograms of the scalar pressure and flow-rate measurements taken from each of the five VPDs are shown above.

The five VPDs presented within this chapter (VPD_H , VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , and VPD_{AAA}) can now be used to train and test ML classifiers. The subsequent chapters of this thesis use the processed healthy and unhealthy data to analysis the ability of ML classifiers to detect arterial disease.

Chapter 6

Application of machine learning classifiers to the physiologically realistic VPDs

6.1 Introduction

In the previous chapter of this thesis a usable healthy VPD, referred to as VPD_H , is created by applying physics advised filters to the raw VPD created in Chapter 4. VPD_H is then used to create four complimentary VPDs containing VPs with the presence of carotid artery stenosis (CAS), subclavian artery stenosis (SAS), peripheral artery disease (PAD), and abdominal aortic aneurysm (AAA). These four unhealthy VPDs are referred to as VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , and VPD_{AAA} , respectively.

This chapter now utilises the aforementioned VPDs to train and set a series of ML classifiers to predict the presence of the four forms of disease, using easily acquirable peripheral measurements. The exploratory stance of the previously completed proof-of-concept (PoC)—see Chapter 3—is adopted in this study. The primary objective of this chapter is to extend the work presented in the PoC to a more physiologically realistic data set. Focus is given to uncovering patterns and behaviours in classification accuracy when using varying haemodynamic measurements, and under various conditions (such as when the severity of disease is reduced) to gain an understanding of both the potential and limitations of the proposed method.

This chapter begins by explaining the methods by which ML classifiers are applied to the five VPDs, including: the available measurements, the configuration of classifiers used, and the optimisation of hyper-parameters. Next the results achieved are presented: ML classifiers are trained and tested using various combinations of pressure and flow-rate measurements; and additional, complimentary, analysis is performed. Finally, this chapter concludes by evaluating what has been discovered about arterial disease detection through easily acquirable measurements.

The majority of work in this chapter is published in: *Jones, G., Parr, J., Nithiarasu, P., Pant S. Machine learning for detection of stenoses and aneurysms: application in a physiologically realistic virtual patient database. Biomechanics and Modeling in Mechanobiology. 2021 July <https://doi.org/10.1007/s10237-021-01497-7>.*

6.2 Methodology

The methods by which ML classifiers are applied to the five VPDs (VPD_H, VPD_{CAS}, VPD_{SAS}, VPD_{PAD}, and VPD_{AAA}) created in Chapter 5 is presented here.

6.2.1 Available measurements

A review of potential measurements that can be acquired in the network is presented in Chapter 4. Based on this, the locations at which time-varying pressure and flow-rate measurements can be acquired are shown in Figure 6.1 and described below. The location of each disease chain (as described in Chapter 5) are also shown in Figure 6.1. The available measurements are:

- **Pressure in the carotid and radial arteries** measured using applanation tonometry [3, 146]. To simplify annotation and description the right and left carotid artery pressures are referred as $P_1^{(R)}$ and $P_1^{(L)}$, respectively. Similarly, the radial artery pressures are referred to $P_3^{(R)}$ and $P_3^{(L)}$, respectively.
- **Pressure in the brachial arteries** estimated through reconstruction of finger arterial pressure [77]. The right and left brachial artery pressures are referred to as $P_2^{(R)}$ and $P_2^{(L)}$ respectively.
- **Flow-rate in the carotid, brachial, and femoral arteries** measured using Doppler ultrasound [27, 144, 164]. The right and left carotid, brachial, and femoral artery flow-rates are referred to as $Q_1^{(R)}$, $Q_1^{(L)}$; $Q_2^{(R)}$, $Q_2^{(L)}$; and $Q_3^{(R)}$, $Q_3^{(L)}$, respectively.

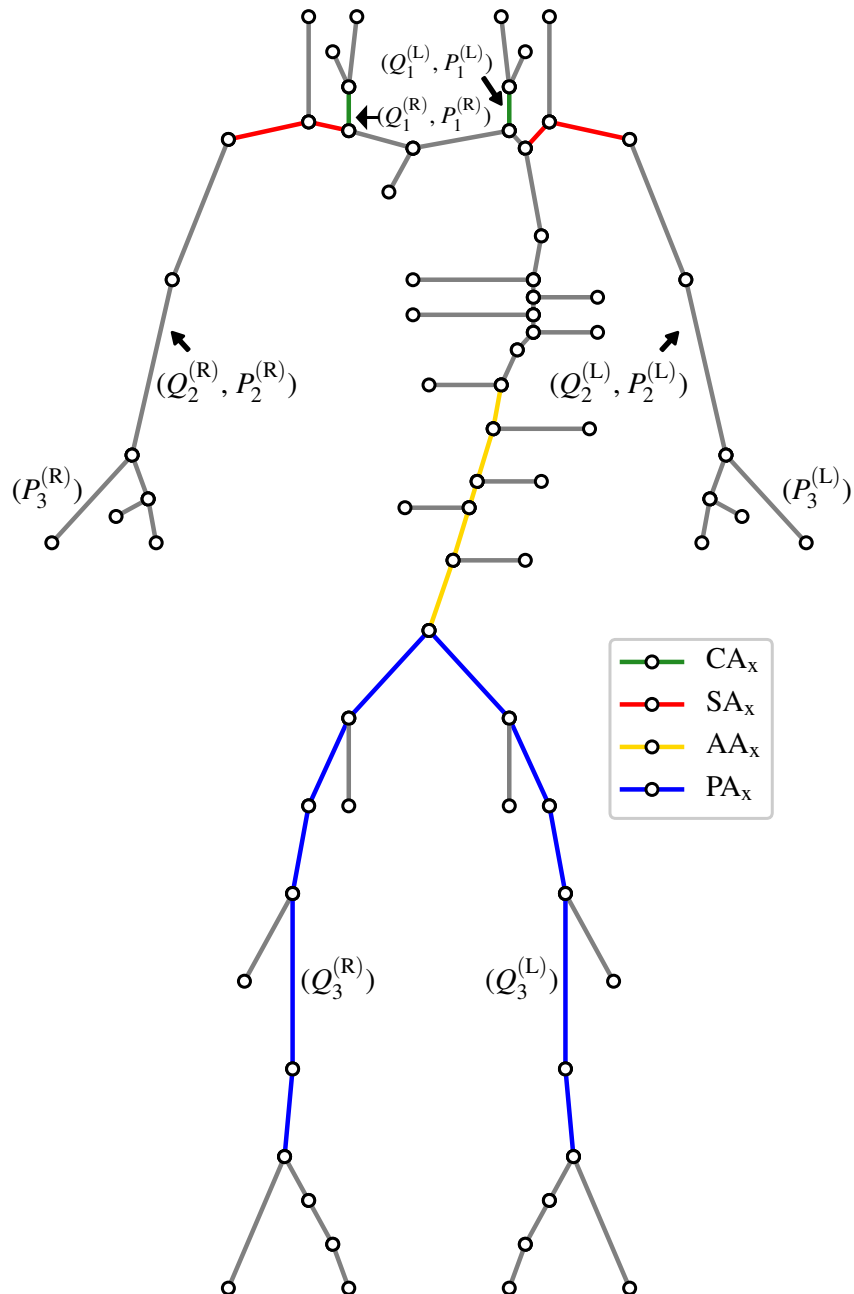


Figure 6.1: The connectivity of the arterial network. The location of the four forms of disease (see Chapter 5); and six pressure and flow-rate measurements (see Section 6.2.1) are highlighted.

Provision of measurements to ML classifiers

Unless specified otherwise, the measurements to ML classifiers are bilateral, *i.e.* when Q_1 is specified it is implied that both right and left carotid flow-rates are used:

$$Q_1 = \{Q_1^{(R)}, Q_1^{(L)}\}. \quad (6.1)$$

There are, therefore, six bilateral measurements—three pressure and three flow-rate. To reduce the dimensionality required to describe each, the periodic profiles are described through a Fourier series (FS), as employed in Chapter 3. Thus, each individual measurement is described by 11 FS coefficients, and each bilateral measurement by 22 FS coefficients (see Chapter 3 for further detail).

6.2.2 Machine learning classifiers

As previously outlined in Chapter 3 a model mapping a vector of input measurements, \mathbf{y} , to a discrete output classification, z , can be described as:

$$z = m(\mathbf{y}) \quad z \in \{\mathcal{C}^{(1)}, \mathcal{C}^{(2)}\}, \quad (6.2)$$

where $\mathcal{C}^{(j)}$ represents the j^{th} possible classification. In the context of this chapter, the measured inputs, \mathbf{y} , represents the FS coefficients of a user defined combination of the haemodynamic measurements $\{Q_1, Q_2, Q_3, P_1, P_2, P_3\}$ (see Section 6.2.1) taken from VPs, and the output classification represents the corresponding health of those VPs : $\mathcal{C}^{(1)}$ = ‘healthy’ and $\mathcal{C}^{(2)}$ = ‘diseased’. As with the PoC study, to account for large differences in magnitudes of the components of \mathbf{y} , they are individually transformed with the Z-score standardisation method [136] to have zero-mean and unit variance.

As stated in Chapter 5, it assumed that disease is limited to only one of the four forms in each patient. As a first exploratory study, the ML classifiers are created for each form independently. All classifiers are therefore binary (see Chapter 3 for details), *i.e.* four independent classifiers are trained to predict the following questions independently: “Does a VP belong to VPD_H or VPD_x ”, where x can be either CAS, SAS, PAD, or AAA.

Training and test sets

Each VP in VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , and VPD_{AAA} shares an identical underlying arterial network, apart from the diseased chain, with the corresponding healthy subject in VPD_H . It is, therefore, important to ensure that the same subset of VPs is not included in both their healthy and diseased forms in the data set used for ML classifiers. As each form of disease is mutually exclusive, four independent training and test sets, each corresponding to one form of disease, are constructed in the following three stages:

- **Step 1:** Half of the available VPs are randomly selected from VPD_H for inclusion within the ML data set; this is referred to as VPD_{H-ML} . The unhealthy VPs corresponding to the remaining unused half are taken from the appropriate unhealthy VPD (VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , or VPD_{AAA}) and incorporated into the ML data set. These data sets are referred to as VPD_{CAS-ML} , VPD_{SAS-ML} , VPD_{PAD-ML} , or VPD_{AAA-ML} .
- **Step 2:** The data sets of Step 1 are combined to create four complete data sets each containing 50% healthy and 50%, unhealthy VPs:
 1. $VPD_{H-ML} \cup VPD_{CAS-ML}$
 2. $VPD_{H-ML} \cup VPD_{SAS-ML}$
 3. $VPD_{H-ML} \cup VPD_{PAD-ML}$
 4. $VPD_{H-ML} \cup VPD_{AAA-ML}$
- **Step 3:** The four data sets of Step 2 are randomly split into a training set containing 2/3 of all the VPs in the data set, and a test set containing 1/3 of all the VPs.

The performance of all ML classifiers is evaluated using a five fold validation (see Chapter 3 for details). For each fold, the same data set from Step 2 is used but different subsets are sampled in Step 3 for training and testing.

ML methods

In this chapter all six ML methods (NB, LR, MLP, SVM, GB, and MLP) outlined in Chapter 2 are employed. Note that the last of these, the multi-layer perceptron,

Modelling approach	Non-probabilistic	Probabilistic
Tree-based	RF	GB
Kernel-based	SVM	
Bayesian		NB
Neuron-based		LR, MLP

Table 6.1: The four different modelling approaches and how each classification method aligns with these approaches, are outlined.

may be considered as a deep learning method. These methods are chosen as they encompass a range of probabilistic and non-probabilistic applications of different modelling approaches, see Table 6.1, while requiring minimal problem specific optimisation. All implementations of the above algorithms in the Python package Scikits-learn [156] are used. Some of these methods require optimisation of the hyper-parameters. This is described in Section 6.2.4.

Quantification of results

Classifier performance is assessed by the same two metrics employed in Chapter 3: *sensitivity* and *specificity* in combination; and the F_1 score. It is desirable to have both sensitivities and specificities to be high. Similarly, a higher F_1 score is desirable. Since the F_1 score is a single scalar metric that balances both precision and recall, it is a good metric to compare classifiers when tuning the hyper-parameters of ML algorithms. For a discussion on these metrics and their relevance, refer to Chapter 3.

6.2.3 Adequacy of the number of VPs within the VPD

In the PoC study the number of VPs within the VPD is assessed from both an *a priori* and *a posteriori* stand point. The same analysis is not performed in this large network study. In the PoC, as independent sampling is performed, the number of VPs drawn from the distribution of arterial network parameters is purely driven from a ML perspective, *i.e.* how many VPs are required to train and test ML classifiers. In contrast, the number of VPs sampled from the posterior distribution in Chapter 4 is chosen based on statistical considerations, *i.e.* the likely number of samples for

the MCMC method to approximate the posterior distribution. The 28,868 VPs in the data sets outlined in Section 6.2.2 have an EPV of 145 ($28,868 \div 2 \div 3/2 \div 66$) when used to train binary ML classifiers. This EPV is significantly higher than the rule of thumb value of 10, and thus it is assumed that there are sufficient VPs to train and test ML classifiers. It is possible to reduce the computational expense associated with creating ML classifiers by finding the optimum number of VPs, *i.e.* removing excessive VPs from the VPD, however this process is unpragmatic for the following reasons:

- As the ML methods used are relatively computationally inexpensive, relative to other methods such as deep learning, the reduction in computational expense achieved by reducing the number of VPs is unlikely to be relatively small.
- Determination of the optimum number of VPs is likely to be computationally expensive in its self. A series of ML classifiers would need to be trained using varying numbers of VPs for each form of disease, using each ML method.

Due to the aforementioned reasons, it is chosen to train and test ML classifiers using the full VPD.

6.2.4 Hyper-parameter optimisation

The architecture of LR, NB, and SVM classifiers can all be considered to be problem independent. While these three algorithms are able to undergo varying levels of problem specific optimisation, the underlying structure of the classifier usually does not change. The architectures of RF, MLP, and GB classifiers, however, are dependent on the specific problem. The architecture choices for the classifiers and associated hyper-parameter optimisation is described next. For all six methods, all non-specified hyper-parameters are left in their default Scikits-learn states.

LR, SVM, and NB

For LR, the ‘LIBLINEAR’ solver offered by the Scikits-learn [156] package is chosen. In the case of SVM, a kernel is typically chosen to map the input measurements to a higher order feature space [94]. All SVM classifiers use a radial basis function kernel [184], with the Scikits-learn hyper-parameter ‘gamma’ set to ‘scale’. In the

case of NB, the distribution of input measurements across the data set is chosen to be Normal [139].

Random Forest

In the case of RF, the number of trees in the ensemble and the maximum depth of each tree is optimised. Other hyper-parameters that can be tuned include the minimum number of data points allowed in a leaf node, and the maximum number of different features considered for splitting each node—although the affect of these are not investigated here as they are likely to be less significant. To optimise the two hyper-parameters, a grid search is carried out. A grid is constructed by discretising the possible number of trees within the ensemble between 10 and 400 at intervals of 10; and the possible depth of each tree between 20 and 200 at intervals of 10. RF classifiers are trained using every combination of number of trees and depth of trees, with all six pressure and flow-rate measurements (see Section 6.2.1) across all the four forms of arterial disease. The hyper-parameters describing the architecture that produces the highest F_1 score is found for each form of disease, and this combination of hyper-parameters is then chosen for all subsequent classifiers. The optimal hyper-parameters for each of the four forms of disease are shown in Table 6.2, along with the F_1 score achieved by each. Contour plots of the F_1 scores achieved when using each combination of number of trees and depth of each tree are shown in Appendix J.

It is unlikely that a single architecture will consistently produce the highest accuracy when varying the combination of input measurements. In this study, re-optimisation of the hyper-parameters when varying the input measurement combination is avoided to minimise computational cost. It is found that when all six input measurements are used the F_1 score produced is relatively insensitive to the hyper-parameter combination (see Appendix J for details). Thus, it is likely that a reasonable estimation of the maximum achievable accuracy can be obtained using a fixed architecture when varying the input measurements. It should be noted, however, that some improvements in classification accuracy may be possible with such re-optimisation.

In this study the available data set is partitioned into two components—a training set and a test set. It is more common, however, for data sets to be

Disease	Trees	Depth	F_1
CAS	100	80	0.8878
SAS	150	80	0.8292
PAD	100	100	0.8935
AAA	100	50	0.9912

Table 6.2: The hyper-parameters describing the architecture of the RF classifiers that produce the highest F_1 scores, when using all six pressure and flow-rate measurements.

broken down into three components—training, validation, and test. Under this three component framework all ML classifiers are trained using the training set. The accuracy of ML classifiers, however, is assessed using the validation set when tuning hyper-parameters and the test set when evaluating the final tuned model. This allows for unbiased assessment of the final accuracy. It is chosen to only break down the data set into two components due to the lack of repeated hyper-parameter tuning. Using a single test set to assess both optimum hyper-parameters and final tuned accuracy results in classifiers being bias to the test set when using all six measurements, and thus an inflation in the accuracy. It is *a posteriori* found in Section 6.3, however, that this inflation is minor. There is no blatant inflation in the accuracy of classifiers using six input measurements, relative to those using five measurements (which are expected to exhibit a deflation in accuracy due to the lack of configuration specific hyper-parameter optimisation). This suggests that the results achieved when using six input measurement are a reasonable representation of the likely unbiased accuracies, despite the fact that the final tuned accuracy is assessed on the same test set as that used to determine optimum hyper-parameters.

Gradient Boosting

Similar to RF architecture, the GB architecture is optimised for the problem of this study by varying the number of trees within the ensemble and the maximum depth of each tree. Other hyper-parameters which may be varied, however are not considered here, are the minimum number of data points allowed in a leaf node, the maximum number of different features considered for splitting each node, and the impact of each tree on the final outcome (*i.e.* the learning rate). A grid search is

Disease	Trees	Depth	F_1
CAS	100	6	0.9343
SAS	100	7	0.8574
PAD	100	10	0.9187
AAA	80	7	0.9970

Table 6.3: The hyper-parameters describing the architecture of the GB classifiers that produce the highest F_1 scores, when using all six pressure and flow-rate measurements.

carried out to find the combination producing the highest F_1 score when using all the six input measurements. It is common for GB classifiers to use weaker, shallower decision trees (relative to RF classifiers) to deliberately create high bias and low variance [82]. The possible depth of each tree is, therefore, discretised between 2 and 20 at intervals of 1. As a high number of trees is not required to compensate for over fitting, contrary to the RF method, the possible number of trees within the ensemble is discretised between 10 and 100 at intervals of 10. The optimal hyper-parameters for each of the four forms of disease are shown in Table 6.3, and contour plots shown in Appendix J.

Multi-layer perceptron

As is common with deep learning methods, relative to ML methods, there is significantly more hyper-parameter optimisation that can be performed for MLP classifiers than Gradient Boosting or Random Forest. Examples of hyper-parameters that significantly affect the performance of an MLP classifier include—however are not limited to—batch-size, learning rate, activation functions, drop-out, and individual units per hidden layers. With consideration for the exploratory stance of this study, it is not possible to optimise all of these hyper-parameters. It is chosen to use a logistic activation function for all hidden layers. The number of neurons within each hidden layer, and the number of hidden layers is optimised to create the optimal architecture for the classification problem of this study. For simplicity, it is assumed that all the hidden layers contain an identical number of neurons. It is likely that this simplistic hyper-parameter optimisation will limit the accuracy of classification achieved by MLP classifiers.

Disease	Neurons	Depth	F_1
CAS	60	4	0.7785
SAS	190	2	0.6040
PAD	120	2	0.6681
AAA	30	2	0.9785

Table 6.4: The hyper-parameters describing the architecture of the MLP classifiers that produce the highest F_1 scores, when using all six pressure and flow-rate measurements.

Similar to RF and GB, the hyper-parameters that produce the highest F_1 score are found through a grid search. The number of neurons within each layer is discretised between 10 and 200 at intervals of 10, and the number of hidden layers is discretised between 1 and 6 at intervals of 1. The optimal hyper-parameters found for each of the four forms of disease are shown in Table 6.4, and contour plots in Appendix J. Table 6.4 shows that relative to RF and GB, there is less consistency in the maximum F_1 scores achieved by MLP— it classifies AAA and CAS to high levels of accuracies, but performs relatively poorly for SAS and PAD.

6.3 Results and discussion

6.3.1 Input measurement grid search

There are 63 possible combinations of input measurements that can be provided to an ML classifier from the six bilateral pressure and flow-rate measurements (see Section 6.2.1). A combination search is performed for each of the four forms of disease. For every combination of input measurements all the six ML classification methods are trained, and then subsequently tested to quantify their performance. The average F_1 score, sensitivity, and specificity for each case across five folds are recorded. Combinations of interest are then further analysed.

The full tables of results achieved for CAS, SAS, PAD, and AAA classification are shown in Appendices K, L, M, and N respectively. The F_1 score achieved by each ML method and combination of input measurements are visually shown for CAS, SAS, PAD, and AAA classification in Figures 6.2, 6.3, 6.4, and 6.5 respectively.

They show that for all forms of arterial disease, NB and LR classifiers consistently produce low accuracy. It has previously been shown in the PoC (Chapter 3) that the partition between the pressure and flow-rate profiles taken from healthy and stenosed patients is likely to be non-linear. The fact that LR consistently produces low accuracy results supports this finding, as LR is the only linear classification method used. The finding that NB classifiers also produce low accuracy classification is also consistent with the results of the PoC, which found that the NB method is poorly suited to the problem of distinguishing between haemodynamic profiles. On the contrary, across all the four forms of disease, the tree based methods (RF and GB) consistently produce high accuracy results. This finding is in contradiction to the finding in the PoC, and is likely due to the inadequate architecture optimisation or because of the unsuitability of RF on the smaller network used in the PoC. The fact that both RF and GB classifiers are producing high accuracy classification in this study suggests that not only are tree based methods well suited to distinguishing between haemodynamic profiles, but also emphasises the importance of adequate architecture optimisation.

There is less consistency in the results achieved by SVM and MLP classifiers when detecting different forms of disease. SVM classifiers produce accuracies comparable with RF and GB classifiers in the case of AAA detection, however low accuracy results for the three other forms of disease. MLP classifiers produce accuracies comparable with RF and GB classifiers in the case of CAS and AAA detection, however relatively low accuracy results for SAS and PAD classification. Overall, it is found that tree-based methods of RF and GB perform best, with GB performance slightly superior to that of RF.

Measurement combinations

To investigate the importance of both the number of input measurements provided to the ML algorithms and the specific combination of measurements, the average F_1 scores achieved by all classifiers when providing one, two, three, four, five, or six input measurements are found. In each case, the specific combinations that achieve the maximum and minimum F_1 scores are also recorded. These results for different forms of disease are presented next.

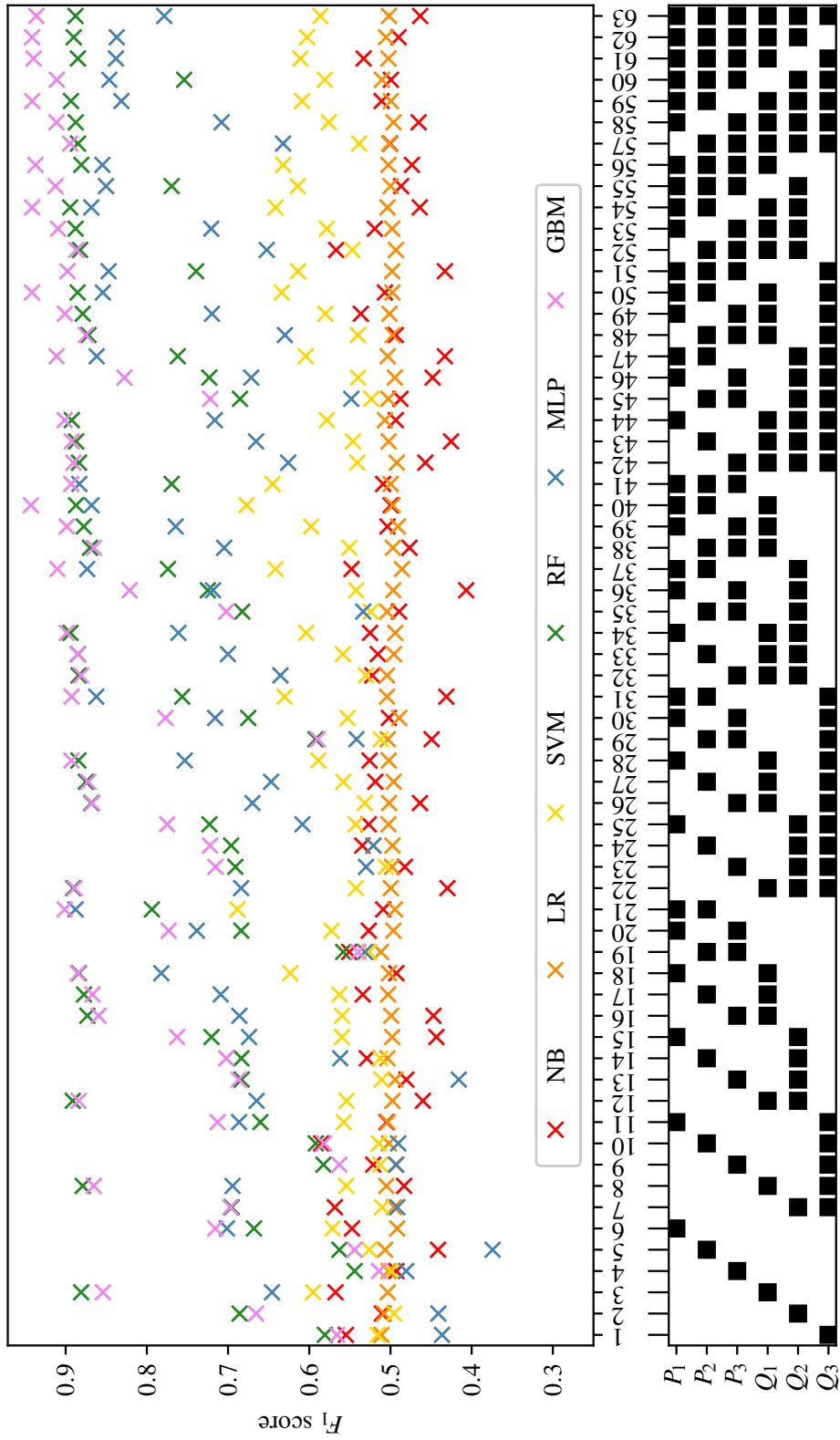


Figure 6.2: The F_1 scores achieved for CAS using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.

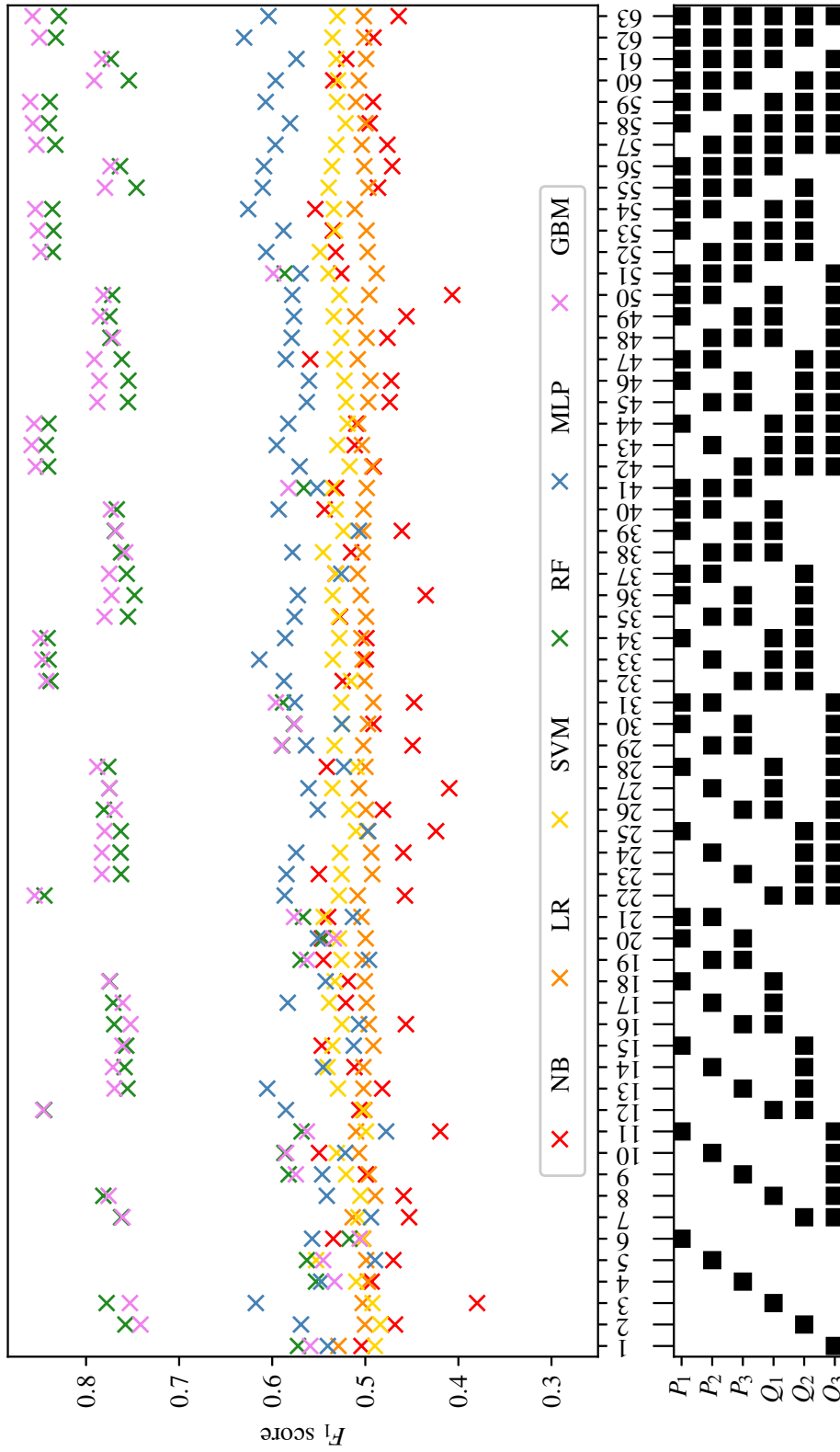


Figure 6.3: The F_1 scores achieved for SAS using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.

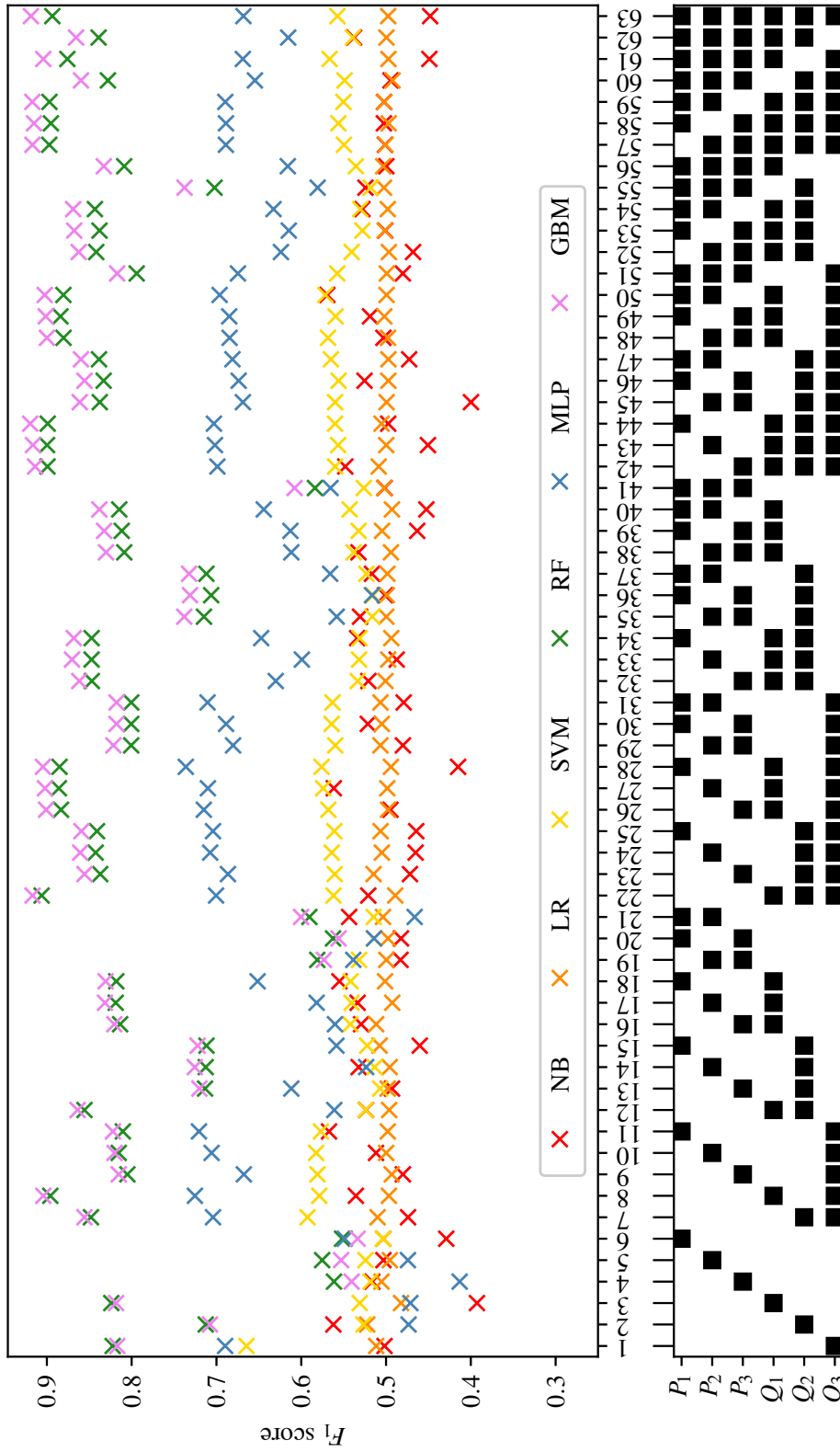


Figure 6.4: The F_1 scores achieved for PAD using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.

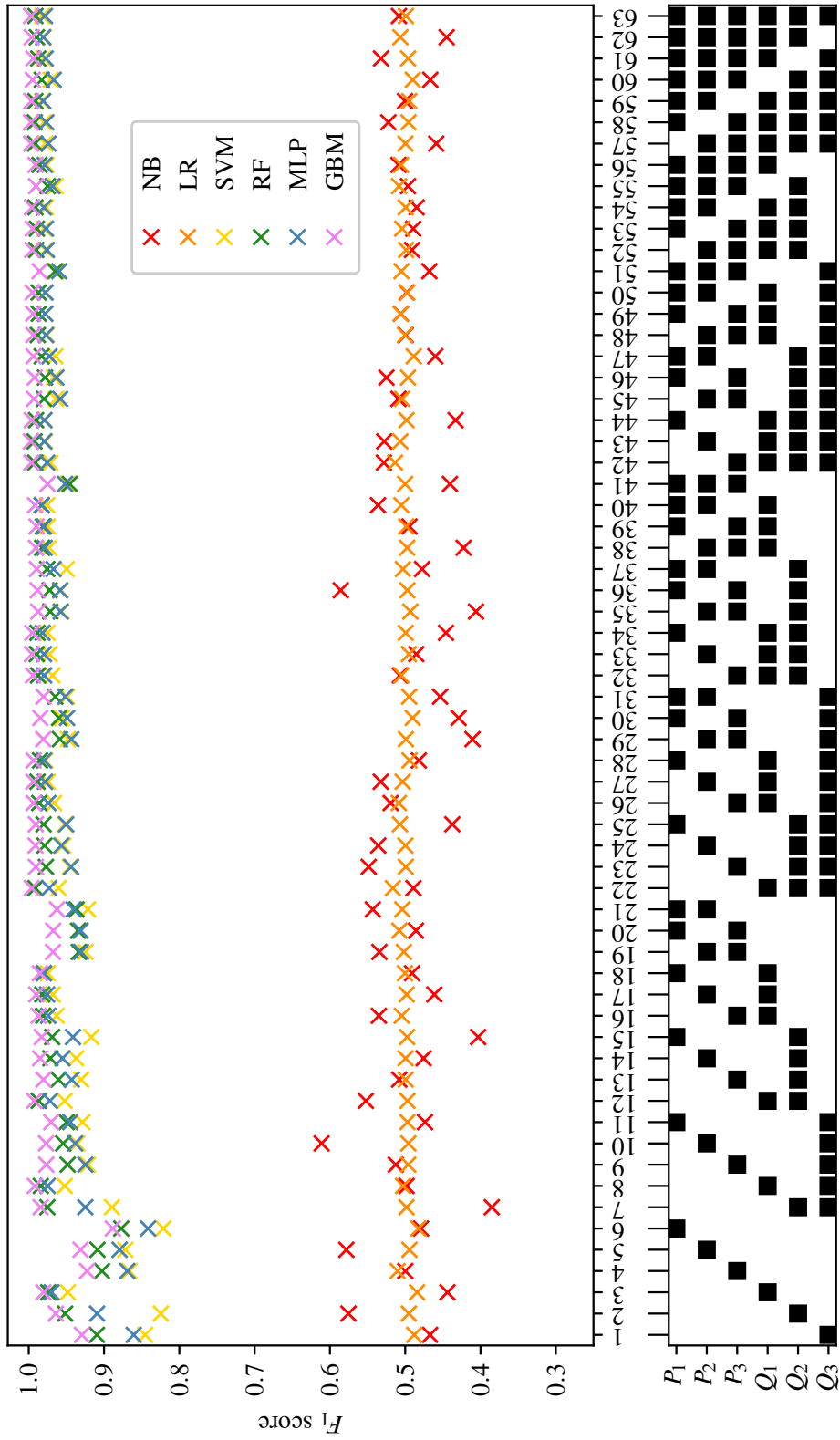


Figure 6.5: The F_1 scores achieved for AAA using each combination of bilateral input measurements are shown. Measurements included within each combination are highlighted with a black square.

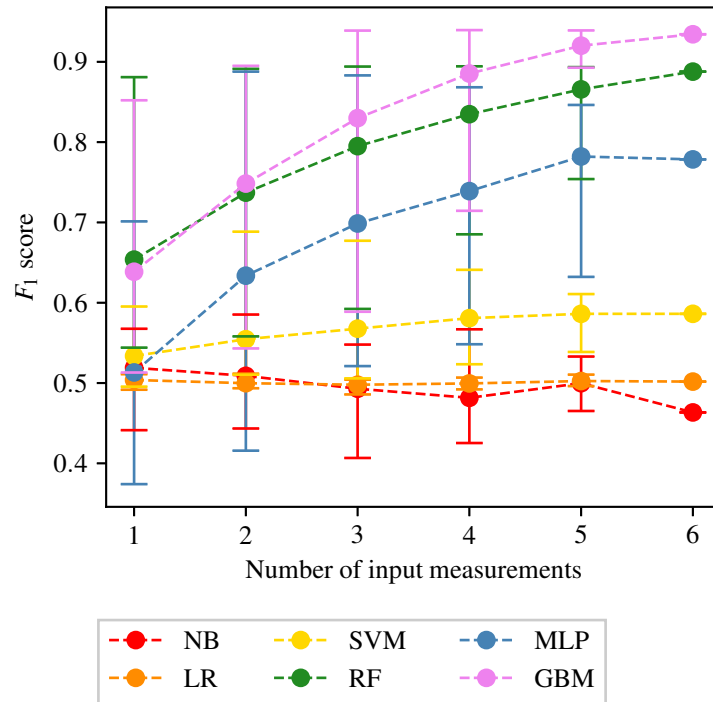


Figure 6.6: The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for carotid artery stenosis classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.

CAS classification: The average, maximum and minimum F_1 score achieved when providing different number of input measurements for CAS classification are shown in Figure 6.6. It shows that NB, and LR classifiers consistently produce an F_1 score of approximately 0.5, which is comparable to naive classification, *i.e.* randomly assigning the health of VPs with an equal probability to each outcome. SVM performs slightly better with F_1 scores averaging 0.5 – 0.6. The other three classification methods (RF, MLP, and GB) perform significantly better with F_1 scores generally averaging between 0.7 and 0.95 and showing a clear increase in the average F_1 score as the number of input measurements increases. While the average and minimum F_1 score achieved by RF and GB classifiers continuously increases, the maximum F_1 score achieved can be seen to quickly reach a plateau (at one input measurement for RF, and three input measurements for GB). For a fixed number of measurements, the wide range of F_1 scores in Figure 6.6 across

all classifiers suggests that specific combinations of measurements may be more important than others for optimal classification. To explore this further, the combinations of input measurements that produce the highest F_1 scores and the corresponding accuracies when employing the RF and GB methods are shown in Table 6.5. Two observations are made from this table. First that for a fixed number of measurements, the best combinations are not identical for the two methods. For example, when two measurements are used the best combination for RF is (Q_2, Q_1) while the best combination for GB is (P_2, P_1) . This suggests that the best combination of measurements is likely dependent on the particular ML method chosen. Second, some patterns stand out with respect to which measurements may be more informative than others. For example, across the Table 6.5 Q_1 appears in 11 out of 12 combinations and P_1 appears in 8 out of 12 combinations. This suggests that Q_1 is most informative about identifying the presence of CAS followed by P_1 . Physiologically, this is not surprising as Q_1 and P_1 are flow-rates and pressures in the carotid arteries and the disease under consideration is carotid artery stenosis. It is encouraging that the ML methods are indeed placing more importance to the relevant physiological measurements. In fact, it is remarkable that RF and GB both achieve F_1 scores above 0.85 and sensitivities and specificities larger than 85% with only Q_1 . Also notable is that these accuracies can be taken to beyond 93% (see GB row for 3 measurements in Table 6.5) when adding two more measurements as long as the additional two measurements are carefully chosen.

An interesting pattern to note is that while the average and minimum F_1 score achieved by MLP classifiers continuously increases in Figure 6.6, the maximum F_1 score decreases beyond three input measurements. The maximum F_1 scores achieved by MLP classifiers, and the corresponding sensitivities and specificities, when using three to six input measurements are shown in Table 6.6. It shows that the decrease in F_1 scores is also accompanied by an associated decrease in both the sensitivities and specificities, as opposed to the balance between them (increase in sensitivity and decrease in specificity and vice versa). This behaviour is unusual as intuitively more input measurements should generally provide more information. This finding may suggest that MLP classifiers are able to extract maximum information from the haemodynamic profiles when using as little as three input measurements, and are susceptible to over fitting when using more than three measurements (hence the decrease in accuracy).

6.3. RESULTS AND DISCUSSION

No. of input measurements	Method	Combination	F_1 score	Sens.	Spec.
1	RF	(Q_1)	0.8809	0.8704	0.8893
	GB	(Q_1)	0.8521	0.8547	0.8502
2	RF	(Q_2, Q_1)	0.8913	0.8765	0.9032
	GB	(P_2, P_1)	0.8950	0.9026	0.8889
3	RF	(Q_2, Q_1, P_1)	0.8941	0.8825	0.9035
	GB	(Q_1, P_2, P_1)	0.9389	0.9433	0.9351
4	RF	(Q_2, Q_1, P_2, P_1)	0.8944	0.8858	0.9015
	GB	(Q_3, Q_1, P_2, P_1)	0.9395	0.9417	0.9376
5	RF	$(Q_3, Q_2, Q_1, P_2, P_1)$	0.8934	0.8858	0.8996
	GB	$(Q_2, Q_1, P_3, P_2, P_1)$	0.9391	0.9416	0.9370
6	RF	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8878	0.8747	0.8984
	GB		0.9343	0.9364	0.9325

Table 6.5: The combinations of input measurements that produce the maximum F_1 scores when providing one to six input measurements and employing the RF and GB methods to detect CAS. The corresponding sensitivities and specificities are also included.

To investigate this hypothesis, the log loss cost across the training and test sets are recorded at each sequential iteration of the training process (up to the 200th iteration) when using each input measurement combination presented in Table 6.6. This analysis is shown in Figure 6.7. MLP classifiers are trained using the principle of gradient descent, *i.e.* each sequential iteration takes a ‘step’ in the opposite direction to the gradient of the cost of predictions. Thus, at a low number of training iterations both the training and test costs are expected to be high (the classifier can neither fit the training data nor generalise to the test data). As the training process progresses, the training and test costs are expected to exponentially decay (maintaining a relatively constant difference between training and test costs), before asymptoting when a global minima is reached. Figure 6.7 shows that when using three input measurements the training cost continuously decreases, while the test cost quickly reaches a minimum which is then approximately maintained (with some minor gradual increase in test cost). This is a close to desired profile. It is seen that as the number of input measurements increases i) the training cost still continuously decreases, ii) the minimum test cost achieved increases, and iii) the rate of increase in test cost after the minimum has been achieved significantly

No. of input measurements	Combination	F_1 score	Sensitivity	Specificity
3	(P_3, P_2, P_1)	0.8831	0.8731	0.8911
4	(Q_3, Q_1, P_2, P_1)	0.8683	0.8538	0.8545
5	$(Q_3, Q_2, P_3, P_2, P_1)$	0.8463	0.8308	0.8577
6	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.7785	0.7916	0.7703

Table 6.6: The combinations of input measurements that produce the maximum F_1 scores when providing three to six input measurements and employing the MLP method to detect CAS. The corresponding sensitivities and specificities are also included.

increases. These behaviours are all signs of overfitting and suggest that, as is *a priori* expected, the more intensive nature of MLP classifiers make them less suited to this exploratory study.

For comparison the training and test cost profiles of the GB and RF classifiers, when using the combination of three to six input measurements that are found to produce the highest F_1 scores, are also included in Figure 6.7. GB classifiers are also trained using the principle of gradient descent, however with each new tree in the ensemble taking a step down the gradient (rather than training iteration, like in the case of MLP). Thus, the desired training and test cost profiles of GB classifiers are identical to that of MLP, with the exception of number of training iterations being replaced with number of trees. RF classifiers are not based on the principle of gradient descent. Each decision tree within an RF ensemble should be a self contained, albeit high variance, classifier. Thus, with a low number of trees within the ensemble the RF classifier is expected to produce relatively low training accuracies, however high test accuracies (*i.e.* the classifier begins in a state of over-fitting). As more trees are included within the ensemble, the training costs are expected to quickly reach there asymptotic minimum. The test costs, however, will experience a ‘lag’, before enough trees are eventually included to allow the aggregation of the predictions of different trees to counteract the high variance of each individual tree. Over-fitting is shown in the training and test cost profiles of RF classifiers by a significant, maintained difference between the training and test costs. Figure 6.7 shows that both GB and RF classifiers are producing desired profiles when using all numbers of input measurements, with no signs of over-fitting

present.

Although the primary purpose of this study is not to perform classifier specific optimisation, to achieve increasingly higher accuracies, an investigation into the achievable MLP performance increase, without the need for more complex hyper-parameter optimisation, is presented in Section 6.3.5.

SAS classification: The results of the analysis for SAS classification are shown in Figure 6.8. As is seen in the case of CAS classification, Figure 6.8 shows that NB, LR, and SVM classifiers consistently produce accuracies comparable to naive classification, irrespective of the number of input measurements used. A clear difference between Figures 6.6 and 6.8 is the accuracy achieved by MLP classifiers. Compared to the CAS case, the MLP performance is further degraded for SAS, while still being better than NB, LR, and SVM, although only marginally.

A high degree of similarity can be seen between the behaviours of RF and GB classifiers for CAS and SAS. Figure 6.8 shows that the average and minimum F_1 score achieved by RF and GB classifiers continuously increases as the number of input measurements used increases. The maximum F_1 score achieved is seen to quickly reach an asymptotic limit (at three input measurements for both RF and GB classifiers). Peak F_1 score of approximately 0.85 is achieved by GB, along with sensitivities and specificities higher than 85%.

The combination of input measurements that produce the highest F_1 scores and the corresponding accuracies are shown in Table 6.7. It shows a higher degree of consistency between the best combinations for the two methods relative to the case for CAS, *i.e.* the best combinations are generally identical (or with minimal differences) between RF and GB. It also shows that Q_1 is particularly informative, with this measurement appearing in all of the best combinations. Physiologically this may be due to its proximity to the disease location.

PAD classification: The results for PAD classification are shown in Figure 6.9. Comparing Figures 6.8 and 6.9, a high degree of similarity is seen between the behaviours of SAS and PAD classification. As is previously seen for SAS classification, Figure 6.9 shows that the NB, LR, and SVM methods are all consistently producing accuracies comparable to naive classification. While the MLP method performs slightly better than naive classification, the accuracy still

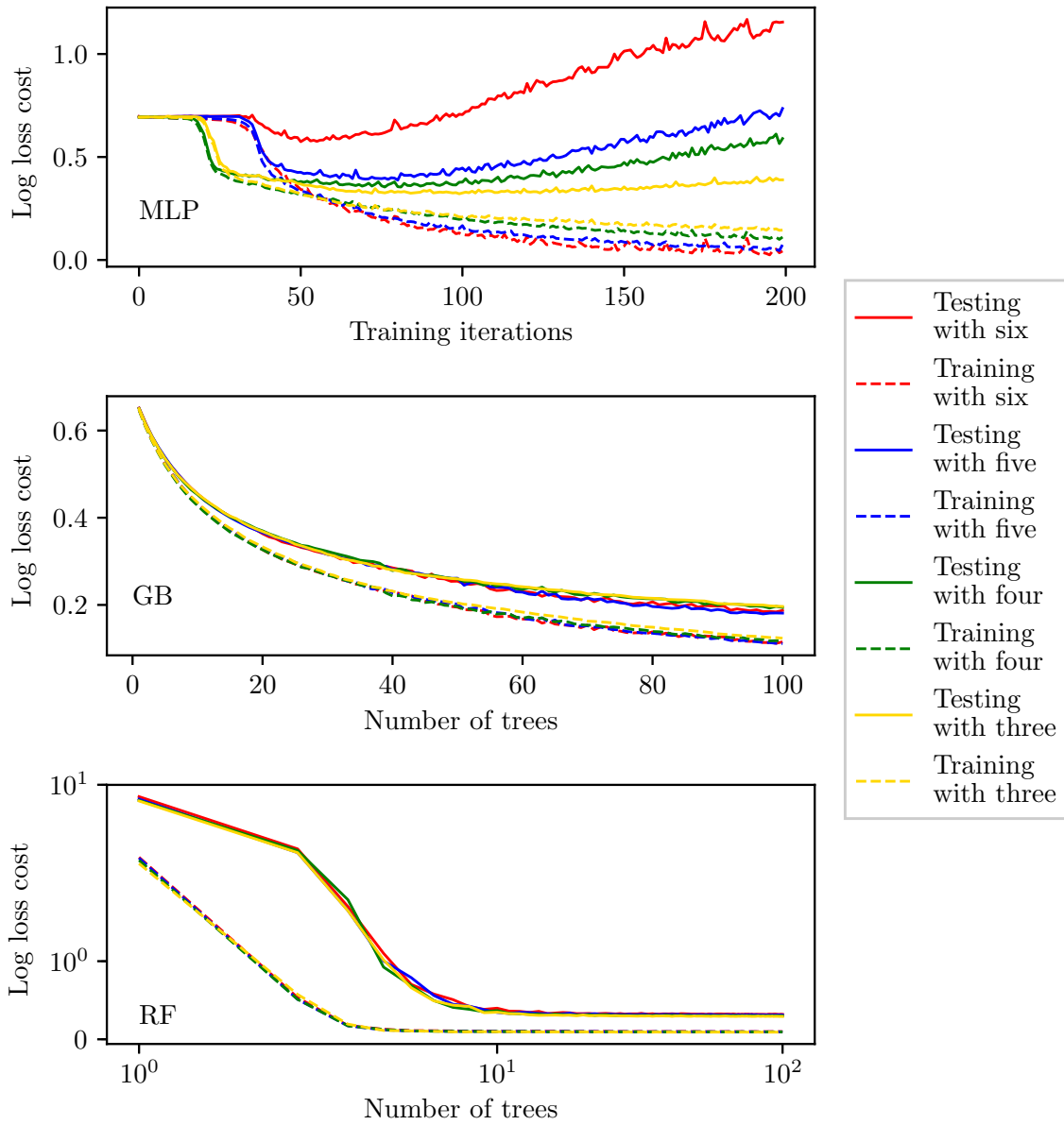


Figure 6.7: The average log loss cost across the training and test sets when using the combination of six to three input measurements that achieve the highest accuracies (as detailed in Tables 6.5 and 6.6).

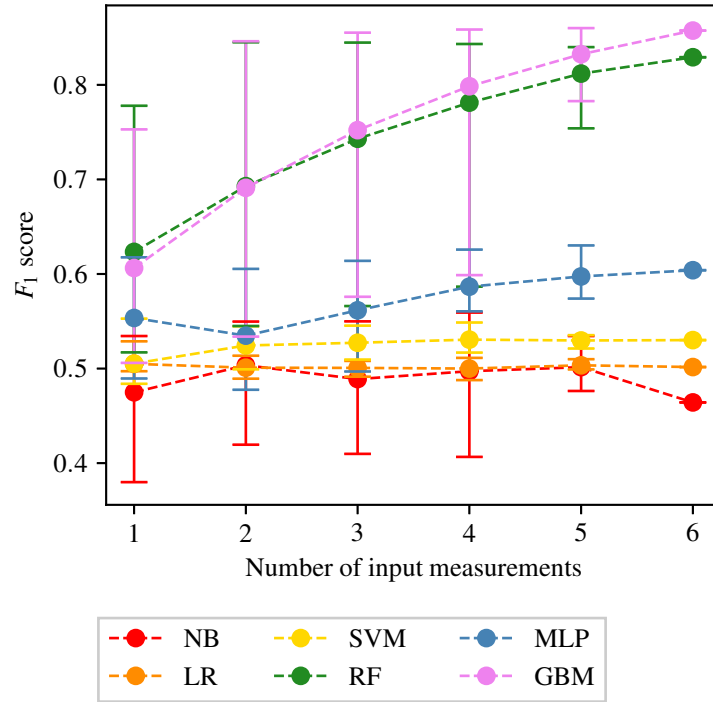


Figure 6.8: The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for SAS classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.

remains relatively low. High accuracy can be seen in Figure 6.9 for the two tree based methods of RF and GB. As has been previously seen for CAS and SAS, while the average and minimum F_1 score achieved by the RF and GB methods increases as the number of input measurements increases, the maximum F_1 score achieved quickly reaches an asymptotic limit (at 3 input measurements for both the RF and GB methods).

The combination of input measurements that produce the highest F_1 scores for PAD classification when employing the RF and GB methods are shown in Table 6.8. Table 6.8 not only shows good consistency between the combinations of input measurements that produce the highest F_1 scores when employing each of the two ML methods, but also good agreement with the combinations presented in Table 6.7. Very similar combinations of input measurements (with some minor differences) can be seen to produce the highest F_1 score when providing all numbers of input

No. of input measurements	Method	Combination	F_1 score	Sens.	Spec.
1	RF	(Q_1)	0.7779	0.7582	0.7905
	GB	(Q_1)	0.7529	0.7224	0.7714
2	RF	(Q_2, Q_1)	0.8450	0.8374	0.8507
	GB	(Q_2, Q_1)	0.8461	0.8293	0.8585
3	RF	(Q_3, Q_2, Q_1)	0.8447	0.8271	0.8576
	GB	(Q_3, Q_2, Q_1)	0.8552	0.8453	0.8626
4	RF	(Q_3, Q_2, Q_1, P_2)	0.8432	0.8303	0.8527
	GB	(Q_3, Q_2, Q_1, P_2)	0.8585	0.8487	0.8660
5	RF	$(Q_3, Q_2, Q_1, P_3, P_1)$	0.8399	0.8256	0.8504
	GB	$(Q_3, Q_2, Q_1, P_2, P_1)$	0.8600	0.8525	0.8657
6	RF	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8292	0.8102	0.8427
	GB	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8574	0.8504	0.8627

Table 6.7: The combinations of input measurements that produce the maximum F_1 scores when providing one to six input measurements and employing the RF and GB methods to detect SAS. The corresponding sensitivities and specificities are also included.

measurements. As has previously been observed in Tables 6.5 and 6.7, the input measurement Q_1 appears to be most informative, appearing in all the best scoring classifiers. Since the location of Q_1 is far from the location of disease, it is not obvious why this measurement is particularly informative of PAD.

AAA classification: The results for AAA classification are shown in Figure 6.10. As has been previously seen for all of the three other forms of disease the NB, and LR classifiers consistently produce accuracies comparable to naive classification, irrespective of the number of input measurements used. The consistency of this finding (as seen in Figures 6.6, 6.8, and 6.9), irrespective of the form of disease being classified, highlights both the importance of non-linear partitions between healthy and unhealthy VPs and the unsuitability of the NB method for distinction between haemodynamic profiles.

In the case of AAA classification the SVM, RF, MLP, and GB methods consistently produce good accuracies. Figure 6.10 shows that these methods produce high accuracies even with a single input measurement. While there is some increase in the average F_1 score as the number of input measurements increases, due to the

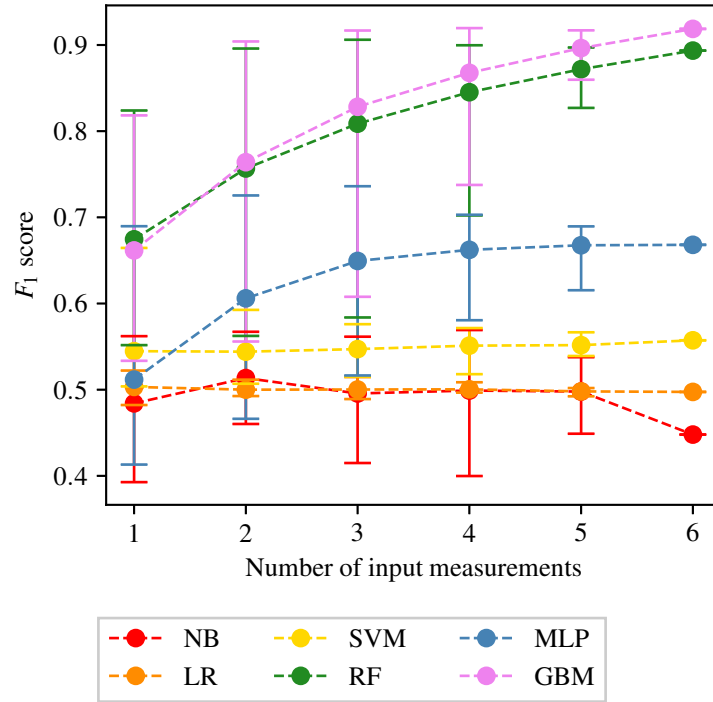


Figure 6.9: The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for PAD classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.

very high initial average F_1 score achieved (when using a single input measurement) this increase is limited (as the F_1 score can not exceed 1). Two possible reasons of the higher accuracies in aneurysm classification relative to stenosis classification are:

- Aneurysms, owing to an increase in area as opposed to decrease in the area for stenoses, may actually produce more significant or consistent biomarkers in the pressure and flow-rate profiles. This hypothesis is supported by [127], which found that even low severity AAAs have a global impact on the pressure and flow-rate profiles.
- While the severities of aneurysms cannot be directly compared to severities of stenosis, it may be that the severity of aneurysms in VPD_{AAA} are disproportionately large relative to the severities of stenoses. The significance

No. of input measurements	Method	Combination	F_1 score	Sens.	Spec.
1	RF	(Q_1)	0.8240	0.8959	0.8320
	GB	(Q_1)	0.8183	0.8126	0.8214
2	RF	(Q_3, Q_1)	0.8140	0.8825	0.9068
	GB	(Q_3, Q_1)	0.9041	0.8950	0.9117
3	RF	(Q_3, Q_2, Q_1)	0.9061	0.8885	0.9208
	GB	(Q_3, Q_2, Q_1)	0.9168	0.9055	0.9265
4	RF	(Q_3, Q_2, Q_1, P_2)	0.8997	0.8868	0.9104
	GB	(Q_3, Q_2, Q_1, P_1)	0.9196	0.9068	0.9306
5	RF	$(Q_3, Q_2, Q_1, P_3, P_2)$	0.8971	0.8802	0.9110
	GB	$(Q_3, Q_2, Q_1, P_2, P_1)$	0.9170	0.9041	0.9281
6	RF	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.8935	0.8813	0.9035
	GB	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.9187	0.9102	0.9261

Table 6.8: The combinations of input measurements that produce the maximum F_1 scores when providing one to six input measurements and employing the RF and GB methods to detect PAD. The corresponding sensitivities and specificities are also included.

of any indicative biomarkers introduced into pressure and flow-rate profiles is likely to be proportional to the severity of the change in area. This implies that the increase in vessel area of 712%–2,593% in VPD_{AAA} is perhaps on the extreme end of aneurysm severity, thereby making the classifications relatively easier. This is further explored in section 6.3.3.

The combination of input measurements that produce the highest F_1 scores when providing one to six input measurements and employing the RF and GB methods are shown for AAA classification in Table 6.9. It shows that F_1 scores range from 0.97–0.997 and sensitivities and specificities range from 99% to 99.8%. Due to the high accuracies across all the number of measurements, the analysis of specific combinations is not very meaningful. However, the measurement Q_1 again appears in all the best combinations. It should also be noted that the high accuracies for AAA classification are also consistent with those reported in [33]— where deep-learning methods are applied on a VPD created by varying seven network parameters, and classification accuracies of $\approx 99.9\%$ are reported—and [214]—where machine learning methods are applied on a VPD, and sensitivities and specificities of $\approx 86\%$

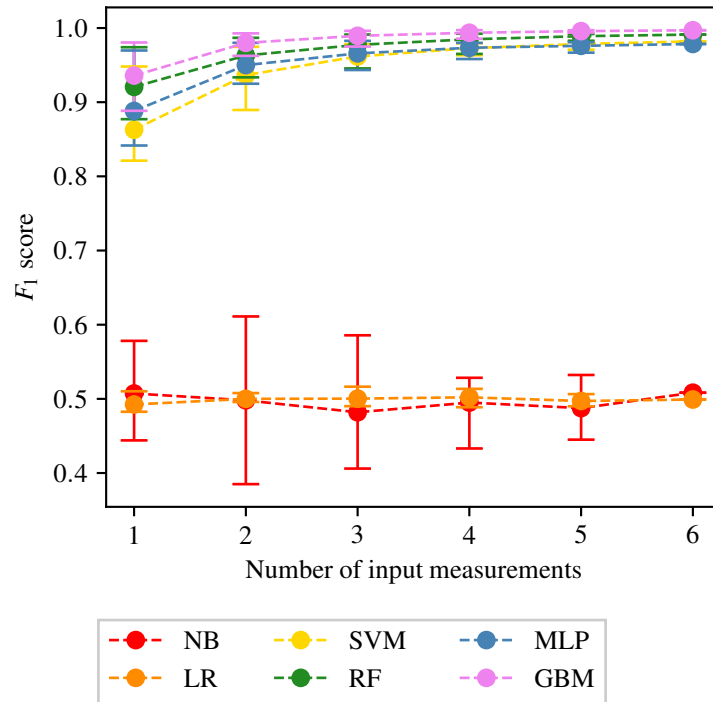


Figure 6.10: The average, maximum, and minimum F_1 score achieved by all classifiers trained using different numbers of input measurements are shown for AAA classification. The central markers represent the average score achieved, while the error bars indicate the upper and lower limits.

are reported.

Importance of carotid artery flow-rate

Appendices K–N, along with the above analysis show that classifiers trained using flow-rates in the common carotid arteries (Q_1) consistently produce the highest accuracy. To analyse this further, the F_1 scores of classifiers with combinations that do and do not contain Q_1 are separated and compared for CAS, SAS, PAD, and AAA in Figures 6.11, 6.12, 6.13, and 6.14 respectively. For each disease form, results are only shown for the classification methods consistently producing reasonable accuracy results. Generally the aforementioned figures shown a clear positive shift in the histograms when Q_1 is included, pointing to the particularly informative nature of Q_1 . Important behaviours that are observed include:

No. of input measurements	Method	Combination	F_1 score	Sens.	Spec.
1	RF	(Q_1)	0.9741	0.9654	0.9825
	GB	(Q_1)	0.9805	0.9799	0.9811
2	RF	(Q_2, Q_1)	0.9868	0.9810	0.9926
	GB	(Q_2, Q_1)	0.9928	0.9919	0.9938
3	RF	(Q_3, Q_2, Q_1)	0.9912	0.9864	0.9961
	GB	(Q_3, Q_2, Q_1)	0.9962	0.9954	0.9970
4	RF	(Q_3, Q_2, Q_1, P_2)	0.9923	0.9879	0.9967
	GB	(Q_3, Q_2, Q_1, P_2)	0.9972	0.9959	0.9986
5	RF	$(Q_3, Q_2, Q_1, P_3, P_1)$	0.9920	0.9873	0.9967
	GB	$(Q_3, Q_2, Q_1, P_3, P_2)$	0.9970	0.9959	0.9981
		$(Q_3, Q_2, Q_1, P_3, P_1)$	0.9970	0.9963	0.9978
6	RF	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.9912	0.9861	0.9964
	GB	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.9970	0.9959	0.9981

Table 6.9: The combinations of input measurements that produce the maximum F_1 scores when providing one to six input measurements and employing the RF and GB methods to detect AAA. The corresponding sensitivities and specificities are also included.

- While there is generally an increase in F_1 score achieved when including Q_1 , it is also simultaneously observed that the maximum accuracy achieved are relatively less sensitive to the inclusion of Q_1 .
- The greatest distinction between F_1 scores when including or excluding Q_1 is seen for CAS classification when using the RF method. There is no overlap between the two RF histograms in Figure 6.11.
- Observing the lower plots in Figures 6.12 and 6.13, a clear subgroup of low-accuracy classifiers can be seen when excluding Q_1 for SAS and PAD, which does not exist when including Q_1 .

6.3.2 Feature importance

An important aspect of the GB method is that the measurement importance, which determines the influence that individual measurements have towards classification, can be computed. This split-improvement feature importance [224] of a feature

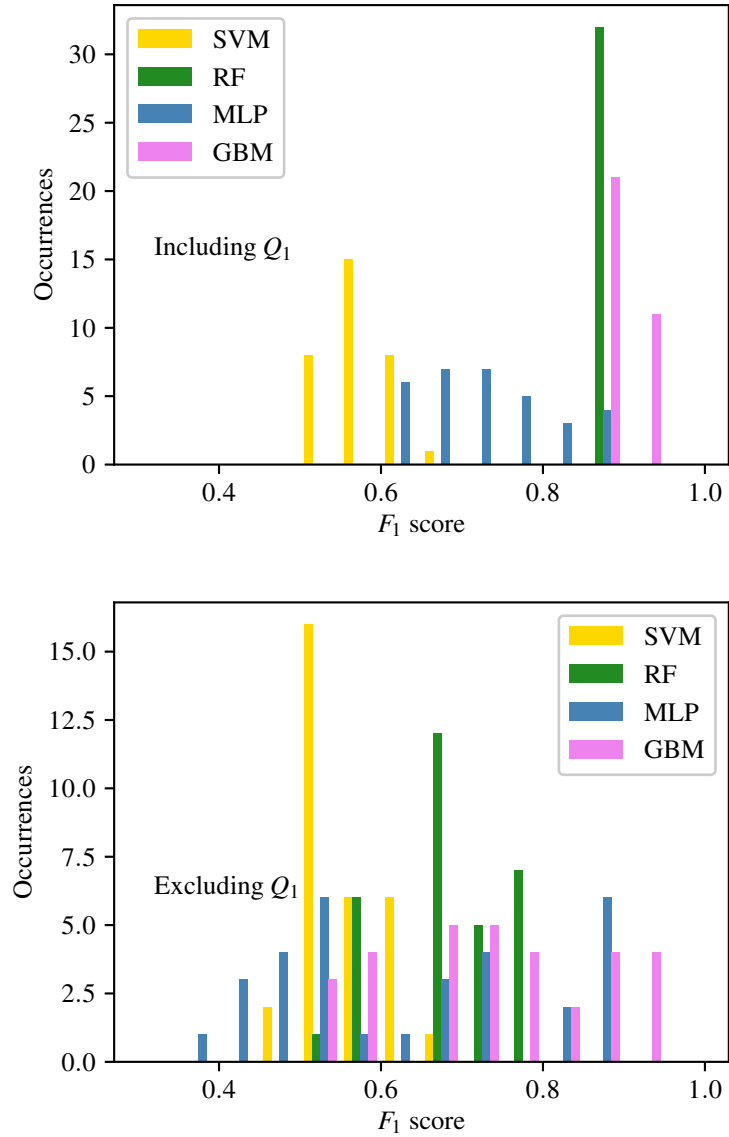


Figure 6.11: The histograms of the F_1 scores achieved for CAS classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.

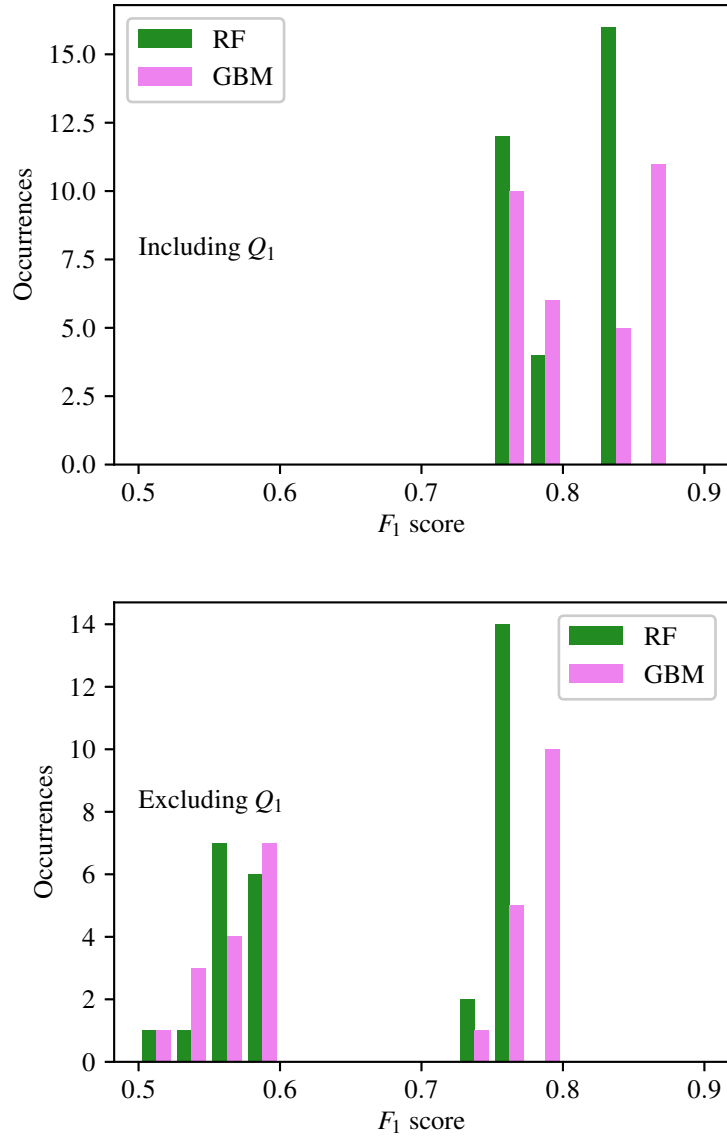


Figure 6.12: The histograms of the F_1 scores achieved for SAS classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.

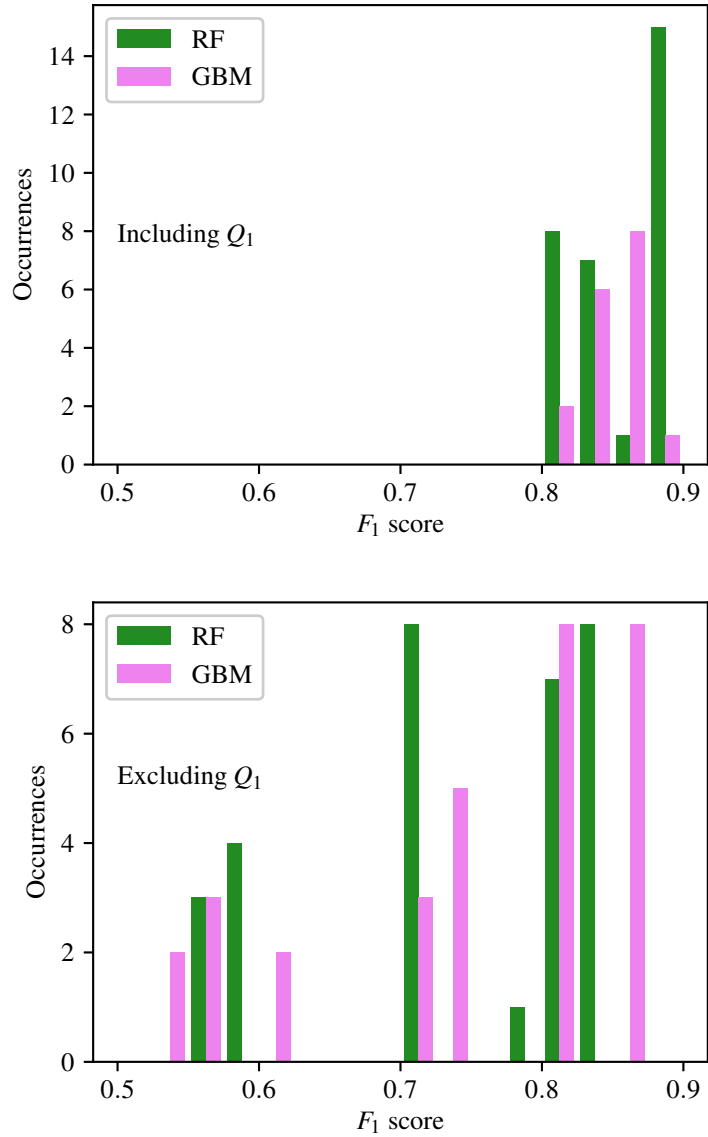


Figure 6.13: The histograms of the F_1 scores achieved for PAD classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.

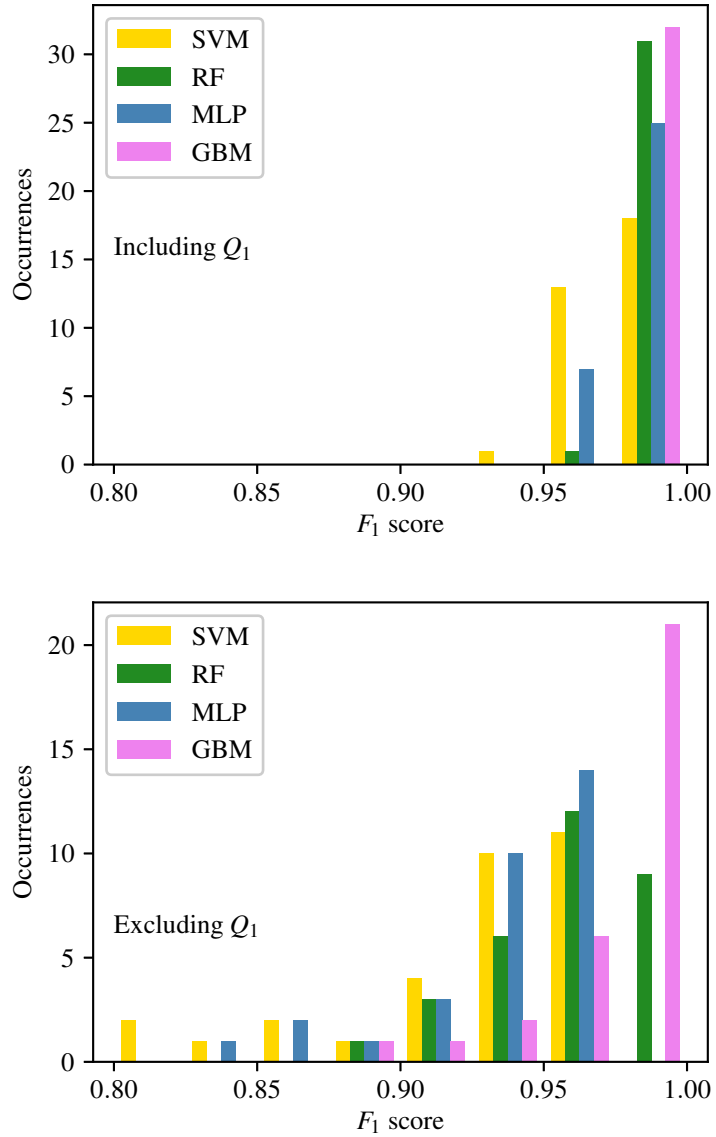


Figure 6.14: The histograms of the F_1 scores achieved for AAA classification are shown for all input measurement combinations that include Q_1 in the upper plot, and exclude Q_1 in the lower plot.

	Q_1 (%)	Q_2 (%)	Q_3 (%)	P_1 (%)	P_2 (%)	P_3 (%)
CAS	67.38	8.02	3.89	11.07	7.692	1.93
SAS	41.90	29.98	8.40	6.80	5.97	6.921
PAD	38.01	15.98	31.11	4.62	4.63	5.62
AAA	69.34	19.10	4.95	2.41	2.61	1.55

Table 6.10: The total importance of each input measurement, based on the GB classifiers provided with all six measurements.

can be thought of as the contribution of that feature to the total information gain achieved in a decision tree, averaged across all the trees in the ensemble. A high feature importance suggests that the given feature is contributing heavily to the classification accuracies achieved. As the features provided to the GB classifiers are the FS coefficients describing the haemodynamic profiles, the total importance of each bilateral pressure or flow-rate measurement is found by summing the feature importance of the associated 22 FS coefficients. The total importance of each input measurement for each disease form is shown in Table 6.10. Three important observations from this table are:

- The input measurement Q_1 consistently produces the highest importance for all forms of disease. This finding supports the findings of Section 6.3.1.
- The importance of each input measurement changes between disease forms based on the spatial proximity to the disease location. Generally, the measurements in close proximity to the disease location have higher importance. For example the importance of Q_3 (flow-rate in the femoral arteries) is highest for PAD classification (see Figure 6.1 for locations of disease and measurements). Similarly, P_1 (pressure in carotid arteries) has highest importance for CAS and SAS.
- The feature importance, when viewed in collection, also shed some light on why Q_1 is important for PAD even though the measurement location is far from the disease location. For PAD the two most informative measurements are Q_1 and Q_3 . From Figure 6.1, it is clear that this combination forms a pair of flow-rate before and at the disease location. Thus, the measurement locations bound the disease location to provide more information on the presence of disease.

6.3.3 Lower severity aneurysms

In Section 6.3.1 it is found that AAAs can be classified to a very high levels of accuracy with only one input measurement. Whether these accuracies are affected when lower severity aneurysms are considered is assessed here. For this assessment, a new lower severity AAA VPD, referred to as VPD_{AAA-L} , is created in an identical manner to the other diseased databases (see Chapter 5), with the following two differences:

- The severity of aneurysms introduced into the virtual subjects is sampled from a uniform distribution bounded as follows: $3.0 \leq \mathcal{S}_{aneurysm} \leq 7.0$.
- To reduce the computational expense associated with the creation of virtual patients, the size of VPD_{AAA-L} is restricted to 5,000 VPs.

A combination search is carried out with only the GB method as it is the best overall method. The F_1 scores, sensitivities, and specificities achieved by all the measurement combinations are presented in Appendix O. For comparison, the GB F_1 scores for all forms of disease (including AAA-L) are shown in Appendix P. The ratios of the GB F_1 scores achieved for AAA-L classification relative to AAA classification are shown in Figure 6.15. The observations from this figure are:

- The F_1 scores for AAA-L classification are consistently lower (ranging from 1% to 10% lower) than that for AAA classification. This finding supports the physiological expectation that the significance of biomarkers in pressure and flow-rate profiles is proportion to the severity.
- The ratios of F_1 scores are lowest for combinations of inputs that predominantly rely on pressure measurements. This suggests that pressure measurements are, in general, less informative about disease severity. This is in support of the, generally, lower feature importance of pressure measurements in Table 6.10.
- The F_1 score ratios are highest for input combinations that include Q_1 . This finding further suggests that Q_1 contains consistent biomarkers.
- The ratios range between 0.9 and 0.99, implying a maximum degradation of only 10% relative to high-severity classification accuracies. Thus, even in

the low-severity aneurysms many combinations of classifiers achieve F_1 scores higher than 0.95 and corresponding sensitivities and specificities larger than 95%.

6.3.4 Unilateral aneurysm measurement tests

Hitherto, all ML classifiers used bilateral measurements, *i.e.* both the right and left instances of each measurement are simultaneously provided. Here, the ability of unilateral measurements, *i.e.* only the right or left instance of a measurement, to detect AAAs is assessed. This analysis is restricted to the GB method as it consistently outperforms other methods. GB classifiers are trained and tested to detect AAAs using four different unilateral measurements:

- **Flow-rate in the right carotid artery**, shown in Figure 6.1 as $Q_1^{(R)}$.
- **Flow-rate in the left carotid artery**, shown in Figure 6.1 as $Q_1^{(L)}$.
- **Pressure in the right radial artery**, shown in Figure 6.1 as $P_3^{(R)}$.
- **Pressure in the left radial artery**, shown in Figure 6.1 as $P_3^{(L)}$.

Carotid artery flow-rate is chosen as it has been shown to be the best measurement for disease classification. Radial artery pressure is chosen due to the location of the radial artery on the human wrist. Recent advancements have resulted in wearable devices capable of measuring continuous radial pressure profiles, such as the TLT Sapphire monitor (Tarilian Laser Technologies, Welwyn Garden City, U.K.) [125], and thus if AAAs can be detected to satisfactory accuracies using these measurements, it may suggest the possibility of future home monitoring of abdominal aortic health through such wearables. The sensitivities and specificities achieved by the four unilateral GB classifiers are shown in Table 6.11. It shows that relative to the bilateral case, while there is a decrease in the classification accuracies, the magnitude of the decrease is less than 10%. This finding suggests that there may be sufficient biomarkers of AAA presence captured within the intra-measurement details of a single pressure or flow-rate profile. The fact that similar accuracies are achieved with either the right or left instances of any measurement is likely due to physiological symmetry. While there are some minor asymmetries between the right

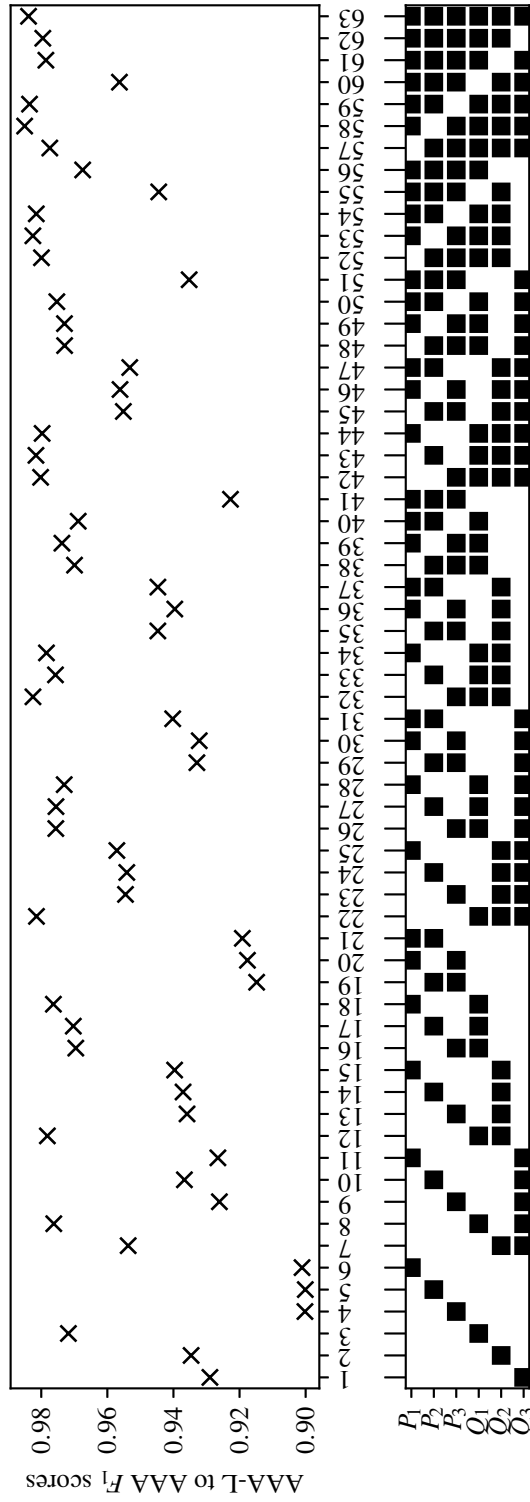


Figure 6.15: The ratios of the F_1 scores for AAA-L classification relative to AAA classification, when providing each combination of input measurements are shown. Measurements included within each combination are highlighted with a black square.

	Side	Sensitivity	Specificity
Carotid flow-rate (Q_1)	Right	0.9369	0.9161
	Left	0.9065	0.9146
	Both	0.9799	0.9811
Radial pressure (P_3)	Right	0.8356	0.8533
	Left	0.8633	0.8605
	Both	0.9202	0.9248

Table 6.11: The sensitivities and specificities achieved when using the measurements of flow-rate in the right, left, and both CAs and pressure in the right, left, and both radial arteries.

and left upper extremities, due to the topology of the arterial network (as shown in Figure 6.1) changes to the cross sectional area of the abdominal aorta are expected to produce relatively consistent changes in both the right and left side of the body.

6.3.5 MLP over-fitting

It is shown in Section 6.3.1 that the accuracy of MLP classifiers is hindered by the presence of significant over-fitting. One method to reduce the presence of over-fitting is to perform more complex hyper-parameter optimisation (such as introducing regularisation). This method is prohibitive, however, due to the increased dimensionality, and so consequently complexity, associated with the tuning of hyper-parameters. Here, the impact on the accuracy of classification by reducing over-fitting without introducing further hyperparameters is investigated by employing an early-stopping criteria to the training process [160, 220]. Essentially, rather than splitting the available data in two (as described in Section 6.2.2), training the MLP classifier until a stopping criteria is met on the training set, and then assessing the accuracy on the test set:

- A third partition to the available data (the validation set) is introduced. The combined healthy and unhealthy data sets in 6.2.2 are split so that the training set contains 50%, the validation set 25%, and the test set 25% of the available data.
- Classifiers are trained on the training set, however with the stopping criteria

being assessed on the validation set. At each sequential iteration in the training process, the average log loss cost is computed across the validation set. If more than 75 iterations have been performed, and the improvement in log loss cost across the validation set between two consecutive iterations is less than 1×10^{-3} , training is stopped.

- The final accuracy is assessed on the test set.

This analysis is performed for CAS and AAA, as the behaviour of SAS and PAD is very similar to that of CAS. MLP classifiers are trained using early-stopping and the combination of input measurements found to produce the highest F_1 score when providing three to six input measurements. As the available data set is now being partitioned into three components the hyper-parameter optimisation performed in Section 6.2.4 is re-completed, however with the three amendments that:

- Early-stopping is employed throughout the grid search.
- Optimum architecture is chosen based on the F_1 score achieved across the validation set (rather than the test set previously used). This results in the test set being completely unseen by the classifiers until the assessment of the final accuracy, removing any bias that maybe present.
- As there are significantly fewer combinations of input measurements being tested (four CAS and four AAA classifiers), hyper-parameter optimisation is performed for each individual classifier.

The hyper-parameters describing the optimum architecture when using each number of input measurements are shown for CAS in Table 6.12, and AAA in Table 6.13. The cost profiles of the CAS classifiers are shown in Figure 6.16, and AAA classifiers in Figure 6.17. A comparison of the F_1 scores achieved with and without early-stopping are shown for CAS classification in Table 6.14, and AAA classification in 6.15.

Table 6.12 shows an extremely high, almost remarkable, degree of consistency between the optimum hyper-parameters for each number of input measurements. Besides some relatively minor differences for the case of three input measurements, all other classifiers return an identical optimum architecture. This finding strongly supports the previous simplification of using a single architecture for all classifiers. An interesting finding to note, however, is that there is less consistency with the

No. measurements	Neurons	Depth	F_1
3	140	3	0.8817
4	180	4	0.8824
5	180	4	0.8355
6	180	4	0.8464

Table 6.12: The hyper-parameters describing the architecture of the MLP classifiers that produce the highest F_1 scores on the validation set with early-stopping for CAS classification, when using the best performing combination of three to six input measurements.

previous optimum hyper-parameters presented in Table 6.4, which found that four layers containing 60 neurons produced the highest F_1 score when providing six input measurements. There is also a relatively high degree of consistency in the optimum hyper-parameters for AAA classification (shown in Table 6.13), however to a lesser extent than CAS classification. Comparing Tables 6.13 and 6.4, it is again seen that the new hyper-parameters are inconsistent with the old. Initially this finding may seem to undermine the results achieved, however while the optimum hyper-parameters are inconsistent the F_1 scores achieved show much more similarity—0.9785 in Table 6.4 and 0.9870 in Table 6.13. The consistency in F_1 scores, despite the inconsistency in hyper-parameters, may suggest an unsusceptibility to the architecture used (*i.e.* the F_1 score plane in the two-dimensional grid search space is relatively flat, with statistical differences in the data sets used to train and test playing a more significant role than the hyper-parameters chosen). This again supports the simplification of using a single architecture for all classifiers.

Figure 6.16 shows that generally the early stopping criteria fulfils its purpose of stopping the training process near to the minimum validation cost point, minimising over-training (and hence over-fitting). It is seen that for all numbers of input measurements, training is stopped as soon as the 75 minimum iterations have been completed. While this early stopping criteria greatly reduces over-fitting in all cases, it is seen that the minimum number of training iterations (75) is too high for the six measurement case (the validation cost has already started to significantly rise), suggesting further refinement may reduce the validation and test costs even further. Looking at Table 6.14 it is seen that while early stopping has reduced the log loss cost

No. measurements	Neurons	Depth	F_1
3	140	2	0.9889
4	60	2	0.9858
5	150	1	0.9915
6	160	1	0.9870

Table 6.13: The hyper-parameters describing the architecture of the MLP classifiers that produce the highest F_1 scores on the validation set with early-stopping for AAA classification, when using the best performing combination of three to six input measurements.

No. of input measurements	Combination	F_1 score	
		Without early-stopping	With early-stopping
3	(P_3, P_2, P_1)	0.8831	0.8621
4	(Q_3, Q_1, P_2, P_1)	0.8683	0.8693
5	$(Q_3, Q_2, P_3, P_2, P_1)$	0.8463	0.7975
6	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.7785	0.8394

Table 6.14: F_1 scores achieved when using the three to six input measurement combinations found to produce the highest accuracies for CAS, when assessing convergence on the training set (*i.e.* the original results presented in Table 6.6) and on the validation set (*i.e.* employing early stopping).

across the test set, this does not generally translate to a difference in the F_1 score. The log loss cost will decrease without increasing the F_1 score if easy to classify patients are predicted with a higher degree of certainty (for example predicting 95% rather than 75%), however no new additional patients are correctly classified. Some increase in F_1 score is seen in Table 6.14 for the case when all six input measurements are used.

Figure 6.17 shows no major signs of over-fitting when using MLP classifiers to detect AAA. As a result, the employment of an early-stopping criteria has little affect on the final log loss cost achieved across the test set. From Table 6.15 it is seen that, as no real differences are achieved in the log loss costs, there is no significant differences in the F_1 scores achieved when employing early-stopping for AAA classification. These findings, for both CAS and AAA classification, suggest

No. of input measurements	Combination	F_1 score	
		Without early-stopping	With early-stopping
3	(Q_1, P_2, P_1)	0.9827	0.9852
4	(Q_2, Q_1, P_2, P_1)	0.9800	0.9784
5	$(Q_2, Q_1, P_3, P_2, P_1)$	0.9808	0.9876
6	$(Q_3, Q_2, Q_1, P_3, P_2, P_1)$	0.9785	0.9836

Table 6.15: F_1 scores achieved when using the three to six input measurement combinations found to produce the highest accuracies for AAA, when assessing convergence on the training set and on the validation set.

that to significantly improve the accuracy of MLP classifiers more extensive hyperparameter optimisation is required. This is prohibitive for an initial exploratory study, and hence suggests the unsuitability of the MLP (and more generally deep learning) methods.

6.4 Conclusions

The main conclusion of this study is that machine learning methods are suitable for detection of arterial disease—both stenoses and aneurysms—from peripheral measurements of pressure and flow-rates across the network. Amongst various ML methods, it is found that tree-based methods of Random Forest and Gradient Boosting perform best for this application. Across the different forms of disease, the Gradient Boosting method outperforms Random Forest, Support Vector Machine, Naive Bayes, Logistic Regression, and even the deep learning method of Multi-layer Perceptron.

The results of this study, generally, support those found in the PoC. Similarities between both include: the Naive Bayes method appears poorly suited to haemodynamic distinction; the ability to form non-linear partitions appears to be an important characteristic of the ML method used; certain input measurements are particularly informative; and high accuracy can be achieved when using even a low number of input measurements, if these measurements are chosen judiciously. The most obvious difference seen in this study, relative to the PoC, is the high accuracy achieved by tree based methods. This, however, is likely due to the inadequate

architecture optimisation in the PoC study.

While the maximum scores are for the case when all the six measurements are used, it is also shown that the performance degradation is less than 5% when using only three measurements and less than 10% when using only two measurements, as long as these measurements are carefully chosen in specific combinations. For the case of AAA, it is further demonstrated that when only a single measurement (either on the left or right side) is used, F_1 scores larger than 0.85 and corresponding sensitivities and specificities larger than 85% are achievable. This aspect encourages the application of AAA monitoring and/or screening through the use of a wearable device, such as the TLT Sapphire monitor (Tarilian Laser Technologies, Welwyn Garden City, U.K.) [125], although requires validation when multiple forms of disease are present. Finally, it is shown through the analysis of several classifiers and feature-importance that, amongst the measurements, the carotid artery flow-rate is a particularly informative measurement to detect the presence of all the four forms of disease considered.

In this chapter ML classifiers have been created using representations of the measurable pressure and flowrate profiles. To extend this analysis, the next chapter of this thesis uses these haemodynamic profiles to create easily acquirable physics advised features. These physics advised features are then used both in isolation and in combination with the measurable profiles to train and test ML classifiers.

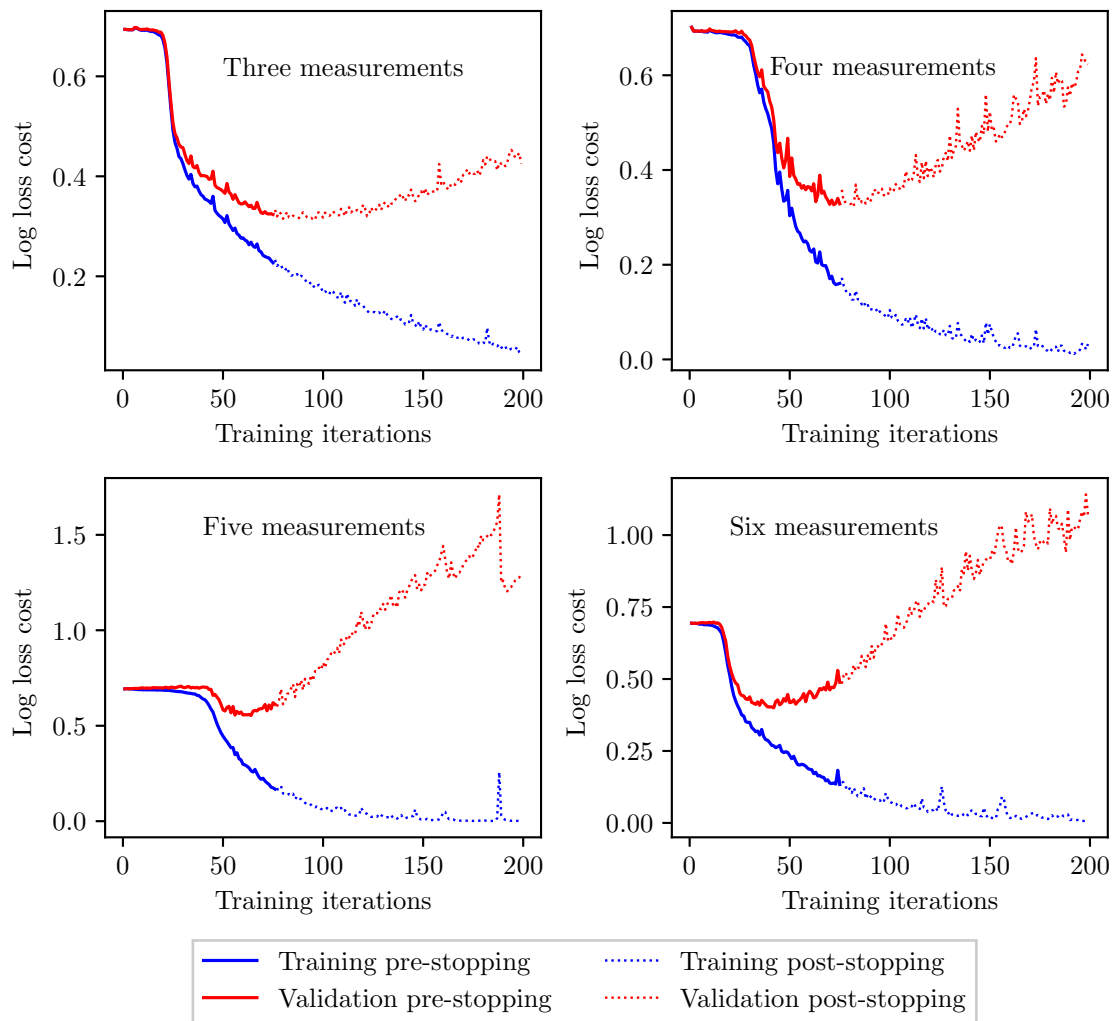


Figure 6.16: The log loss cost profiles across the training and test sets when using the best performing combination containing three to six input measurements for CAS classification and employing early stopping.

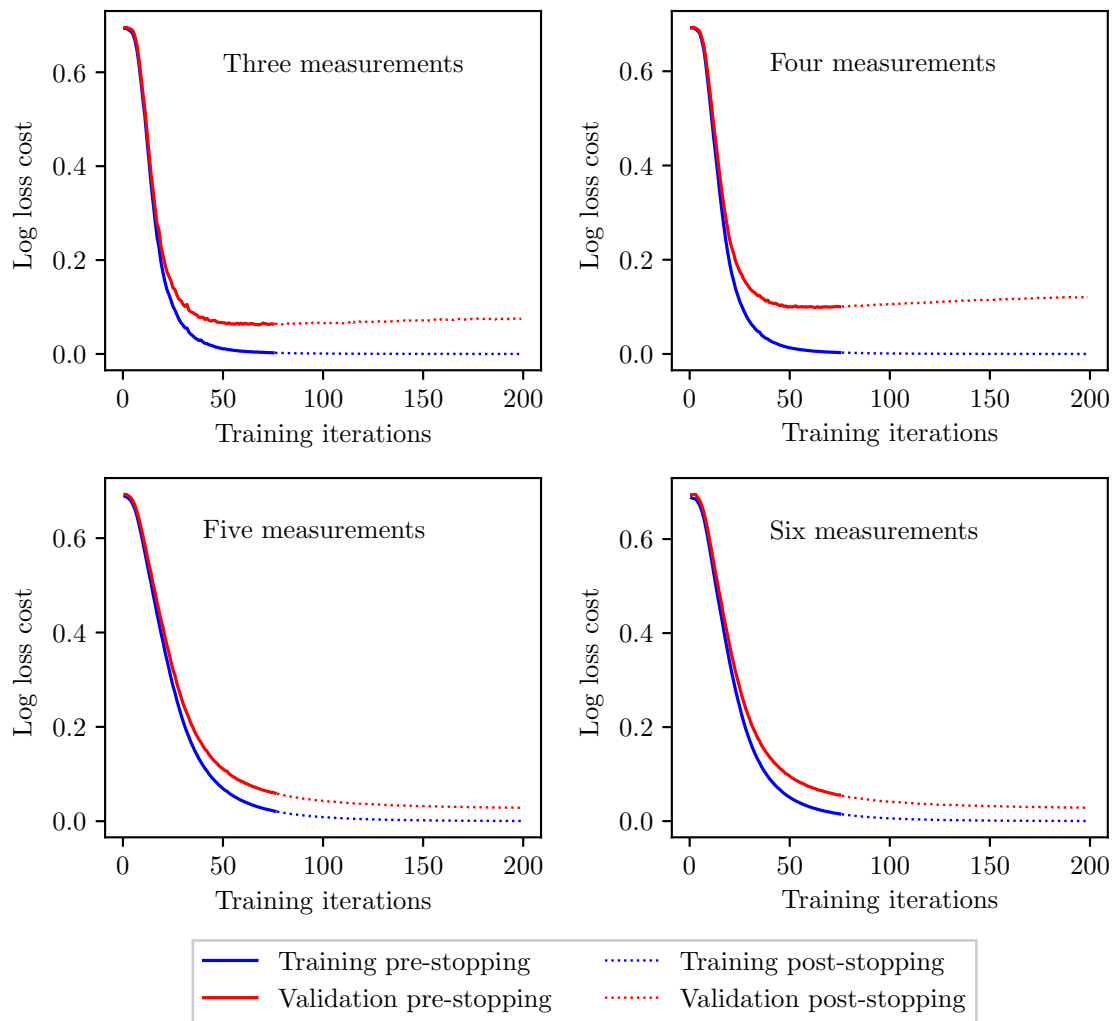


Figure 6.17: The log loss cost profiles across the training and test sets when using the best performing combination containing three to six input measurements for AAA classification and employing early stopping.

Chapter 7

Use of physics advised haemodynamic features

7.1 Introduction

Hitherto, pressure and flow-rate profiles have been used in their raw form—*i.e.* a representation of the measurable haemodynamic profiles—to train and test ML classifiers. It is possible, however, to use these raw pressure and flow-rate profiles to construct complex physics advised features (PAFs). Examples of such PAFs include; ankle-brachial pressure index (ABPI) [5], pulse wave velocity (PWV) [12], and ballistocardiogram waveforms (BCG) [93]. To expand on the results previously achieved, this chapter now uses raw pressure and flow-rate profiles to construct PAFs. ML classifiers are then trained and tested using these PAFs both in isolation, and combination with the raw pressure and flow-rate measurements.

The use of PAFs may increase the maximum accuracy of classification. While some ML methods are capable of combining input measurements to create high order features, the direct provision of PAFs may aid in classification. It is found in Chapter 6 that when using all six measurements maximum F_1 scores larger than 0.9 are achieved for CAS and PAD, larger than 0.85 for SAS, and larger than 0.98 for both low- and high-severity AAAs. Due to the high accuracy achieved when using raw pressure and flow-rate profiles it is unlikely that the use of PAFs will significantly increase the maximum classification accuracy. Instead, the primary advantage is likely to be the need for fewer individual measurements. While it is found in Chapter 6 that AAA can be classified to a high level of accuracy when using a single pressure or flow-rate measurement, the accuracy of CAD, SAS, and PAD is seen to degrade by $\approx 25\%$ when using a single measurement, relative to the maximum accuracy achieved. This suggests that the accuracy of classification of CAD, SAS, and PAD is reliant on complex inter-profile relationships. It maybe possible for PAFs to capture these relationships in a single feature. This would, in turn, allow for high accuracy classification to be achieved using fewer individual

measurements taken from patients.

The remainder of this chapter is structured as follow: first, the PAFs being examined (ABPI, PWV, and BCG) are outlined; next how these features are used to create ML classifiers, using the binary configuration in Chapter 6, is presented; and finally, the accuracy of disease classification is reported and analysed.

7.2 Methodology

7.2.1 Physics advised features

Three easily acquirable PAFs—ABPI, PWV, and BCG—are used to train and test ML classifiers to predict the presence of four forms of arterial disease—CAS, SAS, PAD, and AAA. In this section the three aforementioned features are presented, including: the physiological behaviour they capture; how they are computed from the raw pressure and flowrate profiles; and any *a priori* expectations of which form of disease each feature is likely to be most informative about. The mathematical models describing the PAFs must be carefully chosen to balance physiological realism with complexity. The PAFs, particularly BCG waveforms (see Section 7.2.1), capture complex physiological behaviours. The purpose of this thesis is to perform an initial exploratory study into the potential of using ML classifiers to predict the presence of arterial disease, and thus the mathematical models employed here favour simplicity over realism. This allows for an initial assessment of the viability of using PAFs, which if successful can be extended in future work through the use of more complex mathematical models.

Ankle-brachial pressure index

The ankle-brachial pressure index (ABPI) represents the ratio of the systolic blood pressure in the ankle, most commonly measured in the posterior tibial, to the systolic blood pressure in the brachial artery [5]. The locations of these vessels are shown in Figure 7.1. The ABPI is equal to:

$$\text{ABPI} = \frac{\text{Ankle systolic blood pressure}}{\text{Brachial systolic blood pressure}}. \quad (7.1)$$

ABPI is already clinically used to assess arterial health, most commonly to detect PAD [218]. To predict the health of a patient from their ABPI the Society for Vascular Technology suggests applying the discrete classifications [5]:

- $ABPI > 1.4$, calcification may be present.
- $ABPI > 1.0$, probably no arterial disease.
- $ABPI \in [0.81 - 1.0]$, no significant arterial disease, or mild/insignificant disease.
- $ABPI \in [0.5 - 0.8]$, moderate disease.
- $ABPI < 0.5$, severe disease.
- $ABPI < 0.3$, critical ischemia.

ABPI significantly differs from the raw pressure and flow-rate profiles used in Chapters 3 and 6 as it is a single scalar measurement, rather than time varying. Intuitively it would seem that ABPI would not capture enough information to accurately predict the presence of arterial disease due to the natural variability in systolic pressure throughout the arterial network, and the non-unique nature of ABPI (*i.e.* moderate ankle pressure and high brachial pressure will produce similar values of ABPI to low ankle pressure and moderate brachial pressure). A previous meta-analysis study into the performance of ABPI as a diagnostic method for PAD, however, has shown high levels of specificity (83.3 – 99.0%) and varying levels of sensitivity (15 – 79%) [218]. Thus, it may in fact be possible to predict the presence of at least one form of arterial disease (PAD) using ABPI and ML classifiers.

There are two instances of ABPI per patient, on the right and left side on the body. The right and left instance of ABPI are referred to as $ABPI_R$ and $ABPI_L$, respectively. Computation of either the right or left instance requires measurements of pressure in the corresponding brachial artery and ankle, and thus two individual measurements from a patient.

Pulse wave velocity

Pulse wave velocity (PWV) is a measure of the rate at which pressure waves propagate through the arterial network [47]. PWV is found by recording the time delay (referred to as the transit time) between pressure pulses at different locations in

the arterial network which are at known distances apart. Transit time is often found by recording the time delay between the ‘foot’ waveform (the commencement of the systolic upstroke) [198] at different locations in the arterial network. While PWV can be measured between several different vessels in the network, here the transit time is recorded between the brachial artery and the ankle, as these measurements can be easily obtained [219]. For consistency with ABPI, it is assumed that ankle pressure is measured in the posterior tibial.

It is infeasible to manually determine the foot waveform for all 28,868 VPs. A commonly used automated method to determine the foot waveform—that finds the intersection between the tangent of the minimum pressure and the tangent of the maximum pressure gradient—is presented in [84]. This method is not suitable for the VPD due to the variability in pressure profiles. Examples of brachial artery and ankle pressure profiles taken from VPs for which determination of the footwaves through the aforementioned tangent method is suitable is highlighted in Figure 7.2, and problematic is highlighted in Figure 7.3. One method to improve the reliability of the determination of the foot waveform could be to incorporate physics based knowledge into the geometrical prediction—for example estimating the time point at which the footwave will occur through the use of the physics based model of pulse wave propagation (see Chapter 2), and then examining the geometrical properties of the pressure profile around this estimated time point. Employment of such a method, however, is not feasible within the scope of this thesis. Instead, to simplify implementation, the time point corresponding to peak systolic pressure is used as an easily acquirable surrogate to the foot waveform. The transit time is then estimated to be the time point corresponding to peak systolic pressure in the brachial artery minus the time point corresponding to peak systolic pressure in the ankle. While this method is simple, two issues associated with it are: peak systolic pressure may occur at the same time point in both the brachial artery and ankle in certain VPs; and if the transit time is large enough, so that the systolic pressures are an entire cardiac period out of phase, the transit time will appear negative. Determination of the PWV of VPs with either of these characteristics is not possible through the method being employed, and so these VPs are removed from the VPD when using PWV to train and test ML classifiers. Approximately 40% of each of the four combined data sets described in Section 7.2.3 are discarded due to the above characteristics when using PWV.

It is suggested in [219] that the distance between the brachial and ankle measurement locations can be computed from superficial measurements of a patients height using the equation:

$$L_{ba} = 0.5933h + 14.40139, \quad (7.2)$$

where L_{ba} represents the distance between the brachial arteries and the ankle, and h represents the height of the patient (in cm). The height of each VP is calculated by multiplying the height of the reference ADAN patient (170cm) by the vessel length scaling term applied (as it is assumed that the length of arterial vessels are proportional to the height of the patient, as detailed in Section 4.6.2). From the recorded transit time Δt_{ba} and distance between measurement locations L_{ba} the PWV is computed using the relationship:

$$\text{PWV} = \frac{L_{ba}}{\Delta t_{ba}}. \quad (7.3)$$

There are no strong *a priori* expectations of any form of disease for which PWV should be particularly informative. As with ABPI, there are two instances of PWV per patient, one on each side of the body. The right and left instance of PWV are referred to as PWV_R and PWV_L , respectively. As with ABPI computation of each of these instances requires measurement of pressure in the corresponding brachial artery and ankle.

Ballistocardiogram waveforms

The ballistocardiogram (BCG) waveform represents the reactionary force experienced by the body due to the propagation of blood through the arterial network. A simple mathematical model of the BCG waveform is presented in [108]. This model is created by analysing the forces exerted on the blood in the aorta (the largest arterial vessel in the human body), with the following assumptions: the aorta can be approximated as two segments, a short ascending segment (ascending aorta) and a long descending segment (descending aorta); blood is homogeneous and incompressible; there is little change to the cross sectional area of the arterial vessel (*i.e.* the vessel is rigid and geometric tapering is small); and blood flow is

longitudinal with uniform velocity profile. The resulting mathematical model is:

$$\begin{aligned}
 F_{\text{BCG}}(t) \approx & \\
 & A_D \left[\left(P_1(t) + \rho \left(\frac{Q_1(t)}{A_D} \right)^2 \right) - \left(P_2(t) + \rho \left(\frac{Q_2(t)}{A_D} \right)^2 \right) \right] \\
 & - A_A \left[\left(P_0(t) + \rho \left(\frac{Q_0(t)}{A_A} \right)^2 \right) - \left(P_1(t) + \rho \left(\frac{Q_1(t)}{A_A} \right)^2 \right) \right], \tag{7.4}
 \end{aligned}$$

where $F_{\text{BCG}}(t)$ represents the BCG force at time t ; ρ represents the density of blood; A_D and A_A represent the area of the descending and ascending aortic segments; P_0 , P_1 , and P_2 represent the pressure at the inlet of the ascending segment, the boundary of the ascending and descending segments (*i.e.* the outlet of the ascending and inlet of the descending), and the outlet of the descending segment respectively; and Q_0 , Q_1 , and Q_2 represent the corresponding flow-rate measurements. The locations in the network at which P_0 , P_1 , P_2 , Q_0 , Q_1 , Q_2 are measured is shown in Figure 7.4. Aortic blood velocities are around $0.45\text{--}0.6\text{ms}^{-1}$, and thus $\rho \left(\frac{Q(t)}{A} \right)^2$ yields values of $1.6\text{--}1.9\text{mmHg}$. Blood pressure in the aorta is typically around 100mmHg , and so it is seen that this is primary mechanism for the BCG waveform. The mathematical model in [108] is, therefore, reduced to:

$$F_{\text{BCG}}(t) \approx A_D (P_1(t) - P_2(t)) - A_A (P_0(t) - P_1(t)). \tag{7.5}$$

This model must be adapted for application to VPs to account for: the area of the aorta changes both spatially (along the length of the aorta) and temporally (throughout the cardiac period); and the velocity profile of blood is parabolic in the physics based model of pulse wave propagation. Including these adaptations, Equation (7.4) is written as:

$$\begin{aligned}
 F_{\text{BCG}}(t) \approx & \\
 & \left[A_1(t) \left(P_1(t) + \frac{4}{3}\rho \left(\frac{Q_1(t)}{A_1(t)} \right)^2 \right) - A_2(t) \left(P_2(t) + \frac{4}{3}\rho \left(\frac{Q_2(t)}{A_2(t)} \right)^2 \right) \right] \\
 & - \left[A_0(t) \left(P_0(t) + \frac{4}{3}\rho \left(\frac{Q_0(t)}{A_0(t)} \right)^2 \right) - A_1(t) \left(P_1(t) + \frac{4}{3}\rho \left(\frac{Q_1(t)}{A_1(t)} \right)^2 \right) \right], \tag{7.6}
 \end{aligned}$$

where A_0 , A_1 , and A_2 represent the cross sectional area at the inlet of the ascending segment, the boundary of the ascending and descending segments, and the outlet of the descending segment respectively. The locations at which A_0 , A_1 , and A_2 are measured are included in Figure 7.4. Removing the momentum terms, this model becomes:

$$F_{\text{BCG}}(t) \approx 2A_1(t)P_1(t) - A_0(t)P_0(t) - A_2(t)P_2(t). \quad (7.7)$$

BCG waveforms are computed for five VPs when using the initial model described by Equation (7.5); allowing only spatial cross sectional area variation; and allowing both spatial and temporal cross sectional area variation. This comparison is shown in Figure 7.5. When using constant cross sectional areas A_D and A_A are computed as the average of the mean cross sectional area through out the cardiac period at the inlet and outlet of each segment respectively, *i.e.*:

$$A_A = \frac{1}{2} \left(\frac{1}{N} \sum_{n=1}^N A_0(t = t_n) + \frac{1}{N} \sum_{n=1}^N A_1(t = t_n) \right), \quad (7.8)$$

where $A_x(t = t_n)$ represents the cross sectional area at location x and time t_n , respectively; and N represents the total number of time points in the cardiac period. Similarly, A_D is computed using Equation (7.8), however with A_0 replaced by A_2 . To compute the BCG waveforms when only allowing spatial cross sectional area variation, the mean cross sectional area through out the cardiac period at each location is used. Figure 7.5 shows the BCG waveforms computed using a constant cross sectional area show good consistency with those in [108], which in turn have been compared and show good agreement with measurements from real subjects. When allowing spatial cross sectional area variation the BCG waveform is smoothed and the magnitude of the force increased, relative to the constant area waveform. Very little difference is seen in the BCG waveform produced when either assuming that vessels are rigid, or using the time varying cross sectional area profiles. Thus, it is seen that when more complexity is added to the model describing BCG waveforms the resulting profiles show greater differences to those taken from real subjects. This is counter-intuitive, as the inclusion of more complexity and more physiologically realistic behaviour is expected to produce a more realistic model. This suggests that more work needs to go into creating a comprehensive, realistic model for BCG waveforms. While Equation (7.6) amends the model in [108] to

accommodate parabolic velocity profiles and varying cross sectional areas, it still does not account for: momentum changes within the control volume; the curvature of the aorta; and other large arterial vessels such as the axillary, carotid, and iliac arteries.

Due to the pressure gradient in the aorta being the primary mechanism responsible for the BCG waveform, this PAF is expected to be particularly informative for AAA classification. Unlike ABPI and PWV, the BCG waveform is directly measured (rather than computed from pressure and flow-rate measurements), and so requires a single measurement from patients. Multiple methods have been proposed for non-invasive measurement of BCG waveforms including the use of sensors built into chairs [111], and through modified bathroom scales [70, 216]. These methods suggest strong potential for home monitoring of arterial health through BCG waveforms, if this PAF captures sufficient information about the presence of arterial disease.

7.2.2 Available measurements

In Chapter 6 the locations of six easily acquirable haemodynamic measurements—pressure in the carotid, brachial, and radial arteries; and flow-rate in the carotid, brachial, and femoral arteries—are detailed, and the bilateral provision of these measurements, described through a Fourier series (FS), is presented. These six measurements are made available to ML classifiers in an identical form to Chapter 6—*i.e.* bilaterally and described by a FS—and each measurement referred to by the same notation, Q_1 to Q_3 and P_1 to P_3 (see Chapter 6 for details).

Along with these six bilateral pressure and flow-rate measurements, there are now three PAFs available to ML classifiers, two of which (ABPI and PWV) have a right and left instance. It is chosen to provide PAFs to ML classifiers unilaterally and mutually exclusively—*i.e.* when ABPI is used ML classifiers are provided with $ABPI_x$, where x can be either R or L . These features are provided unilaterally and mutually exclusively as the expected primary advantage of using PAFs is a reduction in the number of individual measurements required. For both ABPI and PWV, computation of either the right or left instance requires measurements of pressure in the corresponding brachial artery and ankle, and thus using either of these features bilaterally requires four individual measurements to be taken from

patients (two in the right side, and two in the left side), significantly mitigating the advantage of using PAFs over raw pressure and flow-rate measurements.

As the BCG is a periodic time varying profile, to reduce the dimensionality this feature is described through a FS truncated at the 5th order (in an identical manner to the raw pressure and flow-rate profiles). Thus, the BCG profile is described by 11 FS coefficients (see Chapter 3 for further detail). Both the ABPI and PWV are scalar values, and so are described by a single dimension. As with the PoC and Chapter 6, to account for large differences in magnitudes of the different inputs provided to ML classifiers, they are individually transformed with the Z-score standardisation method [136] to have zero-mean and unit variance.

7.2.3 Machine learning classifiers

The four unhealthy VPDs (VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , and VPD_{AAA}) and one healthy VPD (VPD_H)—created in Chapters 4 and 5; and employed in Chapter 6—are now used to create ML classifiers for each of the four forms of arterial disease, using the previously described PAFs. In an identical manner to Chapter 6, the ML classifiers are created for each disease form independently. All classifiers are therefore binary (see Chapters 3 and 6 for details), *i.e.* four independent classifiers are trained to predict the following questions independently: “Does a VP belong to VPD_H or VPD_x ”, where x can be either CAS, SAS, PAD, or AAA.

Training, validation, and test sets

Combined data sets, to be provided to ML classifiers, containing 50% healthy and 50% unhealthy VPs are created using an identical methodology to Chapter 6, *i.e.*:

- **Step 1:** Half of the available VPs are randomly selected from VPD_H for inclusion within the ML data set; this is referred to as VPD_{H-ML} . The unhealthy VPs corresponding to the remaining unused half are taken from the appropriate unhealthy VPD (VPD_{CAS} , VPD_{SAS} , VPD_{PAD} , or VPD_{AAA}) and incorporated into the ML data set. These data sets are referred to as VPD_{CAS-ML} , VPD_{SAS-ML} , VPD_{PAD-ML} , or VPD_{AAA-ML} .
- **Step 2:** The data sets of Step 1 are combined to create four complete data sets each containing 50% healthy and 50%, unhealthy VPs:

1. $\text{VPD}_{\text{H-ML}} \cup \text{VPD}_{\text{CAS-ML}}$
2. $\text{VPD}_{\text{H-ML}} \cup \text{VPD}_{\text{SAS-ML}}$
3. $\text{VPD}_{\text{H-ML}} \cup \text{VPD}_{\text{PAD-ML}}$
4. $\text{VPD}_{\text{H-ML}} \cup \text{VPD}_{\text{AAA-ML}}$

In Chapters 6 the four complete data sets of **Step 2** are broken down into two components, a training set (containing 2/3 of all the VPs) and a test set (containing the remaining used 1/3 of all the VPs). It chosen to split the data set into two components, rather than the more commonly used three, as repeated hyper-parameter optimisation is not performed when the input measurement combinations are varied.

In this chapter, one of the primary aims is to assess the accuracy of disease detection when using PAFs in isolation. Due to the lower number of different input combinations available to ML classifiers when using PAFs in isolation, hyper-parameter optimisation is performed for each (see Section 7.2.4). It is, there, important to minimise any inflation in the accuracy of classification achieved when using PAFs in isolation caused by bias towards to the test set. To minimise this, the complete data sets are split into three components:

- A training set, containing 60% of the available data. All ML classifiers are fit on the same training set.
- A validation set containing 20% of the data. When tuning the hyper-parameters describing the architectures of classifiers to each individual PAF in Section 7.2.4, the validation set is used to assess the accuracy.
- A test set containing the remaining 20% of the data. The final accuracies of classification when using both PAFs in isolation and in combination with raw pressure and flow-rate profiles, presented in Section 7.3, are assessed on this test set.

The performance of all ML classifiers is evaluated using a five fold validation (see Chapters 3 and 6 for details).

ML methods

In this chapter only the gradient boosting (GB) method is used, to reduce computational expense. This method is chosen as it found to consistently produce the highest accuracy classification in the previous chapter (Chapter 6). The implementation of the GB algorithm in the Python package Scikits-learn [156] is used. The hyper-parameter optimisation of GB classifiers is described after presenting performance quantification metrics in the next section.

Quantification of results

Classifier performance is assessed by the same two metrics employed in Chapters 3 and 6: *sensitivity* and *specificity* in combination; and the F_1 score. Higher values of all three metrics are desirable. Since the F_1 score is a single scalar metric that balances both precision and recall, it is a good metric to compare classifiers when tuning the hyper-parameters of ML algorithms, described next.

7.2.4 Hyper-parameter optimisation

As detailed in Chapter 6 GB architecture is optimised for a specific problem by varying the number of trees within the ensemble and the maximum depth of each tree. Previously, hyper-parameter optimisation is not performed for each individual configuration of ML classifiers—*i.e.* when the input pressure and flow-rate measurements provided are varied—due to the large number of different input measurement combinations used, and so consequently high computational expense associated with tuning the hyper-parameters to each. When PAFs are used in combination with raw pressure and flow-rate profiles it remains infeasible, within the scope of this thesis, to perform hyper-parameter optimisation for every classifier configuration. Thus, the optimal number of trees and depth of each tree found for each disease form in Chapter 6 are reused here when PAFs are used in combination with raw pressure and flow-rate profiles.

When ML classifiers are created using PAFs in isolation it is feasible to perform configuration specific hyper-parameter optimisation. Three different PAFs are used, two of which (ABPI and PWV) have a left and right instance. Due to imposed condition of PAFs being provided unilaterally and mutually exclusively (see Section

7.2.2) there are five different configurations of PAFs per disease form (ABPI_R, ABPI_L, PWV_R, PWV_L, BCG)—a feasible number to perform hyper-parameter optimisation on. The optimisation of the hyper-parameters to each of these 20 configurations (four disease forms each with five configurations) is presented next.

ABPI

A grid search is carried out to find the combination of hyper-parameters that produces highest F_1 score when using the right and left ABPI in isolation for each form of disease. The possible depth of each tree is discretised between 2 and 20 at intervals of 1, and the possible number of trees within the ensemble is discretised between 10 and 100 at intervals of 10. The optimum hyper-parameter combination for each disease form is detailed in Table 7.1, along with the F_1 scores achieved.

Table 7.1 shows that, generally, the optimum hyper-parameters and the resulting maximum F_1 scores are relatively consistent when using either the right or left ABPI for each disease form. A further observation is that the maximum F_1 score is only marginally better than naive classification (*i.e.* an F_1 score of 0.5) for all disease forms, excluding PAD. Further analysis is not performed on the accuracies presented in Table 7.1, as these results may show a bias to the validation set. Analysis of the accuracy of classification on the unbiased test set is presented in 7.3.1.

Disease	Side	Trees	Depth	F_1
CAS	Right	60	3	0.5388
	Left	40	4	0.5588
SAS	Right	30	2	0.5980
	left	40	4	0.5497
PAD	Right	10	2	0.6435
	Left	10	2	0.6161
AAA	Right	20	3	0.5760
	Left	20	2	0.5851

Table 7.1: The hyper-parameters describing the architecture of the GB classifiers that produce the highest F_1 scores, when using either the right or left ABPI.

PWV

Identical grid searches are carried out to those presented above, however using the right and left PWVs rather than ABPIs. The optimum hyper-parameter combination for each disease form, and corresponding F_1 scores, are detailed in Table 7.2.

Disease	Side	Trees	Depth	F_1
CAS	Right	10	3	0.5109
	Left	10	3	0.62764
SAS	Right	10	2	0.5633
	left	10	2	0.6771
PAD	Right	40	4	0.5997
	Left	40	3	0.5182
AAA	Right	40	4	0.5668
	Left	60	4	0.6545

Table 7.2: The hyper-parameters describing the architecture of the GB classifiers that produce the highest F_1 scores, when using either the right or left PWV.

Table 7.2 shows that, while some improvement has been made relative to ABPI, the maximum F_1 scores achieved are still only marginally better than naive classification for all forms of disease. The fact that both ABPI and PWV are achieving accuracies only marginally better than naive classification suggests that, as is *a priori* expected, there is insufficient information captured in scalar measurements to distinguish between healthy and unhealthy patients.

Table 7.2 shows that while the optimum hyper-parameters are, generally, consistent when using either the right or left instance of PWV, the left instance produces higher F_1 scores for all forms of disease excluding PAD. The exact cause of this imbalance in accuracy is unclear. It is unlikely that the higher accuracies achieved using the left PWV is due to physiological behaviours as the arterial network is, generally, symmetrical. A more likely explanation is that, due to the lack of informative available from the PWVs, the accuracies recorded are susceptible

to statistical errors and the bias of the F_1 score towards sensitivity over specificity. Analysis on the accuracy of classification on the unbiased test set is presented in Section 7.3.2

BCG

As with ABPI and PWV, a grid search is carried out to find the combination of hyper-parameters that produce the highest F_1 score when using BCG waveforms to predict each form of disease (employing the same grids as those in Sections 7.2.4 and 7.2.4). The hyper-parameters describing the optimum architecture for each disease form are outlined in Table 7.3.

Disease	Trees	Depth	F_1
CAS	70	9	0.5413
SAS	50	7	0.5252
PAD	50	8	0.5984
AAA	80	9	0.9741

Table 7.3: The hyper-parameters describing the architecture of the GB classifiers that produce the highest F_1 scores, when using the BCG waveform.

Table 7.3 shows that the accuracy of CAS, SAS, and PAD classification are, again, only marginally better than naive classification. The accuracy of AAA classification, however, is much higher. Further analysis on the unbiased test set is presented in Section 7.3.3.

7.3 Results and discussion

7.3.1 ABPI

ML classifiers are trained and tested using ABPI first in isolation and then in combination with the pressure and flowrate measurements found to produce the highest accuracy when providing one to six measurements (see Chapter 6 for details). To reduce computational expense this analysis is only carried out using

Combination	F_1 score	Sens.	Spec.
(ABPI _R)	0.5332	0.5724	0.4255
(Q_1)	0.8521	0.8547	0.8502
(Q_1 , ABPI _R)	0.8531	0.8530	0.8533
(P_2 , P_1)	0.8950	0.9026	0.8889
(P_2 , P_1 , ABPI _R)	0.8944	0.8877	0.9027
(Q_1 , P_2 , P_1)	0.9389	0.9433	0.9351
(Q_1 , P_2 , P_1 , ABPI _R)	0.9359	0.9350	0.9370
(Q_3 , Q_1 , P_2 , P_1)	0.9395	0.9417	0.9376
(Q_3 , Q_1 , P_2 , P_1 , ABPI _R)	0.9351	0.9302	0.9406
(Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9391	0.9416	0.9370
(Q_2 , Q_1 , P_3 , P_2 , P_1 , ABPI _R)	0.9340	0.9312	0.9372
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9343	0.9364	0.9325
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , ABPI _R)	0.9398	0.9425	0.9369

Table 7.4: The F_1 scores achieved when using ABPI in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for CAS). The corresponding sensitivities and specificities are also included.

the right ABPI, and it is seen in Section 7.2.4 that the right and left instances of ABPI produce similar accuracies. The F_1 scores, sensitivities, and specificities achieved for CAS classification are shown in Table 7.4, SAS classification in Table 7.5, PAD classification in Table 7.6, and AAA classification in Table 7.7. The accuracies achieved in Chapter 6 for each combination of pressure and flow-rate profiles (without the inclusion of ABPI) are included in these tables to allow for comparison.

The four aforementioned tables (Tables 7.4, 7.5, 7.6, and 7.7) show that the ABPI provides no significant additional information about the health of VPs. It is seen that for all forms of disease (including PAD) the ABPI produces accuracies comparable to naive classification when used in isolation. Table 7.6 suggests that the marginally higher accuracies seen for PAD when tuning the hyper-parameters (see Section 7.2.4) are due to statistical errors and bias, rather than a true increase

Combination	F_1 score	Sens.	Spec.
(ABPI _R)	0.5080	0.5228	0.4646
(Q_1)	0.7529	0.7224	0.7714
(Q_1 , ABPI _R)	0.7473	0.7333	0.7709
(Q_2 , Q_1)	0.8461	0.8293	0.8585
(Q_2 , Q_1 , ABPI _R)	0.8408	0.8332	0.8513
(Q_3 , Q_2 , Q_1)	0.8552	0.8453	0.8626
(Q_3 , Q_2 , Q_1 , ABPI _R)	0.8498	0.8423	0.8601
(Q_3 , Q_2 , Q_1 , P_2)	0.8585	0.8487	0.8660
(Q_3 , Q_2 , Q_1 , P_2 , ABPI _R)	0.8514	0.8434	0.8624
(Q_3 , Q_2 , Q_1 , P_2 , P_1)	0.8600	0.8525	0.8657
(Q_3 , Q_2 , Q_1 , P_2 , P_1 , ABPI _R)	0.8608	0.8556	0.8679
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.8574	0.8504	0.8627
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , ABPI _R)	0.8591	0.8525	0.8679

Table 7.5: The F_1 scores achieved when using ABPI in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for SAS). The corresponding sensitivities and specificities are also included.

in accuracy. When the ABPI is used in combination with pressure and flow-rate profiles there is no increase in accuracy, relative to that achieved by the pressure and flow-rate profiles without inclusion of ABPI. These results suggest, that as is expected, a singular scalar measurement is unable to capture enough information to accurately distinguish between healthy and unhealthy VPs.

7.3.2 PWV

Similar to the analysis performed for ABPI, ML classifiers are trained and tested using PWV first in isolation and then in combination with the pressure and flowrate measurements found to produce the highest accuracy when providing one to six measurements. This analysis is performed using the left PWV as it is seen in Section 7.2.4 that the left instances of PWV generally produces higher accuracies than the right. The results achieved for CAS, SAS, PAD, and AAA classification are shown

Combination	F_1 score	Sens.	Spec.
(ABPI _R)	0.5514	0.6136	0.3880
(Q_1)	0.8183	0.8126	0.8214
(Q_1 , ABPI _R)	0.8533	0.8071	0.8286
(Q_3 , Q_1)	0.9041	0.8950	0.9117
(Q_3 , Q_1 , ABPI _R)	0.9030	0.8989	0.9080
(Q_3 , Q_2 , Q_1)	0.9168	0.9055	0.9265
(Q_3 , Q_2 , Q_1 , ABPI _R)	0.9172	0.9106	0.9252
(Q_3 , Q_2 , Q_1 , P_1)	0.9196	0.9068	0.9306
(Q_3 , Q_2 , Q_1 , P_1 , ABPI _R)	0.9141	0.9056	0.9243
(Q_3 , Q_2 , Q_1 , P_2 , P_1)	0.9170	0.9041	0.9281
(Q_3 , Q_2 , Q_1 , P_2 , P_1 , ABPI _R)	0.9179	0.9107	0.9266
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9187	0.9102	0.9261
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , ABPI _R)	0.9131	0.9048	0.9231

Table 7.6: The F_1 scores achieved when using ABPI in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for PAD). The corresponding sensitivities and specificities are also included.

in Tables 7.8, 7.9, 7.10, and 7.11 respectively.

Tables 7.8, 7.9, 7.10, and 7.11 show that the PWV provides very little information about the presence of arterial disease. The accuracies achieved for all forms of disease when using PWV in isolation is seen to be comparable to naive classification. The F_1 score achieved when using PWV to detect CAD is lower than naive classification (an F_1 score of 0.35 is achieved, compared to a naive value of 0.5), however this is seen to be due to an imbalance in the sensitivity and specificity. When using PWV in combination with pressure and flow-rate measurements there is no significant increase in the accuracies achieved, relative to when pressure and flow-rate measurements are used without the inclusion of PWV. The uninformative nature of both ABPI and PWV suggest that, as is *a priori* expected, a single scalar measurement does not capture enough information to allow for the classification of VPs.

Combination	F_1 score	Sens.	Spec.
(ABPI _R)	0.4413	0.3961	0.6012
(Q_1)	0.9805	0.9799	0.9811
(Q_1 , ABPI _R)	0.9806	0.9800	0.9814
(Q_2 , Q_1)	0.9928	0.9919	0.9938
(Q_2 , Q_1 , ABPI _R)	0.9930	0.9926	0.9936
(Q_3 , Q_2 , Q_1)	0.9962	0.9954	0.9970
(Q_3 , Q_2 , Q_1 , ABPI _R)	0.9958	0.9939	0.9975
(Q_3 , Q_2 , Q_1 , P_2)	0.9972	0.9959	0.9986
(Q_3 , Q_2 , Q_1 , P_2 , ABPI _R)	0.9971	0.9964	0.9979
(Q_3 , Q_2 , Q_1 , P_3 , P_2)	0.9970	0.9959	0.9981
(Q_3 , Q_2 , Q_1 , P_3 , P_1)		0.9963	0.9978
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , ABPI _R)	0.9971	0.9962	0.9981
(Q_3 , Q_2 , Q_1 , P_3 , P_1 , ABPI _R)	0.9965	0.9959	0.9973
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9970	0.9959	0.9981
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , ABPI _R)	0.9969	0.9959	0.9980

Table 7.7: The F_1 scores achieved when using ABPI in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for AAA. The corresponding sensitivities and specificities are also included.

7.3.3 BCG

The results achieved when using BCG waveforms both in isolation and in combination with the best performing pressure and flow-rate measurements are shown for CAS in Table 7.12, SAS in Table 7.13, PAD in Table 7.14, and AAA in Table 7.15. Tables 7.12, 7.13, and 7.14 shows that the BCG waveform contains very little information about the presence of CAS, SAS, and PAD. The accuracy of classification for these three forms of disease when using BCG in isolation is, again, only marginally better than naive classification. These tables also show no increase in the accuracy of classification when BCG is used in combination with pressure and flow-rate profiles, relative to the accuracy achieved when the pressure and flow-rate profiles are used without the inclusion of the BCG.

Combination	F_1 score	Sens.	Spec.
(PWV _L)	0.3516	0.2729	0.7207
(Q_1)	0.8521	0.8547	0.8502
(Q_1 , PWV _L)	0.8222	0.8620	0.7652
(P_2 , P_1)	0.8950	0.9026	0.8889
(P_2 , P_1 , PWV _L)	0.8756	0.8800	0.8700
(Q_1 , P_2 , P_1)	0.9389	0.9433	0.9351
(Q_1 , P_2 , P_1 , PWV _L)	0.9282	0.9236	0.9336
(Q_3 , Q_1 , P_2 , P_1)	0.9395	0.9417	0.9376
(Q_3 , Q_1 , P_2 , P_1 , PWV _L)	0.9295	0.9285	0.9308
(Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9391	0.9416	0.9370
(Q_2 , Q_1 , P_3 , P_2 , P_1 , PWV _L)	0.9298	0.9271	0.9331
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9343	0.9364	0.9325
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , PWV _L)	0.9240	0.9408	0.9318

Table 7.8: The F_1 scores achieved when using PWV in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for CAS). The corresponding sensitivities and specificities are also included.

Table 7.15, however, shows significantly better results when the BCG waveforms are used to classify AAA. Due to the very high accuracy achieved for AAA classification when using pressure and flow-rate measurements, there is little improved made when also including the BCG waveforms. The most important observation from Table 7.15 is the very high accuracy (F_1 score of 0.97) achieved when using the BCG waveforms in isolation. The accuracy achieved by the BCG in isolation is comparable to that achieved when using Q_1 , *i.e.* bilateral measurements of carotid flow-rate. It is important to consider the BCG waveform is directly measured, and so requires a single measurement from patients, whereas Q_1 requires measurements of flow-rate in both the right and left carotid artery. BCG waveforms are, therefore, achieving comparable accuracies while requiring half the number of individual measurements from patients.

As the BCG waveform requires a single measurement from patients it is fitting

Combination	F_1 score	Sens.	Spec.
(PWV _L)	0.4351	0.3869	0.6085
(Q ₁)	0.7529	0.7224	0.7714
(Q ₁ , PWV _L)	0.7477	0.7373	0.7652
(Q ₂ , Q ₁)	0.8461	0.8293	0.8585
(Q ₂ , Q ₁ , PWV _L)	0.8366	0.8245	0.8536
(Q ₃ , Q ₂ , Q ₁)	0.8552	0.8453	0.8626
(Q ₃ , Q ₂ , Q ₁ , PWV _L)	0.8436	0.8333	0.8579
(Q ₃ , Q ₂ , Q ₁ , P ₂)	0.8585	0.8487	0.8660
(Q ₃ , Q ₂ , Q ₁ , P ₂ , PWV _L)	0.8518	0.8414	0.8659
(Q ₃ , Q ₂ , Q ₁ , P ₂ , P ₁)	0.8600	0.8525	0.8657
(Q ₃ , Q ₂ , Q ₁ , P ₂ , P ₁ , PWV _L)	0.8548	0.8483	0.8637
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , P ₁)	0.8574	0.8504	0.8627
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , P ₁ , PWV _L)	0.8540	0.8480	0.8621

Table 7.9: The F_1 scores achieved when using PWV in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for SAS). The corresponding sensitivities and specificities are also included.

to compare it to the unilateral classification accuracies achieved in Chapter 6. This comparison is made in Table 7.16. Table 7.16 shows that the BCG out performs all unilateral measurements. These results suggest strong potential for the home monitoring of AAAs through the use of BCG waveforms, especially when considered along the side the numerous non-invasive methods proposed for BCG measurement [111, 70, 216].

7.4 Conclusions

The results achieved in this chapter lead to two conclusions:

1. There is insufficient information captured in a singular scalar measurement to allow for distinction between healthy and unhealthy VPs. the results of Sections 7.3.1 and 7.3.2 suggest that ABPI and PWV provide no additional

Combination	F_1 score	Sens.	Spec.
(PWV _L)	0.4654	0.4322	0.5750
(Q ₁)	0.8183	0.8126	0.8214
(Q ₁ , PWV _L)	0.8190	0.8145	0.8256
(Q ₃ , Q ₁)	0.9041	0.8950	0.9117
(Q ₃ , Q ₁ , PWV _L)	0.8900	0.8854	0.8958
(Q ₃ , Q ₂ , Q ₁)	0.9168	0.9055	0.9265
(Q ₃ , Q ₂ , Q ₁ , PWV _L)	0.9069	0.8977	0.9182
(Q ₃ , Q ₂ , Q ₁ , P ₁)	0.9196	0.9068	0.9306
(Q ₃ , Q ₂ , Q ₁ , P ₁ , PWV _L)	0.9126	0.9063	0.9203
(Q ₃ , Q ₂ , Q ₁ , P ₂ , P ₁)	0.9170	0.9041	0.9281
(Q ₃ , Q ₂ , Q ₁ , P ₂ , P ₁ , PWV _L)	0.9070	0.9013	0.9139
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , P ₁)	0.9187	0.9102	0.9261
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , P ₁ , PWV _L)	0.9068	0.9005	0.9145

Table 7.10: The F_1 scores achieved when using PWV in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for PAD). The corresponding sensitivities and specificities are also included.

information about the presence of arterial disease, relative to that extracted purely from raw pressure and flow-rate profiles. This finding suggests that further investigation into the use of ABPI and PWV, employing the use of more complex mathematical models, is not warranted.

2. There is significant potential in the use of BCG waveforms for the monitoring of AAAs. It is shown in Section 7.3.3 that when BCG waveforms are used in isolation the accuracy achieved is comparable to that when using a singular bilateral measurement and greater than when using a unilateral measurement. The high accuracies achieved using BCG waveforms and the potential ease by which these measurements can be obtained suggest significant potential in the home monitoring of abdominal aortic health. To further investigate this potential the work presented here must be repeated, however using a more realistic mathematical model of the BCG waveforms than that outlined in

Combination	F_1 score	Sens.	Spec.
(PWV _L)	0.4760	0.4522	0.5523
(Q ₁)	0.9805	0.9799	0.9811
(Q ₁ , PWV _L)	0.9798	0.9813	0.9784
(Q ₂ , Q ₁)	0.9928	0.9919	0.9938
(Q ₂ , Q ₁ , PWV _L)	0.9903	0.9886	0.9922
(Q ₃ , Q ₂ , Q ₁)	0.9962	0.9954	0.9970
(Q ₃ , Q ₂ , Q ₁ , PWV _L)	0.9941	0.9921	0.9963
(Q ₃ , Q ₂ , Q ₁ , P ₂)	0.9972	0.9959	0.9986
(Q ₃ , Q ₂ , Q ₁ , P ₂ , PWV _L)	0.9952	0.9937	0.9969
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂)	0.9970	0.9959	0.9981
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₁)	0.9970	0.9963	0.9978
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , PWV _L)	0.9940	0.9928	0.9954
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₁ , PWV _L)	0.9934	0.9929	0.9940
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , P ₁)	0.9970	0.9959	0.9981
(Q ₃ , Q ₂ , Q ₁ , P ₃ , P ₂ , P ₁ , PWV _L)	0.9945	0.9945	0.9946

Table 7.11: The F_1 scores achieved when using PWV in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for AAA). The corresponding sensitivities and specificities are also included.

Section 7.2.1.

Combination	F_1 score	Sens.	Spec.
(BCG)	0.5442	0.5511	0.5261
(Q_1)	0.8521	0.8547	0.8502
(Q_1 , BCG)	0.8681	0.8696	0.8662
(P_2 , P_1)	0.8950	0.9026	0.8889
(P_2 , P_1 , BCG)	0.8951	0.8878	0.9042
(Q_1 , P_2 , P_1)	0.9389	0.9433	0.9351
(Q_1 , P_2 , P_1 , BCG)	0.9355	0.9314	0.9403
(Q_3 , Q_1 , P_2 , P_1)	0.9395	0.9417	0.9376
(Q_3 , Q_1 , P_2 , P_1 , BCG)	0.9368	0.9431	0.9313
(Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9391	0.9416	0.9370
(Q_2 , Q_1 , P_3 , P_2 , P_1 , BCG)	0.9414	0.9387	0.9446
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9343	0.9364	0.9325
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , BCG)	0.9363	0.9321	0.9411

Table 7.12: The F_1 scores achieved when using BCG in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for CAS). The corresponding sensitivities and specificities are also included.

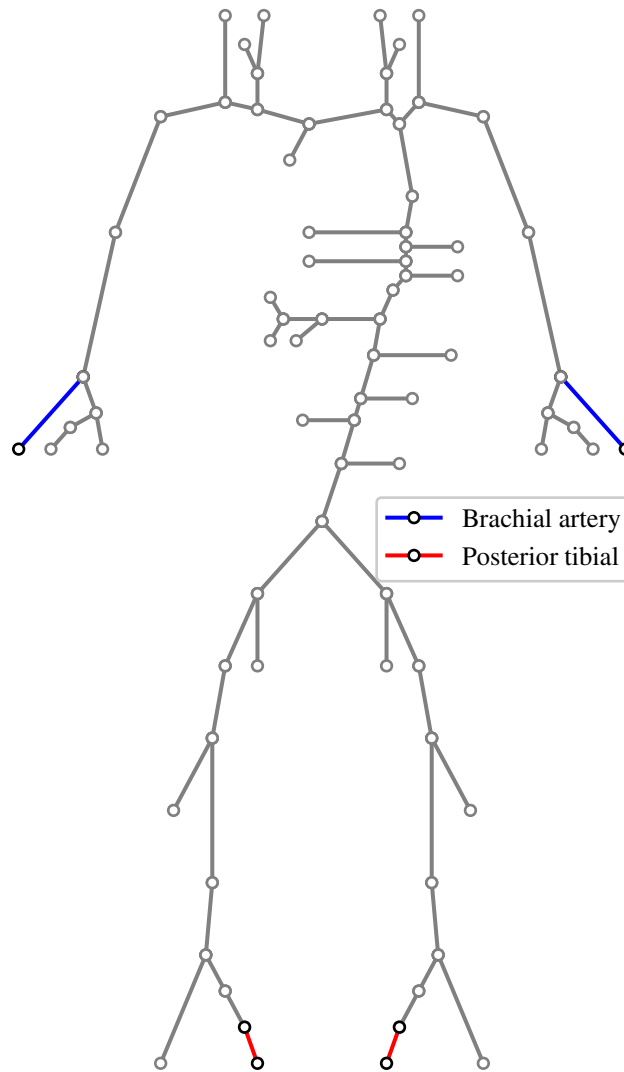


Figure 7.1: The topological positions of the brachial arteries and the posterior tibials.

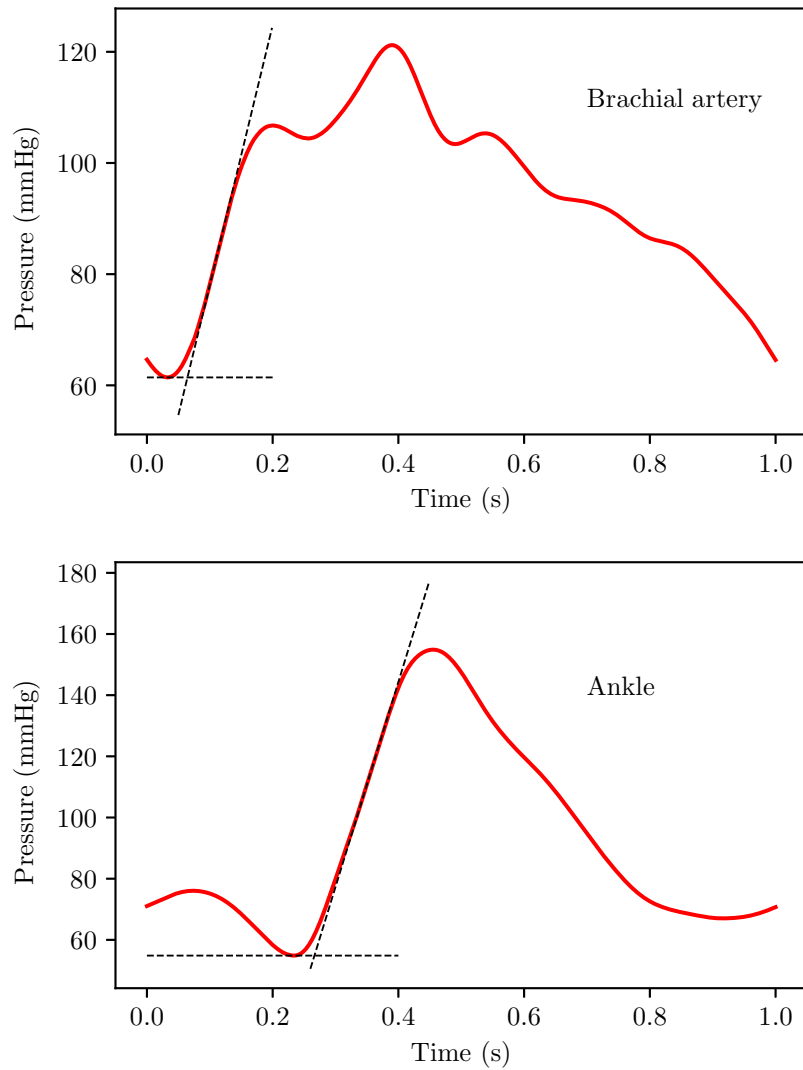


Figure 7.2: An example of brachial artery and ankle pressure profiles taken from a VP for which determination of the footwaves from geometric properties is suitable.

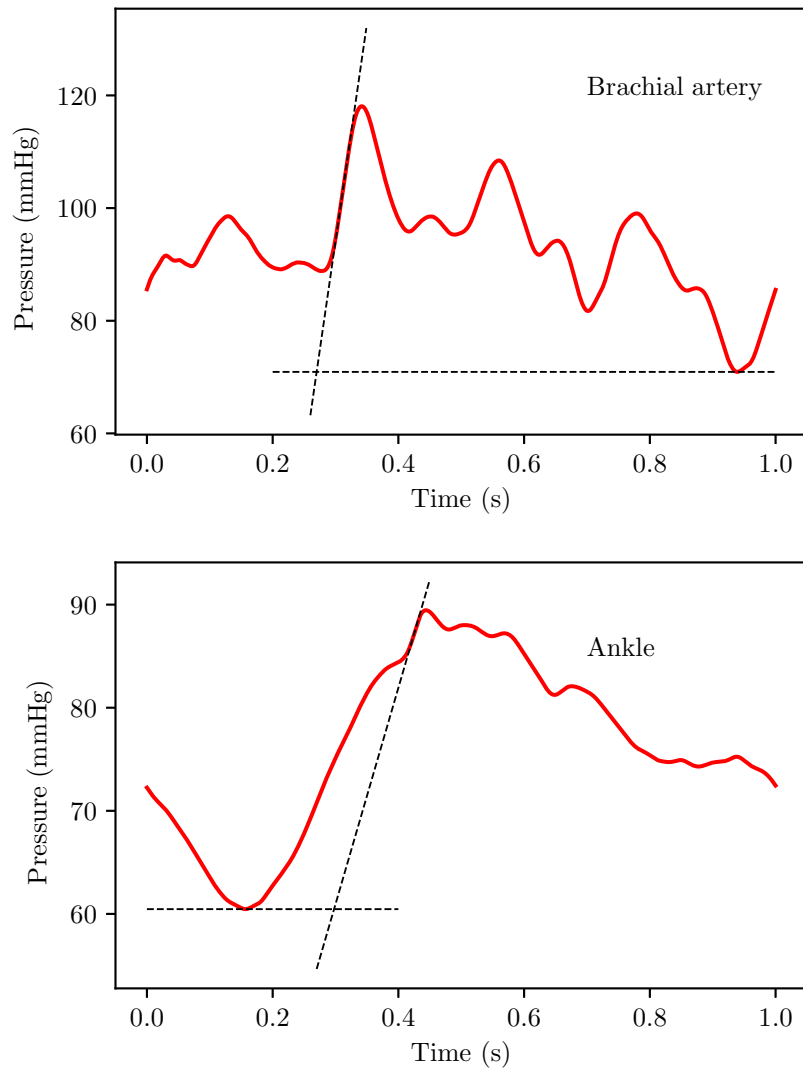


Figure 7.3: An example of brachial artery and ankle pressure profiles taken from a VP for which determination of the footwaves from geometric properties is problematic. Due to the oscillatory pressure profile in the brachial artery, the minimum pressure and maximum pressure gradient occur during different features in the waveform.

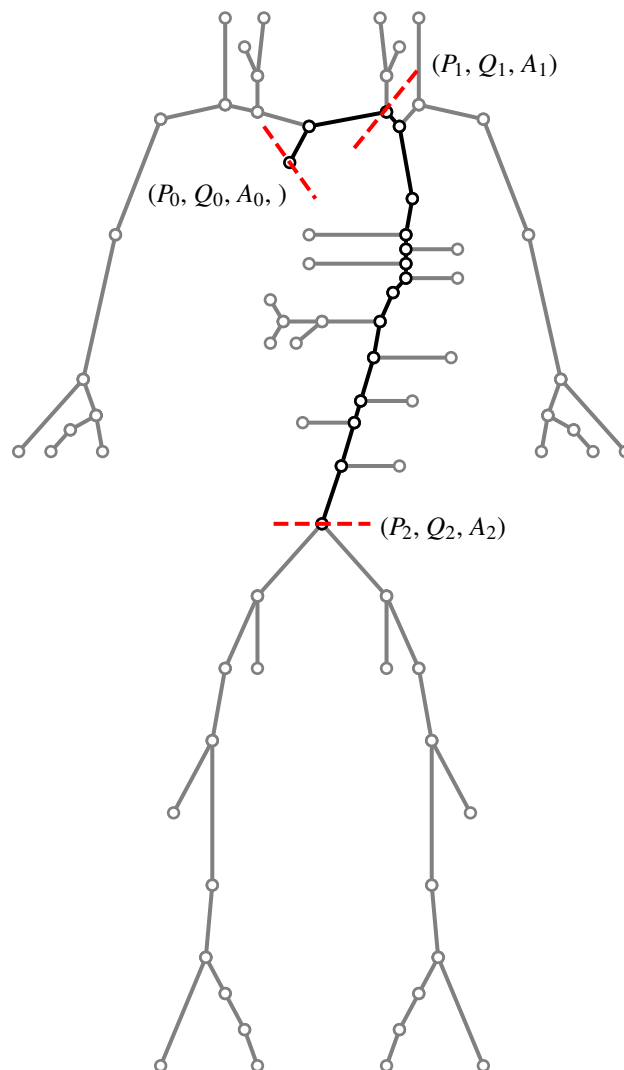


Figure 7.4: The topological positions of the boundaries of the short ascending and long descending aortic segments, used in the mathematical model of the BCG waveform.

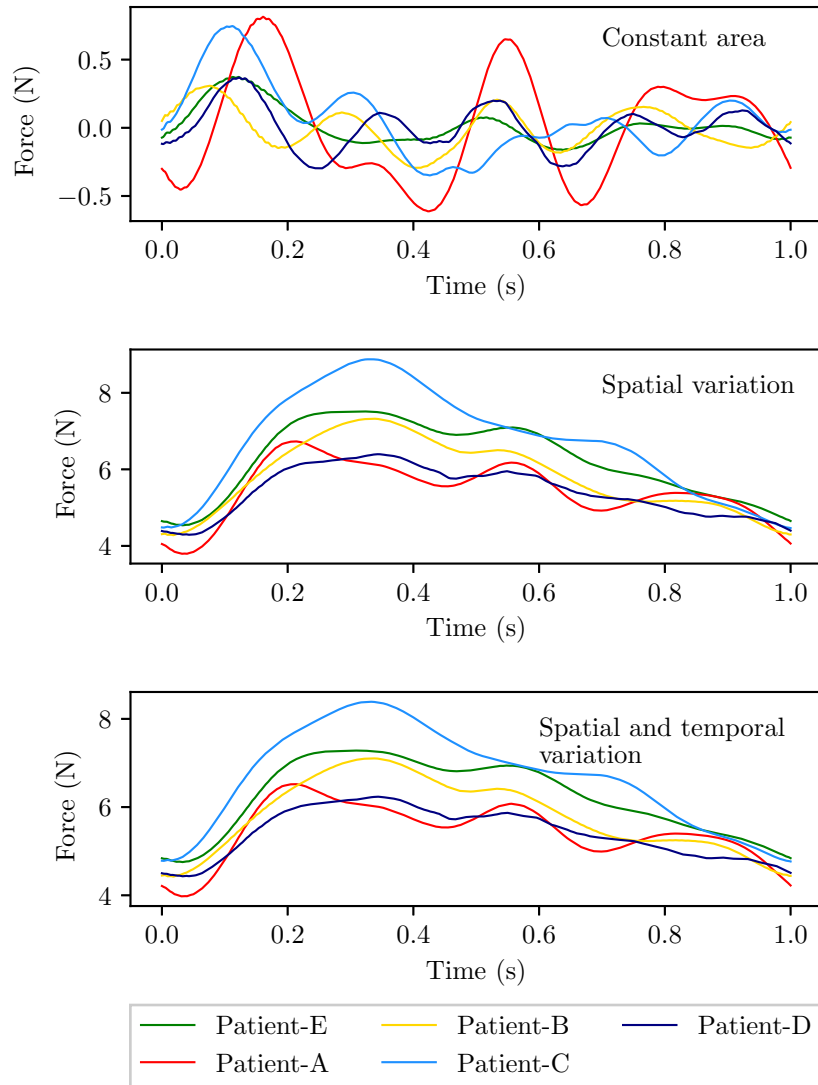


Figure 7.5: Examples of BCG waveforms taken from five VPs when using constant cross sectional areas for the ascending and descending aortas; allowing spatial cross sectional area variation; and allowing spatial and temporal cross sectional area variation. When using constant cross sectional areas the BCG waveform is computed using Equation (7.5), and A_D and A_A are computed as the average of the mean cross sectional area through out the cardiac period at the inlet and outlet of each segment respectively. When spatial cross sectional area variation is allowed the BCG waveform is computed using Equation (7.7) and A_0 , A_1 , A_2 computed as the mean cross sectional area through out the cardiac period at each location.

Combination	F_1 score	Sens.	Spec.
(BCG)	0.5461	0.5609	0.5069
(Q_1)	0.7529	0.7224	0.7714
(Q_1 , BCG)	0.7610	0.7485	0.7814
(Q_2 , Q_1)	0.8461	0.8293	0.8585
(Q_2 , Q_1 , BCG)	0.8470	0.8377	0.8598
(Q_3 , Q_2 , Q_1)	0.8552	0.8453	0.8626
(Q_3 , Q_2 , Q_1 , BCG)	0.8557	0.8486	0.8652
(Q_3 , Q_2 , Q_1 , P_2)	0.8585	0.8487	0.8660
(Q_3 , Q_2 , Q_1 , P_2 , BCG)	0.8738	0.8553	0.8978
(Q_3 , Q_2 , Q_1 , P_2 , P_1)	0.8600	0.8525	0.8657
(Q_3 , Q_2 , Q_1 , P_2 , P_1 , BCG)	0.8624	0.8557	0.8714
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.8574	0.8504	0.8627
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , BCG)	0.8589	0.8550	0.8641

Table 7.13: The F_1 scores achieved when using BCG in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for SAS). The corresponding sensitivities and specificities are also included.

Combination	F_1 score	Sens.	Spec.
(BCG)	0.5582	0.53500	0.6183
(Q_1)	0.8183	0.8126	0.8214
(Q_1 , BCG)	0.8324	0.8310	0.8346
(Q_3 , Q_1)	0.9041	0.8950	0.9117
(Q_3 , Q_1 , BCG)	0.9039	0.8997	0.9090
(Q_3 , Q_2 , Q_1)	0.9168	0.9055	0.9265
(Q_3 , Q_2 , Q_1 , BCG)	0.9156	0.9054	0.9277
(Q_3 , Q_2 , Q_1 , P_1)	0.9196	0.9068	0.9306
(Q_3 , Q_2 , Q_1 , P_1 , BCG)	0.9205	0.9135	0.9289
(Q_3 , Q_2 , Q_1 , P_2 , P_1)	0.9170	0.9041	0.9281
(Q_3 , Q_2 , Q_1 , P_2 , P_1 , BCG)	0.9202	0.9130	0.928
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9187	0.9102	0.9261
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , BCG)	0.9170	0.9095	0.9259

Table 7.14: The F_1 scores achieved when using BCG in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for PAD). The corresponding sensitivities and specificities are also included.

Combination	F_1 score	Sens.	Spec.
(BCG)	0.9742	0.9785	0.9698
(Q_1)	0.9805	0.9799	0.9811
(Q_1 , BCG)	0.9939	0.9943	0.9936
(Q_2 , Q_1)	0.9928	0.9919	0.9938
(Q_2 , Q_1 , BCG)	0.9974	0.9975	0.9973
(Q_3 , Q_2 , Q_1)	0.9962	0.9954	0.9970
(Q_3 , Q_2 , Q_1 , BCG)	0.9984	0.9984	0.9986
(Q_3 , Q_2 , Q_1 , P_2)	0.9972	0.9959	0.9986
(Q_3 , Q_2 , Q_1 , P_2 , BCG)	0.9977	0.9972	0.9984
(Q_3 , Q_2 , Q_1 , P_3 , P_2)	0.9970	0.9959	0.9981
(Q_3 , Q_2 , Q_1 , P_3 , P_1)		0.9963	0.9978
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , BCG)	0.9982	0.9981	0.998
(Q_3 , Q_2 , Q_1 , P_3 , P_1 , BCG)	0.9979	0.9978	0.9982
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1)	0.9970	0.9959	0.9981
(Q_3 , Q_2 , Q_1 , P_3 , P_2 , P_1 , BCG)	0.9980	0.9978	0.9984

Table 7.15: The F_1 scores achieved when using BCG in both isolation and combination with the pressure and flow-rate measurements found to produce the highest F_1 score for each given number of inputs (see Chapter 6, for AAA). The corresponding sensitivities and specificities are also included.

	Side	F_1 score	Sensitivity	Specificity
BCG	-	0.9742	0.9785	0.9698
Carotid flow-rate (Q_1)	Right	0.9272	0.9369	0.9161
	Left	0.9101	0.9065	0.9146
	Both	0.9804	0.9799	0.9811
Radial pressure (P_3)	Right	0.8430	0.8356	0.8533
	Left	0.8620	0.8633	0.8605
	Both	0.9223	0.9202	0.9248

Table 7.16: The F_1 scores, sensitivities, and specificities achieved when using the BCG waveforms; measurements of flow-rate in the right, left, and both carotid arteries; and pressure in the right, left, and both radial arteries.

Chapter 8

Conclusions and contributions

The penultimate chapter of this thesis outlines the contributions and conclusions reached. These contributions and conclusions are split into three components: those relating to the haemodynamic affects of arterial disease, presented in Section 8.1; discussion on the potential for disease detection, presented in Section 8.2; and general tools and frameworks that facilitate future work in Section 8.3.

8.1 Haemodynamic affects of arterial disease

The conclusions reached with regards to the general haemodynamic affects of arterial disease are presented here, with a more specific discussion on the potential for disease detection outlined in Section 8.2. It is shown that the inclusion of a diseased vessel (either stenosis or aneurysm) produces consistent and significant biomarkers in haemodynamic profiles, irrespective of a patients unique underlying arterial network. These biomarkers are found to be differentiable from the natural variability present across a large cohort of patients, showing that arterial disease has a clear and unique affect on pressure and flow-rate profiles.

It is found that the affects of arterial disease on pressure and flow-rate profiles at different locations are dependent on the location of the disease. In Chapter 3 it is shown that classification accuracy increases when the health of each individual vessel is predicted, rather than the health of the entire network. This suggests that the indicative biomarkers of arterial disease differ when disease affects different vessels. This finding is further supported by Chapter 6 which shows that the informativeness—with regards to the presence of arterial disease—of different pressure and flow-rate profiles changes as the disease location changes. An unexpected finding is that the informativeness of different pressure and flow-rate profiles is affected to a greater extent by the topological position at which the profile is measured, than by the spatial distance from the disease location. For example it is shown in Chapter 6 that flow-rate in the carotid artery is highly informative for all forms of disease, despite the large spatial distance from the arterial vessels in which PAD is assumed to affect. It remains unclear why pressure and flow-rate

profiles in certain arterial vessels—particularly the carotid arteries—are so heavily influenced by the presence of arterial disease, irrespective of the form or location of disease, although a hypothesis is presented in Chapter 6.

It is *a priori* expected that the indicative biomarkers of arterial disease consist of micro inter- and intra-measurement details, suggesting the ability to differentiate between healthy and unhealthy patients is reliant on provision of measurements at multiple locations. It is shown in Chapter 6, however, that AAAs can be classified to a high level of accuracy from a single unilateral measurement. This shows that there is sufficient biomarkers of arterial disease captured in intra-profile details to allow for distinction from healthy patients, without the need for inter-profile details. Unilateral measurements have not been tested for the classification of the three forms of stenosis. It is seen in Chapter 6 that the accuracy of stenosis classification is consistently lower than aneurysm classification. Two possible causes for this difference in accuracy are: i) the severity range of aneurysms is disproportional to the severity range of stenosis, causing more severe changes to the pressure and flow-rate profiles; and ii) aneurysm have a more consistent and global affect on pressure and flow-rate profiles. The low levels of degradation seen in the accuracies when classifying low-severity AAAs, relative to high-severity AAAs, suggests that the latter is the more likely cause.

The differentiating partition between the pressure and flow-rate profiles, described through a Fourier series, taken from healthy and diseased patients appears to be non-linear. It is shown in Chapter 3 that SVM classifiers employing an RBF kernel (which allows for non-linear partitions between different classifications) produces significantly higher accuracy classification than both SVM classifiers employing a linear kernel and LR classifiers (both of which do not allow for non-linear partitions). This finding is supported by Chapter 6 which shows that the LR method, the only linear method used in this chapter, consistently produces the lowest accuracy.

8.2 Potential for arterial disease detection through easily acquirable peripheral measurements

This thesis has shown strong potential in the possibility of arterial disease detection from easily acquirable peripheral haemodynamic measurements. It is shown in Chapter 6 that when using six measurements of pressure and flow-rate, maximum F_1 scores larger than 0.9 are achieved for CAS and PAD, larger than 0.85 for SAS, and larger than 0.98 for both low- and high-severity AAAs. Corresponding sensitivities and specificities are larger than 90% for CAS and PAD, larger than 85% for SAS, and larger than 98% for both low- and high-severity AAAs. When reducing the number of measurements, it is found that the performance is degraded by less than 5% when three measurements are used, and less than 10% when only two measurements are used for classification.

The current potential of large scale screening for stenosis through haemodynamic measurements is limited by the fact that the accuracy of classification is degraded by approximately 25% when a single bilateral measurement is used. No analysis has been carried out on the accuracy of stenosis classification when using unilateral measurements. The results achieved in this thesis are likely, however, to be easily improved by performing classifier specific optimisation (discussed in Chapter 9). If the results achieved for stenosis classification when using a single peripheral measurement of pressure or flow-rate is improved the viability of a stenosis screening program, similar to the AAA screening program currently offered by the United Kingdom's National Health Service (NHS), would be significantly improved.

The greatest potential shown in this thesis is for AAA classification. Three viable methods for AAA classification are shown in this thesis:

- Through the use of unilateral carotid artery flow-rate measurements. In Chapter 6 AAAs are classified using unilateral measurements of right and left carotid flow-rate with sensitivities and specificities of 93.69% and 91.61%, respectively; and 90.65% and 91.46%, respectively. Carotid artery flow-rate is easily acquirable in a clinical environment (see Chapter 4 for details), and so this method could possess the potential to be used to expand the AAA

screening program currently offered by the NHS.

- Through the use of unilateral radial pressure measurements. In Chapter 6 classification using unilateral measurements of right and left radial pressure achieves sensitivities and specificities of 83.56% and 85.33%, respectively; and 86.33% and 86.05%, respectively. The use of radial pressure holds potential due to its location in the human wrist. Through wearable devices that can measure continuous radial pressure profiles, such as the TLT Sapphire monitor (Tarilian Laser Technologies, Welwyn Garden City, U.K.) [125], this method holds the potential for constant home monitoring of abdominal aortic health.
- Through the use of BCG waveforms. It is shown in Chapter 7 that AAAs can be classified using BCG waveforms with sensitivities and specificities of 97.85% and 96.98%, respectively. As with unilateral radial pressure measurements, this method holds the potential for constant home monitoring of abdominal aortic health. The potential of BCG waveforms is increased from that of radial pressure measurements as multiple methods already exist that allow for acquisition of the BCG waveform from home (see Chapter 7 for details).

8.3 Tools and frameworks

A discussion of the general tools and frameworks outlined in this thesis, that can be used to facilitate future work is presented in this section.

8.3.1 Creation of a physiologically realistic virtual patient database, for the study of arterial haemodynamic

A major contribution of this thesis is the creation of the physiologically realistic virtual patient database (VPD) in Chapter 4. This VPD contains 28,868 physiologically realistic virtual patients (VPs), each containing 71 arterial vessel segments. The pressure, flow-rate, and area profiles associated with each VP are included. This VPD is made publicly available (<https://doi.org/10.5281/zenodo.4549764>). This database can now be used by others who intend to study arterial haemodynamic over large cohorts (for example statistical and ML studies).

8.3.2 Methodology for the creation of physiologically realistic virtual patients

Aside from the VPD, a further contribution of Chapter 4 is the detailing of a generic methodology for the creation of physiologically realistic VPs. In Chapter 4 this methodology is broken down into the following steps: i) network reduction without compromising relevant behaviour; ii) re-parameterisation to reduce dimensionality; iii) incorporation of geometrical and physiological constraints in the form of a prior; iv) incorporation of literature reported clinical measurements in the form of the likelihood; v) combination of the prior and likelihood to generate the posterior; and vi) sampling from the posterior with MCMC to create the VPD. Examples of the incorporation of both scalar and time varying (through the use of a Gaussian random process) literature reported clinical measurements are presented in Chapter 4. With minor modifications this methodology can now be applied to a range of biological systems, given a mathematical description. Possible adaptations to this generic methodology include: the incorporation of expected physiological inter-parameter relationships into the re-parameterisation, to further reduce dimensionality and impose stricter control over the realisations of VPs sampled; the imposition of a symmetry metric, for example introducing correlations into the prior distribution; and the creation of higher dimensional literature based pseudo measurements (through high dimensional Gaussian random fields). VPs created through the aforementioned methodology can be used for in-silico trials, along with data-mining. The potential of the VP approach for data-mining is highlighted by Section 8.1. A further advantage of the VP approach is that it can be used to *a priori* examine ML effectiveness, to access potential, and guide what measurements should be acquired to tackle clinical problems, such as diagnosis. This benefit is particularly valuable for physiological problems, as obtaining measurements from a real cohort is often invasive and expensive.

8.3.3 Primer for the application of machine learning classifiers

In Chapters 2, 3, 6, and 7 a range of methodologies, considerations, and behaviours are detailed with regards to the application of machine learning classifiers. These

include:

- A description of six different machine learning classification methods, that encompass a range of probabilistic and non-probabilistic applications of different modelling approaches (see Chapter 2 for details).
- The use of events per variable (EPV) to produce an *a priori* estimation of the required quantity of data to accurately train and test classifiers (Chapter 3).
- A method to *a posteriori* assess if the quantity of data is sufficient (Chapter 3).
- Two different methods—sensitivity and specificity in combination; and the F_1 score—to quantify and compare the results achieved by different classifiers. The relevance and limitations of these are also discussed (Chapter 3).
- When and why bias may be introduced due to class imbalance (Chapter 3), and the subset of data on which the accuracy of classification is assessed (Chapters 3, 6, and 7).
- The application of a grid search to optimise the hyper-parameters describing the architecture of a range of different classifiers to a particular problem (Chapters 3, 6, and 7).

All of the above information is presented in a relatively non-intensive intuitive manner, to facilitate their use by those who have mathematical and computational knowledge, however lack experience in the application of machine learning classifiers—such as computational mechanics. This work can, therefore, be thought of as a primer presenting relatively simple however robust and widely applicable methods that can be employed by non-specialists carrying out initial investigations into the application of ML classifiers.

Chapter 9

Future work

The final chapter of this thesis outlines possible future work arising from the results achieved and conclusions reached. The future work suggested in this chapter is broken down into three categories: improvements to the classification accuracies achieved in this thesis, presented in Section 9.1; expansions to the current analysis to build a better understanding of both the potential for arterial disease detection and the haemodynamic affects of arterial disease, presented in Section 9.2; and work that can be completed to validate the results and conclusions reached in this thesis, presented in Section 9.3.

9.1 Improving classification accuracies

Throughout this thesis an exploratory stance has been taken to the training and testing of ML classifiers, with focus being on uncovering patterns and behaviours, rather than optimisation for increased accuracy. This has allowed for a wide range of classifiers to be created using different combinations of haemodynamic measurements—in the input combination search presented in Chapter 6 ML classifiers are created for four forms of disease using 63 combinations of pressure and flow-rate measurements over five folds of the data sets, thus totalling $1,260$ ($4 \times 63 \times 5$) ML classifiers trained and tested. This width, however, has come at the cost of the “depth” of each individual classifier. Key classifiers can now be identified—either due to the patterns and behaviours they show or the potential they possess—and optimisation performed to improve the accuracies presented in this thesis. This optimisation may involve the use of more complex classification methods, measurement specific hyper-parameter optimisation, feature selection (*i.e.* identifying the most informative FS coefficients), and changing the representation of the pressure and flow-rate profiles used. Classifiers which warrant further optimisation, and the reasons why each are of interest, are:

- Carotid artery stenosis (CAS) classification when using only the measurement of common carotid artery flow-rate (Q_1). It is seen in Chapter 6 that a

sensitivity and specificity of 87.04% and 88.93%, respectively, are achieved for CAS classification when using only the measurement Q_1 . This shows that there is sufficient information captured in the inter-profile features of carotid flow-rate to allow for high accuracy classification. Optimisation of this classifier could be performed to find the maximum accuracy to which CAS can be classified using a single measurement.

- Subclavian artery stenosis (SAS) classification when using only the measurement Q_1 . As with CAS classification, reasonable accuracies are achieved in Chapter 6 by this classifier (sensitivity and specificity of 75.82% and 79.05%, respectively), and so further work could be carried out to find the maximum accuracy achievable.
- Peripheral arterial disease (PAD) classification when using only the measurement Q_1 . Sensitivity and specificity of 89.59% and 83.20%, respectively, are achieved by this classifier in Chapter 6. As with CAS and SAS, further optimisation could be performed to find the maximum accuracy achievable.
- Abdominal aortic aneurysm (AAA) classification using unilateral measurements of carotid flow-rate. When AAAs are classified using unilateral measurements of right and left carotid flow-rate sensitivities and specificities of 93.69% and 91.61%; and 90.65% and 91.46% are achieved, respectively. This compares to a sensitivity and specificity of 97.99% and 98.11%, respectively, when both the right and left carotid artery flow-rates are used bilaterally. This suggests that i) there is sufficient information in unilateral carotid flow-rate measurements to classify AAA to a high accuracy, and ii) near perfect accuracy can be achieved using a single unilateral measurements if only marginally more information is available.
- Similar analysis to that presented above could be completed for AAA classification using unilateral measurements of radial pressure. Current accuracies (achieved in Chapter 6) are sensitivities and specificities of 83.56% and 85.33% when using the right radial pressure; 86.33% and 86.05% when using the left radial pressure; and 92.02% and 92.48% when using both the right and left radial pressure, respectively. The use of unilateral radial pressure

measurements holds the additional benefit of the location of the radial artery in the wrist (see Chapter 8).

9.2 Expanding the current analysis

Future work can be performed to gain a deeper understanding of both the potential of different classifiers, and the physical cause of certain behaviours. Patterns and behaviours that warrant further investigation, that can be completed to build further insight into the haemodynamic affects of arterial disease, include:

- Why the ability to create non-linear partitions between classifications appears to be important in the binary classification analysis presented in Chapter 3, however not in the multiclass analysis.
- Why flow-rate in the carotid artery (Q_1) is so informative. This measurement is seen in Chapter 6 to be highly informative for all forms of disease classification, including PAD, irrespective of the distance between the disease and measurement locations. In Chapter 6 it is hypothesised that carotid flow-rate is informative as it acts as a pre-disease reference profile, which when combined with a post-disease measurement bound the disease location to provide more information. This hypothesised could be further investigated.

To further explore the potential of using machine learning classifiers to predict the presence of arterial disease, and expand the analysis carried out in this thesis, the healthy VPD created in Chapter 4 can be used to complete the follow analysis:

- In Chapter 6 ML classifiers are trained to first predict the presence of AAAs with a severity range of 713%–2,593% increase in cross sectional area, and then 300%–700% increase in cross sectional area. High accuracy classification is achieved for both high- and low-severity aneurysm classification individually (see Chapter 6 for details). To expand on these results a new AAA VPD could be created, containing VPs with a severity range of 300%–2,593% (*i.e.* the full range of both low- and high-severities), and classifiers trained and tested using this new VPD. It is likely that classifiers trained to predict both low- and high-severity AAAs in combination will produce lower accuracies than either severity ranges in isolation, as there is likely to be greater variability

in the pressure and flow-rate profiles across the data set. Future work could investigate the degree to which accuracy is degraded when classifying both low- and high-severity AAAs in combination.

- In Chapter 5 it is chosen to create VPs containing only one form of disease (*i.e.* each unhealthy VP has either CAS, SAS, PAD, or AAA). Chapter 6 then furthers this assumption of mutual exclusivity by creating classifiers for each form of disease independently. ML classifiers could now be created to predict the presence of each form of arterial disease when more than one form of disease may be present. This further analysis can be completed in two stages:
 1. The previously created unhealthy VPDs (each containing only one form of disease) can be used to create mixed disease data sets, *i.e.* each VP has only one form of disease however the data set contains multiple forms of disease. Binary ML classifiers can then be created to predict if a VP is subject to a particular form of disease, and multiclass classifiers to determine which form of disease a VP has.
 2. New VPDs can be created, in which each VP may contain more than one form of disease. In this case binary classifiers can be created to predict the presence of each individual form of disease within a VP, and multiclass classifiers to predict the combination of disease forms present.
- While it has been shown in Chapters 6 and 7 that both stenosis and aneurysm can be classified to high levels of accuracy using haemodynamic measurements, no analysis has been carried out on the ability to predict the severity of disease. The potential to not only determine the presence, but also estimate the severity of disease can now be assessed. Initially, predicting the severity of disease can be treated as a multiclass classification problem, with ML classifiers being created to predict if a VP has: no disease present, mild disease present, moderate disease present, or severe disease present. If this is successful ML regression algorithms can be used to further this analysis, and provide a prediction of the change in cross sectional area. This further analysis is particularly fitting for AAA classification using unilateral radial pressure measurements and BCG waveforms (see Chapter 7). Both unilateral radial pressure measurements and BCG waveforms offer the

possibility of home monitoring of abdominal aortic health, and thus showing these measurements can be used to predict the severity of disease suggest the potential for continuous monitoring of disease progression. If it is shown that the severity of AAAs can be continuously monitored through radial pressure measurements or BCG waveforms, it may have significant implications on current clinical practices. Rather than assessing the relative risk posed to a patient based on a single “snapshot” in time, the aforementioned method allows for automatic continuous assessment, independent of human input, based on severity and growth rate. Patients deemed to be at high risk can then be further assessed through traditional clinical methods, such as angiography and Doppler ultrasound.

9.3 Validating results and conclusions

It is shown in Chapter 7 that BCG waveforms have the potential to be used to predict the presence of AAAs to a high level of accuracy from a single measurement. The mathematical model used to compute the BCG waveforms associated with VPs, however, is relatively simplistic and does not account for: momentum changes within the control volume; the curvature of the aorta; and other large arterial vessels such as the axillary, carotid, and iliac arteries. Before further work is carried out on the potential of BCG waveforms, such as the aforementioned regression problem, a more physiologically realistic model should be developed that accounts for and rectifies the simplifications made in Chapter 7. The results achieved when using BCG waveforms in Chapter 7 can then be validated against those achieved when employing a more detailed model.

All ML classifiers created in this thesis have been trained and tested using synthetic data sets. While the VPD created in Chapter 4—used in Chapters 5, 6, and 7—is designed to be as physiologically realistic as possible (incorporating literature reported measurements of pressure and flow-rate; and known physiological restrictions and geometrical constraints), it is likely that the use of this synthetic data set will have an affect on the accuracy of classification, relative to using data obtained from a real cohort. The resulting effect on the classification accuracies achieved, *i.e.* an increase or decrease in accuracy, is unclear and difficult to predict.

9.3. VALIDATING RESULTS AND CONCLUSIONS

To understand both the effect of using synthetic data sets and the “true” accuracy of disease classification, data measured from a real cohort is required. The ML classifiers trained on the synthetic data sets in this thesis can be used to predict the health of real patients, for which the true state classification is already known. The true accuracy of classification (achieved on the data set taken from a real cohort) can then be compared to that achieved on the synthetic data set.

Appendix A

Evaluation of the distribution of pressure and flow-rate profiles across the VPD

During the creation of the VPD, *a priori* distributions are assigned to the parameters describing the arterial networks of VPs. While a post simulation filter has been applied to VPs, to constrain the range of pressure profiles at the inlet of the system, no active consideration has been given to the the distribution of pressure profiles within these ranges. It is instead assumed that due to the cause and effect relationship between arterial networks and the corresponding pressure and flow-rate profiles, that the physiological realism of the distributions of pressure and flow-rate profiles is proportional to the physiological realism of the distributions of arterial network parameters. It is therefore important to examine the *a posteriori* distribution of the pressure and flow-rate profiles across the VPD. The distribution of pressure and flow-rate profiles across the VPD are empirically found at the inlet and two outlets of the arterial system, and are shown in Figure A.1.

Figure A.1 shows that the standard deviation of the inlet flow-rate distribution is significantly lower as a proportion of the mean value than any other pressure or flow-rate profile. This is likely due to the fact that the inlet flow-rate profile prescribed to VPs is solely produced by the FS coefficients describing the profile. This means that the variability in inlet flow-rate profiles seen across the VPD is solely caused by the the distribution of the FS coefficients. The pressure and flow-rate profiles at all other locations within the network are caused by inter-parameters relationships, and so the variability in pressure and flow-rate profiles at all other locations is caused by complex inter-distribution relationships.

Generally the standard deviation of pressure profiles seems to be more consistent than that of flow-rate profiles. It is seen from the lower two left subplots within Figure A.1 that there is a constriction in the standard deviation of the distribution of flow-rate profiles at the two outlets at approximately 0.75 seconds. This temporal position corresponds to the peak systolic pressure.

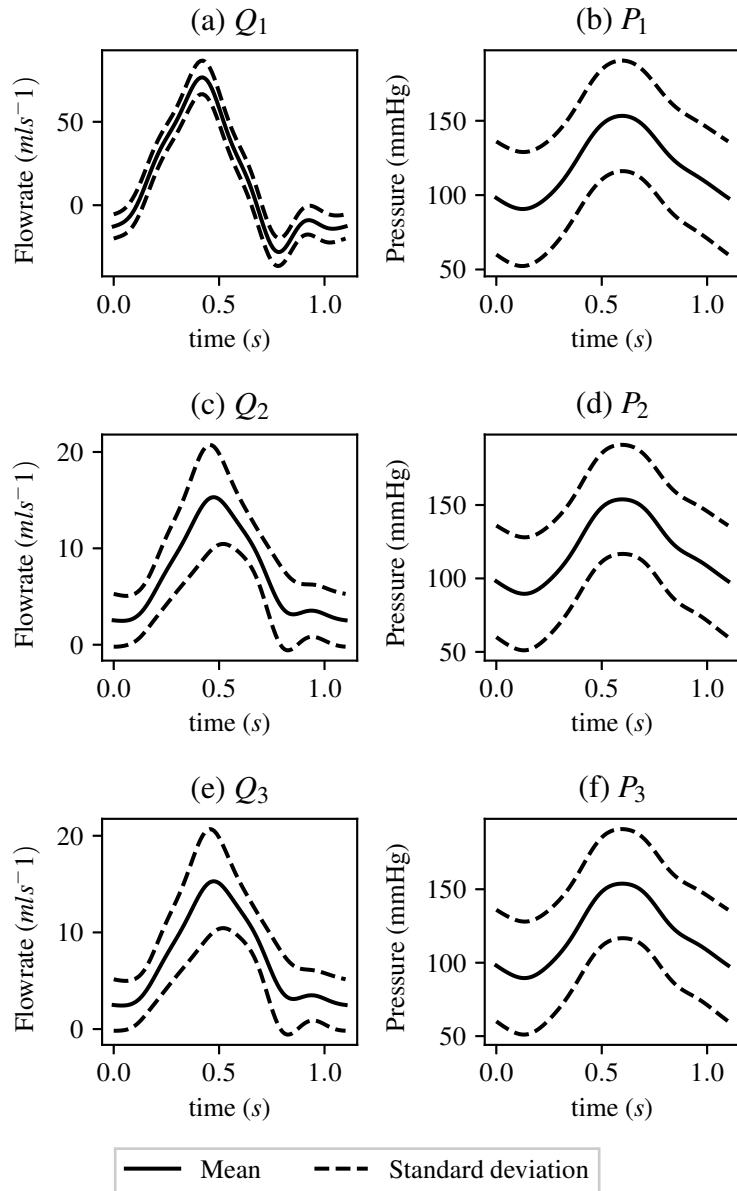


Figure A.1: The mean and standard deviation envelope, *i.e.* the mean \pm one standard deviation, are shown at all measurement obtainment locations. The subplots show the flow-rate at the inlet of the system (a), pressure at the inlet of the system (b), flow-rate at the outlet of the first iliac (c), pressure at the outlet of the first iliac (d), flow-rate at the outlet of the second iliac (e), and pressure at the outlet of the second iliac (f).

As no consideration has been given to the likely distribution of pressure and flow-rate profiles across the resulting VPD, it is hard to pass *a posteriori* judgement on the resulting empirical distributions. Generally, however, the magnitude of the standard deviations of pressure and flow-rate profiles across the VPD is reasonable. A very large standard deviation, as a proportion of the mean value, will likely make it difficult for ML classifiers to distinguish between healthy and unhealthy VPs, as there will be a large amount of naturally variability within the VPD. On the other hand, overly constricted standard deviations will result in the variability of pressure and flow-rate profiles not being representative of a that measured in a real population. Looking at Figure A.1, the distribution of pressure and flow-rate profiles seems to reasonably balance these two opposing behaviours.

Appendix B

Fourier representation

There are two possible forms of a FS. A FS can be described using either a sine-cosine form:

$$\mathbf{q}(x) = \sum_{n=0}^N a_{(n,x)} \sin(n\omega t) + b_{(n,x)} \cos(n\omega t), \quad (\text{B.1})$$

or using a phase-amplitude form:

$$\mathbf{q}(x) = \sum_{n=0}^N \mathcal{Z}_{(n,x)} \cos(n\omega t + \phi_{(n,x)}) \quad \text{for}, \quad (\text{B.2})$$

where:

$$\omega = \frac{2\pi}{T}. \quad (\text{B.3})$$

and $a_{(n,x)}$, $b_{(n,x)}$, $\mathcal{Z}_{(n,x)}$, and $\phi_{(n,x)}$ represents the n^{th} sine, cosine, amplitude, and phase FS coefficients at the spatial position x respectively. Within the above equations $\mathbf{q}(x)$ represents the vector describing the pressure or flow-rate profile at spatial location x , n represents the FS coefficient order, N represent the order at which the FS is being truncated, t represents the temporal position, and T represents the cardiac period.

While these two forms of a FS require the same dimensionality to represent a pressure or flow-rate profile, they do not necessarily capture the same information about the haemodynamic profiles as one another. ML classifiers trained and tested using pressure and flow-rate profiles represented through each of the two forms of a FS will, therefore, exhibit different behaviours and accuracies. To compare the use of these two forms of a FS as input measurements into ML classifiers, a series of LR classifiers are created using each. The first comparison carried out trains two ENBCs, using the pressure and flow-rate measurements at all three locations. The input measurements into the first and second classifier trained are the sine-cosine coefficients and the phase-amplitude coefficients of the FS describing each pressure or flow-rate profile respectively.

For this comparison the custom implementation of a LR method is used as, unlike Scikit-learn, it allows the cost of the predicted probabilities (as described by

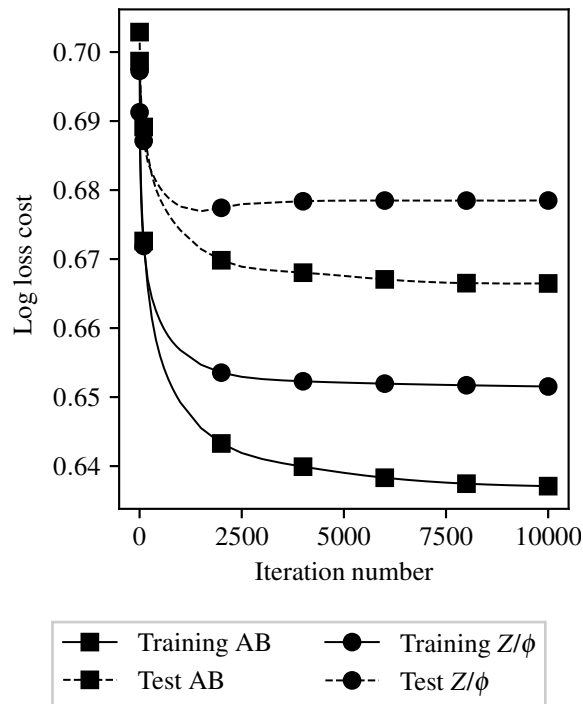


Figure B.1: Comparison of the use of the A/B and Z/ϕ forms of a FS to describe the pressure and flow-rate profiles used as inputs into ENBC classifiers.

Equation (3.20)) to be recorded at every sequential iteration within the training process across the training and test sets. This allows for not only a comparison of the end accuracy, but also an evaluate of the differences in behaviours of training when using each of the two FS forms.

The highest accuracy results achieved by the custom LR implementation with Section 3.4.7 are obtained when using a tanh activation function and the Adam update function. Based on this finding, both instance of a LR algorithm within this test are trained for ten thousand iterations of Adam, using a tanh activation function. The average logloss cost across the training and test set are recorded at regular iterations during the training of each LR algorithms. The average logloss cost profiles, over the five folds of the VPD, are shown in Figure B.1.

Figure B.1 shows that the A/B FS coefficients produce lower costs across both the training and test sets. The cost profiles show no obvious signs of over-fitting. While there is a relatively large discrepancy between training and test costs, the test

cost profile of both representations can be seen to smoothly reduce, before reaching a plateau. This is a desired profile.

To further compare the use of A/B and \mathcal{Z}/ϕ forms of a FS as inputs into ML classifiers three IVBCs are created using each representation of a FS. The three IVBCs created using each representation predict the health of the aorta, the first iliac, and the second iliac respectively. As in the case of the ENBCs, these classifiers are trained and tested using each representation of the FS describing all six measurements. The training and test profiles of the three IVBCs trained using each representation are shown in Figure B.2.

Figure B.2 shows that all three IVBC produce very consistent training and test cost profiles when using each of the two forms of a FS. When using either of the two forms of a FS for all three IVBCs there is a continuous smooth decrease in the training and test costs as the number of sequential training iterations increases. In the case of aortic classification, shown within the upper subplot of Figure B.2, it is seen that both the A/B and \mathcal{Z}/ϕ forms of a FS quickly reach an asymptotic minimum training and test cost. Neither form of a FS appear to have reached their asymptotic minimum training and test cost in the case of the first and second common iliac classifiers, shown in the middle and bottom subplot of Figure B.2. For all three binary classifiers, it is seen that after a very low number of sequential training iterations the A/B form of a FS is producing lower training and test costs than the \mathcal{Z}/ϕ form. The lower costs produced by the A/B form of a FS are maintained throughout the training process.

From the comparison of the two forms of a FS outlined above, it can be seen that using an A/B form of the FS describing the pressure and flow-rate profiles taken from VPs produces lower training and test costs. An A/B form of the FS describing the pressure and flow-rate profiles taken from VPs is used as the input measurements into all future ML classifiers created.

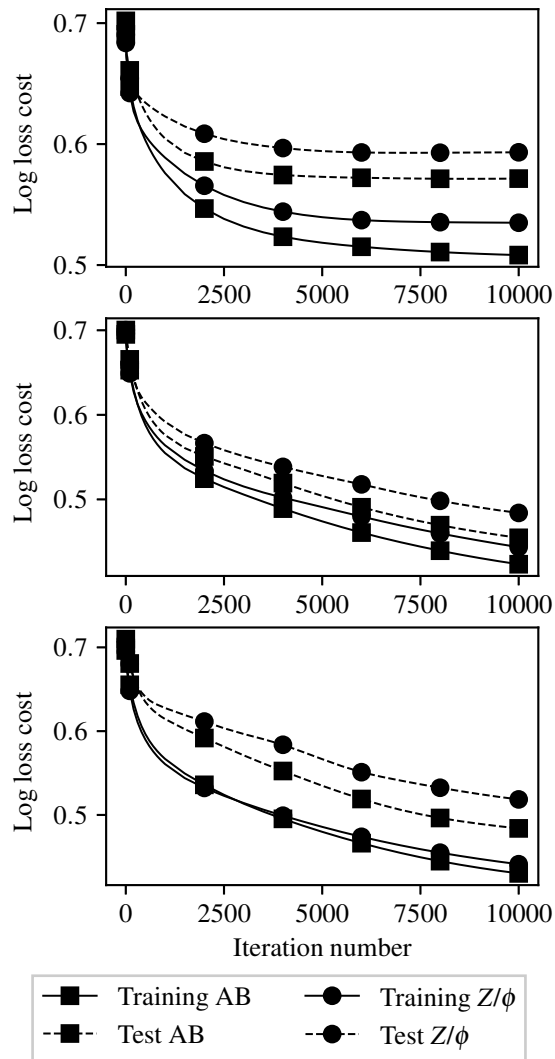


Figure B.2: Comparison of the use of the A/B and Z/ϕ forms of a FS to describe the pressure and flow-rate profiles used as inputs into IVBC classifiers. Comparison is made on the classifiers trained to predict the health of the aorta within the top subplot, the first iliac within the middle subplot, and second iliac within the bottom subplot.

Appendix C

ENBC combination search results

The F_1 scores, sensitivities, and specificities achieved across the ENBC combination search are shown in Tables C.1, C.2, and C.3, respectively.

Input combination	Classification method			
	NB	LR	SVM	RF
Q_3	0.3494	0.4915	0.5138	0.4789
Q_2	0.5318	0.4824	0.4989	0.4965
Q_1	0.3008	0.4932	0.5621	0.4540
P_3	0.4328	0.4908	0.5292	0.4926
P_2	0.4413	0.5060	0.5287	0.5122
P_1	0.3059	0.4924	0.5307	0.4705
Q_3, Q_2	0.4930	0.4878	0.5510	0.4852
Q_3, Q_1	0.3126	0.5136	0.6015	0.4756
Q_3, P_3	0.4244	0.4989	0.5710	0.5053
Q_3, P_2	0.4342	0.5032	0.5757	0.5109
Q_3, P_1	0.3208	0.5077	0.5801	0.4910
Q_2, Q_1	0.4228	0.4962	0.5892	0.4916
Q_2, P_3	0.4916	0.5057	0.5559	0.5080
Q_2, P_2	0.4934	0.5081	0.5479	0.5046
Q_2, P_1	0.3997	0.5054	0.5570	0.4861
Q_1, P_3	0.3698	0.5163	0.6050	0.4956
Q_1, P_2	0.3806	0.5316	0.6121	0.5086
Q_1, P_1	0.3140	0.5190	0.6152	0.4776
P_3, P_2	0.4378	0.5052	0.5391	0.5200
P_3, P_1	0.3668	0.5267	0.5617	0.5065
P_2, P_1	0.3729	0.5397	0.5620	0.5106
Q_3, Q_2, Q_1	0.4181	0.5091	0.6098	0.4901
Q_3, Q_2, P_3	0.4739	0.5079	0.5883	0.5080
Q_3, Q_2, P_2	0.4778	0.5092	0.5824	0.5090
Q_3, Q_2, P_1	0.3957	0.5104	0.5918	0.4945
Q_3, Q_1, P_3	0.3728	0.5292	0.6240	0.4957
Q_3, Q_1, P_2	0.3840	0.5292	0.6279	0.5138
Q_3, Q_1, P_1	0.3205	0.5290	0.6356	0.4909
Q_3, P_3, P_2	0.4360	0.5041	0.5769	0.5085
Q_3, P_3, P_1	0.3702	0.5420	0.5983	0.5049
Q_3, P_2, P_1	0.3770	0.5435	0.5985	0.5160
Q_2, Q_1, P_3	0.4444	0.5223	0.6117	0.5036
Q_2, Q_1, P_2	0.4437	0.5348	0.6105	0.5013
Q_2, Q_1, P_1	0.3780	0.5262	0.6182	0.4901
Q_2, P_3, P_2	0.4741	0.5090	0.5629	0.5179
Q_2, P_3, P_1	0.4119	0.5378	0.5769	0.5103
Q_2, P_2, P_1	0.4165	0.5478	0.5761	0.5153
Q_1, P_3, P_2	0.4121	0.5344	0.6143	0.5196
Q_1, P_3, P_1	0.3433	0.5470	0.6221	0.4948
Q_1, P_2, P_1	0.3507	0.5549	0.6228	0.5146

Input combination	Classification method			
	NB	LR	SVM	RF
P_3, P_2, P_1	0.3938	0.5518	0.5740	0.5239
Q_3, Q_2, Q_1, P_3	0.4376	0.5292	0.6280	0.5023
Q_3, Q_2, Q_1, P_2	0.4395	0.5384	0.6273	0.5060
Q_3, Q_2, Q_1, P_1	0.3797	0.5350	0.6368	0.4947
Q_3, Q_2, P_3, P_2	0.4696	0.5098	0.5929	0.5183
Q_3, Q_2, P_3, P_1	0.4105	0.5466	0.6100	0.5119
Q_3, Q_2, P_2, P_1	0.4144	0.5482	0.6078	0.5175
Q_3, Q_1, P_3, P_2	0.4104	0.5364	0.6240	0.5079
Q_3, Q_1, P_3, P_1	0.3516	0.5588	0.6429	0.5043
Q_3, Q_1, P_2, P_1	0.3540	0.5587	0.6407	0.5103
Q_3, P_3, P_2, P_1	0.3956	0.5596	0.6025	0.5163
Q_2, Q_1, P_3, P_2	0.4511	0.5387	0.6153	0.5093
Q_2, Q_1, P_3, P_1	0.3956	0.5538	0.6268	0.5053
Q_2, Q_1, P_2, P_1	0.3982	0.5676	0.6274	0.5170
Q_2, P_3, P_2, P_1	0.4238	0.5597	0.5836	0.5229
Q_1, P_3, P_2, P_1	0.3773	0.5698	0.6277	0.5180
Q_3, Q_2, Q_1, P_3, P_2	0.4455	0.5397	0.6275	0.5161
Q_3, Q_2, Q_1, P_3, P_1	0.3955	0.5670	0.6469	0.5129
Q_3, Q_2, Q_1, P_2, P_1	0.4000	0.5686	0.6432	0.5139
Q_3, Q_2, P_3, P_2, P_1	0.4251	0.5595	0.6140	0.5127
Q_3, Q_1, P_3, P_2, P_1	0.3791	0.5695	0.6434	0.5191
Q_2, Q_1, P_3, P_2, P_1	0.4100	0.5718	0.6306	0.5188
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4139	0.5815	0.6454	0.5246

Table C.1: The F_1 scores achieved across the combination search by each of the four classification methods.

Input combination	Classification method			
	NB	LR	SVM	RF
Q_3	0.7431	0.5516	0.6868	0.5961
Q_2	0.4624	0.5896	0.6932	0.5669
Q_1	0.8321	0.5348	0.7154	0.5956
P_3	0.6755	0.5833	0.7289	0.5654
P_2	0.6732	0.6038	0.7445	0.5681
P_1	0.8094	0.6309	0.7634	0.6168
Q_3, Q_2	0.5447	0.5686	0.7186	0.59611
Q_3, Q_1	0.8127	0.5355	0.7220	0.6413
Q_3, P_3	0.6817	0.5738	0.7144	0.5741
Q_3, P_2	0.6710	0.5928	0.7183	0.5704
Q_3, P_1	0.7803	0.6239	0.7603	0.6121
Q_2, Q_1	0.6912	0.5684	0.7387	0.6150
Q_2, P_3	0.5907	0.5840	0.7311	0.5791
Q_2, P_2	0.5941	0.5879	0.7466	0.5812
Q_2, P_1	0.7303	0.6213	0.8000	0.6344
Q_1, P_3	0.7664	0.5754	0.7532	0.5930
Q_1, P_2	0.7657	0.5946	0.7595	0.5926
Q_1, P_1	0.8299	0.6406	0.7934	0.6283
P_3, P_2	0.6731	0.5984	0.7402	0.5729
P_3, P_1	0.7607	0.7386	0.8009	0.6027
P_2, P_1	0.7631	0.7349	0.8067	0.6047
Q_3, Q_2, Q_1	0.6952	0.5706	0.7693	0.6200
Q_3, Q_2, P_3	0.6100	0.5784	0.7379	0.5835
Q_3, Q_2, P_2	0.6075	0.5798	0.7378	0.5880
Q_3, Q_2, P_1	0.7167	0.6201	0.7854	0.6255
Q_3, Q_1, P_3	0.7560	0.5708	0.7516	0.6052
Q_3, Q_1, P_2	0.7507	0.6032	0.7607	0.6034
Q_3, Q_1, P_1	0.8129	0.6330	0.7857	0.6217
Q_3, P_3, P_2	0.6699	0.5949	0.7297	0.5813
Q_3, P_3, P_1	0.7507	0.7209	0.7966	0.6173
Q_3, P_2, P_1	0.7485	0.7198	0.7910	0.6094
Q_2, Q_1, P_3	0.6925	0.5963	0.7723	0.5982
Q_2, Q_1, P_2	0.6950	0.5896	0.7773	0.5999
Q_2, Q_1, P_1	0.7711	0.6376	0.8059	0.6420
Q_2, P_3, P_2	0.6308	0.5890	0.7418	0.5811
Q_2, P_3, P_1	0.7187	0.7151	0.7992	0.6095
Q_2, P_2, P_1	0.7181	0.7290	0.8165	0.6140
Q_1, P_3, P_2	0.7330	0.5935	0.7623	0.5968
Q_1, P_3, P_1	0.7951	0.7096	0.7934	0.6220
Q_1, P_2, P_1	0.7926	0.7060	0.8023	0.6264
P_3, P_2, P_1	0.7391	0.7388	0.8016	0.5958
Q_3, Q_2, Q_1, P_3	0.6903	0.6116	0.7872	0.6062
Q_3, Q_2, Q_1, P_2	0.6872	0.6169	0.7861	0.5911
Q_3, Q_2, Q_1, P_1	0.7593	0.6470	0.8098	0.6410
Q_3, Q_2, P_3, P_2	0.6325	0.5846	0.7370	0.5868
Q_3, Q_2, P_3, P_1	0.7089	0.7115	0.7963	0.6096
Q_3, Q_2, P_2, P_1	0.7081	0.7219	0.7970	0.6134
Q_3, Q_1, P_3, P_2	0.7266	0.6026	0.7680	0.6088
Q_3, Q_1, P_3, P_1	0.7760	0.6973	0.7994	0.6221

Input combination	Classification method			
	NB	LR	SVM	RF
Q_3, Q_1, P_2, P_1	0.7754	0.6911	0.7965	0.6211
Q_3, P_3, P_2, P_1	0.7314	0.7321	0.7962	0.6056
Q_2, Q_1, P_3, P_2	0.6908	0.6039	0.7773	0.5978
Q_2, Q_1, P_3, P_1	0.7517	0.6999	0.8006	0.6297
Q_2, Q_1, P_2, P_1	0.7514	0.7156	0.8152	0.6207
Q_2, P_3, P_2, P_1	0.7081	0.7386	0.8059	0.6005
Q_1, P_3, P_2, P_1	0.7682	0.7115	0.7982	0.6213
Q_3, Q_2, Q_1, P_3, P_2	0.6872	0.6186	0.7858	0.6013
Q_3, Q_2, Q_1, P_3, P_1	0.7402	0.7121	0.8115	0.6127
Q_3, Q_2, Q_1, P_2, P_1	0.7394	0.7175	0.8103	0.6386
Q_3, Q_2, P_3, P_2, P_1	0.7022	0.7310	0.7947	0.6029
Q_3, Q_1, P_3, P_2, P_1	0.7587	0.7016	0.7968	0.6182
Q_2, Q_1, P_3, P_2, P_1	0.7390	0.7234	0.8022	0.6149
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.7322	0.7267	0.8050	0.6106

Table C.2: The sensitivities achieved across the combination search by each of the four classification methods.

Input combination	Classification method			
	NB	LR	SVM	RF
Q_3	0.2660	0.4720	0.4540	0.4421
Q_2	0.5570	0.4484	0.4344	0.4733
Q_1	0.2067	0.4797	0.5022	0.4125
P_3	0.3657	0.4608	0.4574	0.4689
P_2	0.3757	0.4730	0.4512	0.4930
P_1	0.2150	0.4472	0.4467	0.4255
Q_3, Q_2	0.4762	0.4618	0.4873	0.4497
Q_3, Q_1	0.2199	0.5061	0.5497	0.4239
Q_3, P_3	0.3550	0.4741	0.5137	0.4821
Q_3, P_2	0.3685	0.4732	0.5181	0.4906
Q_3, P_1	0.2331	0.4682	0.5065	0.4517
Q_2, Q_1	0.3508	0.4724	0.5268	0.4515
Q_2, P_3	0.4594	0.4793	0.4885	0.4838
Q_2, P_2	0.4604	0.4810	0.4729	0.4788
Q_2, P_1	0.3171	0.4663	0.4633	0.4385
Q_1, P_3	0.2799	0.4958	0.5407	0.4635
Q_1, P_2	0.2900	0.5089	0.5471	0.4800
Q_1, P_1	0.2179	0.4764	0.5361	0.4304
P_3, P_2	0.3718	0.4738	0.4649	0.5015
P_3, P_1	0.2784	0.4510	0.4683	0.4740
P_2, P_1	0.2834	0.4676	0.4664	0.4784
Q_3, Q_2, Q_1	0.3448	0.4882	0.5399	0.4480
Q_3, Q_2, P_3	0.4317	0.4840	0.5260	0.4823
Q_3, Q_2, P_2	0.4372	0.4852	0.5186	0.4820
Q_3, Q_2, P_1	0.3166	0.4729	0.5104	0.4515
Q_3, Q_1, P_3	0.2850	0.5143	0.5662	0.4597
Q_3, Q_1, P_2	0.2969	0.5027	0.5672	0.4829
Q_3, Q_1, P_1	0.2265	0.4917	0.5658	0.4484
Q_3, P_3, P_2	0.3707	0.4736	0.5150	0.4838
Q_3, P_3, P_1	0.2837	0.4756	0.5137	0.4670
Q_3, P_2, P_1	0.2908	0.4778	0.5164	0.4835
Q_2, Q_1, P_3	0.3734	0.4962	0.5410	0.4718
Q_2, Q_1, P_2	0.3721	0.5149	0.5373	0.4684
Q_2, Q_1, P_1	0.2863	0.4865	0.5343	0.4408
Q_2, P_3, P_2	0.4254	0.4817	0.4929	0.4959
Q_2, P_3, P_1	0.3324	0.4726	0.4869	0.4764
Q_2, P_2, P_1	0.3371	0.4795	0.4789	0.4811
Q_1, P_3, P_2	0.3288	0.5129	0.5487	0.4926
Q_1, P_3, P_1	0.2497	0.4859	0.5448	0.4530
Q_1, P_2, P_1	0.2568	0.4970	0.5417	0.4759
P_3, P_2, P_1	0.3092	0.4806	0.4825	0.4985
Q_3, Q_2, Q_1, P_3	0.3668	0.4996	0.5551	0.4675
Q_3, Q_2, Q_1, P_2	0.3697	0.5095	0.5548	0.4773
Q_3, Q_2, Q_1, P_1	0.2908	0.4942	0.5560	0.4466
Q_3, Q_2, P_3, P_2	0.4196	0.4843	0.5322	0.4944
Q_3, Q_2, P_3, P_1	0.3335	0.4847	0.5282	0.4784
Q_3, Q_2, P_2, P_1	0.3376	0.4827	0.5252	0.4841
Q_3, Q_1, P_3, P_2	0.3287	0.5122	0.5587	0.4736
Q_3, Q_1, P_3, P_1	0.2610	0.5051	0.5688	0.4646

Input combination	Classification method			
	NB	LR	SVM	RF
Q_3, Q_1, P_2, P_1	0.2634	0.5074	0.5673	0.4724
Q_3, P_3, P_2, P_1	0.3127	0.4927	0.5190	0.4853
Q_2, Q_1, P_3, P_2	0.3813	0.5147	0.5434	0.4791
Q_2, Q_1, P_3, P_1	0.3078	0.4979	0.5475	0.4633
Q_2, Q_1, P_2, P_1	0.3104	0.5090	0.5416	0.4809
Q_2, P_3, P_2, P_1	0.3473	0.4903	0.4920	0.4954
Q_1, P_3, P_2, P_1	0.2865	0.5134	0.5497	0.4819
Q_3, Q_2, Q_1, P_3, P_2	0.3762	0.5106	0.5552	0.4865
Q_3, Q_2, Q_1, P_3, P_1	0.3105	0.5096	0.5683	0.4785
Q_3, Q_2, Q_1, P_2, P_1	0.3151	0.5095	0.5640	0.4708
Q_3, Q_2, P_3, P_2, P_1	0.3504	0.4930	0.5340	0.4816
Q_3, Q_1, P_3, P_2, P_1	0.2902	0.5170	0.5707	0.4844
Q_2, Q_1, P_3, P_2, P_1	0.3251	0.5112	0.5516	0.4852
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.3308	0.5221	0.5694	0.4941

Table C.3: The specificities achieved across the combination search by each of the four classification methods.

Appendix D

IVBC combination search results

The F_1 scores, sensitivities, and specificities achieved across the IVBC combination searches are shown in Tables D.1, D.2, and D.3, respectively.

Input combination	LR			SVM		
	Aortic	Iliac 1	Iliac 2	Aortic	Iliac 1	Iliac 2
Q_3	0.5588	0.4356	0.4974	0.6431	0.5043	0.6056
Q_2	0.5661	0.4953	0.4810	0.6515	0.6057	0.5750
Q_1	0.5423	0.4895	0.5103	0.7010	0.6452	0.6686
P_3	0.5664	0.5226	0.5659	0.6700	0.6354	0.6647
P_2	0.5666	0.5650	0.5233	0.6716	0.6596	0.6309
P_1	0.6395	0.5065	0.5171	0.7332	0.6143	0.6035
Q_3, Q_2	0.5622	0.4891	0.5144	0.6909	0.6326	0.6654
Q_3, Q_1	0.5626	0.4816	0.5266	0.7323	0.6715	0.7166
Q_3, P_3	0.5654	0.5210	0.5631	0.6939	0.6519	0.6869
Q_3, P_2	0.5701	0.5759	0.5401	0.6941	0.7083	0.6894
Q_3, P_1	0.6391	0.5081	0.5283	0.7717	0.6444	0.6632
Q_2, Q_1	0.5629	0.5033	0.5168	0.7339	0.6878	0.6844
Q_2, P_3	0.5638	0.5273	0.5806	0.6981	0.6782	0.7108
Q_2, P_2	0.5622	0.5652	0.5205	0.6965	0.6850	0.6528
Q_2, P_1	0.6405	0.5208	0.5199	0.7733	0.6584	0.6534
Q_1, P_3	0.5832	0.5226	0.5834	0.7586	0.7118	0.7456
Q_1, P_2	0.5893	0.5865	0.5320	0.7590	0.7448	0.7098
Q_1, P_1	0.6843	0.5040	0.5125	0.8301	0.6996	0.7059
P_3, P_2	0.5658	0.7746	0.7800	0.6829	0.7478	0.7437
P_3, P_1	0.7233	0.5425	0.6456	0.7853	0.6477	0.7149
P_2, P_1	0.7235	0.6392	0.5303	0.7854	0.7156	0.6270
Q_3, Q_2, Q_1	0.5628	0.5014	0.5374	0.7572	0.7192	0.7422
Q_3, Q_2, P_3	0.5651	0.5369	0.5783	0.7221	0.7069	0.7328
Q_3, Q_2, P_2	0.5675	0.5754	0.5498	0.7210	0.7355	0.7175
Q_3, Q_2, P_1	0.6417	0.5144	0.5369	0.7935	0.6844	0.7020
Q_3, Q_1, P_3	0.5806	0.5271	0.5794	0.7693	0.7267	0.7687
Q_3, Q_1, P_2	0.5949	0.6066	0.5367	0.7739	0.7763	0.7475
Q_3, Q_1, P_1	0.6844	0.5028	0.5327	0.8346	0.7149	0.7409
Q_3, P_3, P_2	0.5745	0.7821	0.7680	0.7024	0.7825	0.7728
Q_3, P_3, P_1	0.7300	0.5465	0.6477	0.7980	0.6670	0.7329
Q_3, P_2, P_1	0.7201	0.6455	0.5400	0.7940	0.7368	0.6892
Q_2, Q_1, P_3	0.5881	0.5240	0.6055	0.7745	0.7331	0.7824
Q_2, Q_1, P_2	0.5864	0.5815	0.5354	0.7701	0.7546	0.7239
Q_2, Q_1, P_1	0.6901	0.5169	0.5151	0.8329	0.7246	0.7182
Q_2, P_3, P_2	0.5620	0.7650	0.7857	0.7076	0.7680	0.7721
Q_2, P_3, P_1	0.7220	0.5492	0.6507	0.7953	0.6869	0.7366
Q_2, P_2, P_1	0.7373	0.6330	0.5290	0.7960	0.7272	0.6632
Q_1, P_3, P_2	0.5926	0.7646	0.7647	0.7631	0.7775	0.7717
Q_1, P_3, P_1	0.7291	0.5439	0.6408	0.8259	0.7084	0.7606
Q_1, P_2, P_1	0.7329	0.6481	0.5392	0.8249	0.7614	0.7052

Input combination	LR			SVM		
	Aortic	Iliac 1	Iliac 2	Aortic	Iliac 1	Iliac 2
P_3, P_2, P_1	0.7265	0.7698	0.7728	0.7869	0.7499	0.7414
Q_3, Q_2, Q_1, P_3	0.5866	0.5330	0.6016	0.7857	0.7489	0.8039
Q_3, Q_2, Q_1, P_2	0.5947	0.6009	0.5501	0.7858	0.7918	0.7641
Q_3, Q_2, Q_1, P_1	0.6898	0.5156	0.5412	0.8437	0.7388	0.7538
Q_3, Q_2, P_3, P_2	0.5693	0.8255	0.8167	0.7259	0.8256	0.8163
Q_3, Q_2, P_3, P_1	0.7288	0.5538	0.6481	0.8053	0.7107	0.7550
Q_3, Q_2, P_2, P_1	0.7358	0.6402	0.5517	0.8049	0.7557	0.7175
Q_3, Q_1, P_3, P_2	0.5975	0.7783	0.7686	0.7769	0.8188	0.8115
Q_3, Q_1, P_3, P_1	0.7358	0.5490	0.6459	0.8322	0.7266	0.7857
Q_3, Q_1, P_2, P_1	0.7352	0.6461	0.5562	0.8309	0.7851	0.7465
Q_3, P_3, P_2, P_1	0.7309	0.7762	0.7693	0.7967	0.7880	0.7910
Q_2, Q_1, P_3, P_2	0.5932	0.7752	0.7818	0.7789	0.8033	0.8144
Q_2, Q_1, P_3, P_1	0.7325	0.5551	0.6498	0.8310	0.7359	0.7857
Q_2, Q_1, P_2, P_1	0.7438	0.6422	0.5397	0.8322	0.7717	0.7197
Q_2, P_3, P_2, P_1	0.7353	0.7655	0.7771	0.7968	0.7898	0.7784
Q_1, P_3, P_2, P_1	0.7358	0.7606	0.7583	0.8213	0.7759	0.7707
Q_3, Q_2, Q_1, P_3, P_2	0.5932	0.8147	0.8069	0.7875	0.8387	0.8407
Q_3, Q_2, Q_1, P_3, P_1	0.7412	0.5568	0.6498	0.8387	0.7550	0.8028
Q_3, Q_2, Q_1, P_2, P_1	0.7466	0.6401	0.5616	0.8391	0.7961	0.7622
Q_3, Q_2, P_3, P_2, P_1	0.7352	0.8241	0.8139	0.8051	0.8219	0.8180
Q_3, Q_1, P_3, P_2, P_1	0.7391	0.7764	0.7728	0.8283	0.8208	0.8190
Q_2, Q_1, P_3, P_2, P_1	0.7440	0.7738	0.7733	0.8269	0.8096	0.8125
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.7461	0.8208	0.8086	0.8363	0.8348	0.8364

Table D.1: The F_1 scores achieved across the IVBC combination searches by the LR and SVM classification methods.

Input combination	LR			SVM		
	Aortic	Iliac 1	Iliac 2	Aortic	Iliac 1	Iliac 2
Q_3	0.5937	0.3853	0.4850	0.6992	0.4245	0.5942
Q_2	0.6166	0.4791	0.4844	0.7289	0.6135	0.5500
Q_1	0.5691	0.4990	0.5216	0.7338	0.5971	0.6517
P_3	0.6065	0.5310	0.5868	0.7239	0.6156	0.6873
P_2	0.6096	0.5847	0.5390	0.7212	0.6410	0.6332
P_1	0.6932	0.5064	0.5276	0.7930	0.5808	0.5705
Q_3, Q_2	0.6020	0.4694	0.5210	0.7591	0.5777	0.6411
Q_3, Q_1	0.5861	0.4751	0.5303	0.7549	0.6045	0.7036
Q_3, P_3	0.6055	0.5236	0.5784	0.7258	0.6110	0.6891
Q_3, P_2	0.6117	0.6006	0.5435	0.7248	0.7045	0.6733
Q_3, P_1	0.6958	0.5048	0.5442	0.8258	0.5967	0.6457
Q_2, Q_1	0.5957	0.4969	0.5320	0.7694	0.6324	0.6541
Q_2, P_3	0.6086	0.5149	0.6119	0.7536	0.6376	0.7327
Q_2, P_2	0.6030	0.5772	0.5377	0.7538	0.6744	0.6407
Q_2, P_1	0.6978	0.5183	0.5374	0.8416	0.6327	0.6285
Q_1, P_3	0.6262	0.5325	0.6023	0.7755	0.6488	0.7341
Q_1, P_2	0.6302	0.5988	0.5343	0.7764	0.7207	0.6854
Q_1, P_1	0.7549	0.4968	0.5078	0.8779	0.6385	0.6629
P_3, P_2	0.6108	0.8456	0.8583	0.7330	0.7637	0.7817
P_3, P_1	0.8102	0.4944	0.7059	0.8471	0.6036	0.7590
P_2, P_1	0.8065	0.6963	0.4919	0.8454	0.7481	0.5741
Q_3, Q_2, Q_1	0.5902	0.4941	0.5501	0.7899	0.6566	0.7202
Q_3, Q_2, P_3	0.6072	0.5351	0.6046	0.7635	0.6667	0.7376
Q_3, Q_2, P_2	0.6101	0.5912	0.5629	0.7614	0.7339	0.6940
Q_3, Q_2, P_1	0.6986	0.5144	0.5577	0.8508	0.6394	0.6857
Q_3, Q_1, P_3	0.6177	0.5405	0.5851	0.7826	0.6695	0.7583
Q_3, Q_1, P_2	0.6400	0.6358	0.5318	0.7919	0.7631	0.7136
Q_3, Q_1, P_1	0.7532	0.4911	0.5272	0.8790	0.6522	0.7067
Q_3, P_3, P_2	0.6239	0.8500	0.7952	0.7356	0.8219	0.7660
Q_3, P_3, P_1	0.8069	0.4988	0.6968	0.8548	0.6122	0.7529
Q_3, P_2, P_1	0.8023	0.6937	0.5076	0.8503	0.7523	0.6437
Q_2, Q_1, P_3	0.6329	0.5124	0.6436	0.8043	0.6743	0.7861
Q_2, Q_1, P_2	0.6292	0.5812	0.5465	0.7978	0.7328	0.6938
Q_2, Q_1, P_1	0.7612	0.5049	0.5232	0.8811	0.6683	0.6797
Q_2, P_3, P_2	0.6059	0.7882	0.8550	0.7607	0.7520	0.8307
Q_2, P_3, P_1	0.8055	0.5029	0.7119	0.8524	0.6340	0.7709
Q_2, P_2, P_1	0.8169	0.6692	0.5051	0.8576	0.7400	0.6165
Q_1, P_3, P_2	0.6354	0.8278	0.8265	0.7876	0.7660	0.7731
Q_1, P_3, P_1	0.8070	0.5115	0.6879	0.8700	0.6476	0.7670
Q_1, P_2, P_1	0.8049	0.6871	0.5168	0.8662	0.7595	0.6538
P_3, P_2, P_1	0.8067	0.8150	0.8229	0.8487	0.7658	0.7633
Q_3, Q_2, Q_1, P_3	0.6289	0.5288	0.6266	0.8076	0.6904	0.8054
Q_3, Q_2, Q_1, P_2	0.6365	0.6110	0.5596	0.8077	0.7937	0.7322
Q_3, Q_2, Q_1, P_1	0.7537	0.5016	0.5509	0.8893	0.6819	0.7169
Q_3, Q_2, P_3, P_2	0.6162	0.8627	0.8549	0.7631	0.8439	0.8303
Q_3, Q_2, P_3, P_1	0.8086	0.5132	0.6983	0.8581	0.6552	0.7652
Q_3, Q_2, P_2, P_1	0.8138	0.6736	0.5337	0.8574	0.7631	0.6761
Q_3, Q_1, P_3, P_2	0.6424	0.8407	0.7974	0.7949	0.8245	0.8028
Q_3, Q_1, P_3, P_1	0.8091	0.5181	0.6845	0.8729	0.6630	0.7847

Input combination	LR			SVM		
	Aortic	Iliac 1	Iliac 2	Aortic	Iliac 1	Iliac 2
Q_3, Q_1, P_2, P_1	0.8083	0.6830	0.5379	0.8686	0.7800	0.7047
Q_3, P_3, P_2, P_1	0.8054	0.8194	0.7916	0.8514	0.8050	0.7969
Q_2, Q_1, P_3, P_2	0.6359	0.8010	0.8413	0.8105	0.7784	0.8351
Q_2, Q_1, P_3, P_1	0.8104	0.5246	0.7019	0.8735	0.6765	0.7940
Q_2, Q_1, P_2, P_1	0.8219	0.6682	0.5249	0.8762	0.7607	0.6734
Q_2, P_3, P_2, P_1	0.8133	0.7799	0.8236	0.8561	0.7863	0.8126
Q_1, P_3, P_2, P_1	0.8093	0.8070	0.8024	0.8633	0.7673	0.7711
Q_3, Q_2, Q_1, P_3, P_2	0.6387	0.8516	0.8452	0.8109	0.8333	0.8406
Q_3, Q_2, Q_1, P_3, P_1	0.8164	0.5283	0.6906	0.8796	0.6936	0.7987
Q_3, Q_2, Q_1, P_2, P_1	0.8213	0.6673	0.5517	0.8775	0.7907	0.7172
Q_3, Q_2, P_3, P_2, P_1	0.8133	0.8496	0.8457	0.8571	0.8300	0.8187
Q_3, Q_1, P_3, P_2, P_1	0.8103	0.8190	0.7974	0.8658	0.8197	0.8138
Q_2, Q_1, P_3, P_2, P_1	0.8218	0.7905	0.8176	0.8696	0.7923	0.8236
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.8197	0.8451	0.8408	0.8734	0.8255	0.8276

Table D.2: The sensitivities achieved across the IVBC combination searches by the LR and SVM classification methods.

Input combination	LR			SVM		
	Aortic	Iliac 1	Iliac 2	Aortic	Iliac 1	Iliac 2
Q_3	0.4688	0.6165	0.5352	0.5250	0.7410	0.6320
Q_2	0.4382	0.5448	0.4703	0.4915	0.5879	0.6371
Q_1	0.4703	0.4602	0.4777	0.6405	0.7463	0.7023
P_3	0.4652	0.4989	0.5133	0.5633	0.6781	0.6195
P_2	0.4580	0.5150	0.4793	0.5735	0.6976	0.6260
P_1	0.5255	0.5070	0.4871	0.6300	0.6899	0.6799
Q_3, Q_2	0.4605	0.5502	0.4956	0.5619	0.7514	0.7144
Q_3, Q_1	0.5026	0.5023	0.5166	0.6934	0.8042	0.7401
Q_3, P_3	0.4638	0.5139	0.5242	0.6341	0.7365	0.6828
Q_3, P_2	0.4658	0.5150	0.5312	0.6366	0.7153	0.7202
Q_3, P_1	0.5184	0.5181	0.4842	0.6857	0.7448	0.6985
Q_2, Q_1	0.4793	0.5225	0.4734	0.6729	0.7937	0.7428
Q_2, P_3	0.4500	0.5623	0.5041	0.5949	0.7574	0.6711
Q_2, P_2	0.4581	0.5348	0.4718	0.5893	0.7056	0.6780
Q_2, P_1	0.5189	0.5281	0.4703	0.6652	0.7109	0.7049
Q_1, P_3	0.4791	0.4947	0.5377	0.7312	0.8260	0.7651
Q_1, P_2	0.4917	0.5571	0.5258	0.7307	0.7856	0.7544
Q_1, P_1	0.5487	0.5256	0.5265	0.7628	0.8132	0.7849
P_3, P_2	0.4521	0.6623	0.6576	0.5865	0.7214	0.6797
P_3, P_1	0.5700	0.6718	0.5194	0.6899	0.7400	0.6358
P_2, P_1	0.5771	0.5179	0.6370	0.6927	0.6573	0.7430
Q_3, Q_2, Q_1	0.4931	0.5235	0.5032	0.7038	0.8308	0.7795
Q_3, Q_2, P_3	0.4582	0.5420	0.5137	0.6489	0.7805	0.7247
Q_3, Q_2, P_2	0.4601	0.5363	0.5153	0.6496	0.7383	0.7596
Q_3, Q_2, P_1	0.5214	0.5146	0.4805	0.7065	0.7711	0.7323
Q_3, Q_1, P_3	0.4902	0.4900	0.5656	0.7481	0.8271	0.7855
Q_3, Q_1, P_2	0.4885	0.5397	0.5501	0.7456	0.7972	0.8045
Q_3, Q_1, P_1	0.5524	0.5380	0.5481	0.7727	0.8278	0.7991
Q_3, P_3, P_2	0.4522	0.6766	0.7246	0.6412	0.7213	0.7837
Q_3, P_3, P_1	0.5963	0.6736	0.5452	0.7127	0.7767	0.6986
Q_3, P_2, P_1	0.5742	0.5445	0.6277	0.7087	0.7104	0.7759
Q_2, Q_1, P_3	0.4807	0.5569	0.5180	0.7275	0.8349	0.7768
Q_2, Q_1, P_2	0.4833	0.5823	0.5051	0.7259	0.7907	0.7772
Q_2, Q_1, P_1	0.5553	0.5517	0.4920	0.7654	0.8239	0.7870
Q_2, P_3, P_2	0.4499	0.7278	0.6786	0.6109	0.7937	0.6791
Q_2, P_3, P_1	0.5744	0.6717	0.5240	0.7090	0.7882	0.6779
Q_2, P_2, P_1	0.6012	0.5550	0.5956	0.7030	0.7050	0.7576
Q_1, P_3, P_2	0.4910	0.6625	0.6649	0.7236	0.7956	0.7697
Q_1, P_3, P_1	0.5935	0.6307	0.5409	0.7633	0.8193	0.7503
Q_1, P_2, P_1	0.6086	0.5670	0.6002	0.7662	0.7647	0.7998
P_3, P_2, P_1	0.5862	0.6977	0.6933	0.6918	0.7235	0.7045
Q_3, Q_2, Q_1, P_3	0.4847	0.5449	0.5435	0.7519	0.8467	0.8018
Q_3, Q_2, Q_1, P_2	0.4962	0.5774	0.5252	0.7521	0.7891	0.8157
Q_3, Q_2, Q_1, P_1	0.5685	0.5562	0.5154	0.7814	0.8360	0.8150
Q_3, Q_2, P_3, P_2	0.4518	0.7727	0.7614	0.6608	0.7996	0.7961
Q_3, Q_2, P_3, P_1	0.5898	0.6600	0.5434	0.7270	0.8114	0.7383
Q_3, Q_2, P_2, P_1	0.6020	0.5693	0.5993	0.7271	0.7436	0.7916
Q_3, Q_1, P_3, P_2	0.4922	0.6805	0.7225	0.7488	0.8107	0.8244
Q_3, Q_1, P_3, P_1	0.6100	0.6309	0.5653	0.7752	0.8381	0.7874

Input combination	LR			SVM		
	Aortic	Iliac 1	Iliac 2	Aortic	Iliac 1	Iliac 2
Q_3, Q_1, P_2, P_1	0.6096	0.5688	0.6038	0.7780	0.7930	0.8169
Q_3, P_3, P_2, P_1	0.6017	0.7082	0.7338	0.7141	0.7619	0.7821
Q_2, Q_1, P_3, P_2	0.4921	0.7347	0.6891	0.7296	0.8405	0.7843
Q_2, Q_1, P_3, P_1	0.5978	0.6345	0.5418	0.7713	0.8381	0.7730
Q_2, Q_1, P_2, P_1	0.6119	0.5873	0.5798	0.7707	0.7893	0.8022
Q_2, P_3, P_2, P_1	0.6012	0.7425	0.7040	0.7074	0.7954	0.7248
Q_1, P_3, P_2, P_1	0.6096	0.6850	0.6861	0.7612	0.7895	0.7702
Q_3, Q_2, Q_1, P_3, P_2	0.4856	0.7612	0.7505	0.7516	0.8464	0.8409
Q_3, Q_2, Q_1, P_3, P_1	0.6137	0.6308	0.5653	0.7821	0.8563	0.8090
Q_3, Q_2, Q_1, P_2, P_1	0.6212	0.5825	0.5870	0.7862	0.8045	0.8354
Q_3, Q_2, P_3, P_2, P_1	0.6010	0.7878	0.7677	0.7281	0.8103	0.8172
Q_3, Q_1, P_3, P_2, P_1	0.6177	0.7093	0.7338	0.7755	0.8226	0.8266
Q_2, Q_1, P_3, P_2, P_1	0.6128	0.7475	0.7032	0.7665	0.8352	0.7964
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.6225	0.7859	0.7614	0.7847	0.8479	0.8488

Table D.3: The specificities achieved across the IVBC combination searches by the LR and SVM classification methods.

Appendix E

OVA and OVO multiclass combination search results

The F_1 scores, sensitivities, and specificities achieved across the multiclass combination searches are shown in Tables E.1, E.2, and E.3, respectively.

Input combination	LR			SVM		
	Aorta	Iliac 1	Iliac 2	Aorta	Iliac 1	Iliac 2
Q_3	0.4459	0.4396	0.3545	0.3521	0.4645	0.4344
Q_2	0.3697	0.4555	0.3907	0.3490	0.4349	0.4754
Q_1	0.4515	0.3722	0.3856	0.3888	0.4332	0.4123
P_3	0.3521	0.4618	0.4951	0.3764	0.4661	0.4613
P_2	0.3732	0.4887	0.4749	0.3904	0.4662	0.4217
P_1	0.5676	0.3791	0.3967	0.5337	0.4054	0.4315
Q_3, Q_2	0.3791	0.4551	0.4065	0.3712	0.4871	0.4852
Q_3, Q_1	0.4615	0.3904	0.4242	0.4258	0.4491	0.4345
Q_3, P_3	0.3657	0.4607	0.4865	0.3699	0.4666	0.4561
Q_3, P_2	0.3791	0.4875	0.4956	0.4044	0.5050	0.4929
Q_3, P_1	0.5676	0.4327	0.3974	0.5346	0.4596	0.4467
Q_2, Q_1	0.4422	0.4358	0.3847	0.4178	0.4640	0.4118
Q_2, P_3	0.3295	0.5138	0.4980	0.3945	0.5351	0.4794
Q_2, P_2	0.3306	0.5012	0.4671	0.3757	0.4723	0.4625
Q_2, P_1	0.5507	0.4148	0.4068	0.5175	0.4672	0.4713
Q_1, P_3	0.4217	0.4756	0.5004	0.4472	0.4789	0.4627
Q_1, P_2	0.4284	0.5190	0.4848	0.4462	0.4712	0.4508
Q_1, P_1	0.6028	0.4289	0.4309	0.5673	0.4341	0.4314
P_3, P_2	0.4763	0.7395	0.7516	0.4524	0.5658	0.5556
P_3, P_1	0.6246	0.6222	0.5357	0.5619	0.5459	0.4938
P_2, P_1	0.6240	0.5302	0.5962	0.5609	0.4701	0.5302
Q_3, Q_2, Q_1	0.4294	0.4232	0.4302	0.4250	0.4651	0.4474
Q_3, Q_2, P_3	0.3417	0.5018	0.4983	0.4011	0.5232	0.4928
Q_3, Q_2, P_2	0.3369	0.5117	0.4904	0.3939	0.5181	0.5222
Q_3, Q_2, P_1	0.5543	0.4454	0.4167	0.5300	0.4997	0.4815
Q_3, Q_1, P_3	0.4230	0.4602	0.5103	0.4498	0.4748	0.4736
Q_3, Q_1, P_2	0.4319	0.5069	0.5180	0.4659	0.5099	0.4859
Q_3, Q_1, P_1	0.6000	0.4341	0.4511	0.5530	0.4684	0.4634
Q_3, P_3, P_2	0.4685	0.7450	0.7662	0.4535	0.5703	0.5994
Q_3, P_3, P_1	0.6401	0.6107	0.5529	0.5333	0.5383	0.4943
Q_3, P_2, P_1	0.6166	0.5482	0.5994	0.5485	0.5211	0.5364
Q_2, Q_1, P_3	0.4111	0.5345	0.4884	0.4503	0.5155	0.4941
Q_2, Q_1, P_2	0.3912	0.5354	0.4782	0.4338	0.4757	0.4505
Q_2, Q_1, P_1	0.5974	0.4746	0.4337	0.5398	0.4612	0.4321
Q_2, P_3, P_2	0.4690	0.7608	0.7523	0.4638	0.6110	0.5518
Q_2, P_3, P_1	0.6202	0.6228	0.5397	0.5537	0.5641	0.5098
Q_2, P_2, P_1	0.6457	0.5608	0.5730	0.5351	0.4829	0.5308

Input combination	LR			SVM		
	Aorta	Iliac 1	Iliac 2	Aorta	Iliac 1	Iliac 2
Q_1, P_3, P_2	0.5185	0.7395	0.7433	0.4691	0.5443	0.5281
Q_1, P_3, P_1	0.6414	0.5945	0.5607	0.5686	0.5272	0.5021
Q_1, P_2, P_1	0.6538	0.5631	0.5780	0.5651	0.4877	0.4960
P_3, P_2, P_1	0.6681	0.7511	0.7486	0.5722	0.5958	0.5775
Q_3, Q_2, Q_1, P_3	0.4076	0.5209	0.5042	0.4437	0.5234	0.5074
Q_3, Q_2, Q_1, P_2	0.4058	0.5354	0.5093	0.4629	0.5129	0.5018
Q_3, Q_2, Q_1, P_1	0.5996	0.4683	0.4357	0.5398	0.4964	0.4747
Q_3, Q_2, P_3, P_2	0.5214	0.8043	0.7888	0.4883	0.6390	0.6024
Q_3, Q_2, P_3, P_1	0.6307	0.6179	0.5470	0.5365	0.5617	0.5106
Q_3, Q_2, P_2, P_1	0.6440	0.5660	0.5840	0.5373	0.5320	0.5430
Q_3, Q_1, P_3, P_2	0.5198	0.7382	0.7614	0.5031	0.5801	0.5721
Q_3, Q_1, P_3, P_1	0.6626	0.5818	0.5780	0.5535	0.5252	0.5003
Q_3, Q_1, P_2, P_1	0.6592	0.5710	0.5896	0.5515	0.5264	0.5162
Q_3, P_3, P_2, P_1	0.6808	0.7548	0.7687	0.5636	0.6330	0.6196
Q_2, Q_1, P_3, P_2	0.5172	0.7670	0.7399	0.4945	0.5996	0.5562
Q_2, Q_1, P_3, P_1	0.6422	0.6018	0.5615	0.5682	0.5383	0.5177
Q_2, Q_1, P_2, P_1	0.6619	0.5812	0.5633	0.5644	0.4877	0.4961
Q_2, P_3, P_2, P_1	0.6682	0.7728	0.7422	0.5839	0.6371	0.6040
Q_1, P_3, P_2, P_1	0.6865	0.7493	0.7457	0.5723	0.5625	0.5468
Q_3, Q_2, Q_1, P_3, P_2	0.5572	0.7951	0.7884	0.5021	0.6162	0.5877
Q_3, Q_2, Q_1, P_3, P_1	0.6593	0.5875	0.5795	0.5507	0.5507	0.5187
Q_3, Q_2, Q_1, P_2, P_1	0.6639	0.5857	0.5719	0.5477	0.5243	0.5265
Q_3, Q_2, P_3, P_2, P_1	0.7053	0.8060	0.7893	0.5780	0.6584	0.6244
Q_3, Q_1, P_3, P_2, P_1	0.6921	0.7478	0.7678	0.5735	0.6159	0.5917
Q_2, Q_1, P_3, P_2, P_1	0.6943	0.7741	0.7398	0.5905	0.6184	0.5814
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.7171	0.7993	0.7857	0.5811	0.6284	0.6024

Table E.1: The F_1 scores achieved across the multiclass combination search on the stenosis VPD by the LR and SVM classification methods.

Input combination	LR				SVM			
	Healthy	Aorta	Iliac 1	Iliac 2	Healthy	Aorta	Iliac 1	Iliac 2
Q_3	0.0008	0.3792	0.3918	0.2754	0.0116	0.2619	0.4325	0.3661
Q_2	0	0.2845	0.4134	0.3252	0.0073	0.2593	0.3648	0.4481
Q_1	0	0.3920	0.3048	0.3155	0.0095	0.3063	0.3801	0.3474
P_3	0	0.2620	0.4205	0.4436	0.0127	0.2883	0.4265	0.3947
P_2	0	0.2822	0.4320	0.4372	0.0102	0.3049	0.4064	0.3680
P_1	0.0006	0.5216	0.3074	0.3311	0.0102	0.4420	0.3405	0.3852
Q_3, Q_2	0.0005	0.2977	0.4074	0.3406	0.0260	0.2831	0.4382	0.4286
Q_3, Q_1	0.0009	0.3997	0.3217	0.3566	0.0313	0.3455	0.4002	0.3559
Q_3, P_3	0.008	0.2788	0.4163	0.4299	0.0270	0.2860	0.4227	0.3842
Q_3, P_2	0.0005	0.2895	0.4227	0.4660	0.02.0	0.3251	0.4402	0.4383
Q_3, P_1	0.0017	0.5147	0.3743	0.3228	0.0349	0.4361	0.4112	0.3862
Q_2, Q_1	0	0.3705	0.3772	0.3164	0.0233	0.3318	0.4081	0.3494
Q_2, P_3	0	0.2402	0.4913	0.4398	0.0192	0.3099	0.5058	0.4052
Q_2, P_2	0.0005	0.2437	0.4505	0.4260	0.0158	0.2872	0.4075	0.4201
Q_2, P_1	0.0009	0.4894	0.3462	0.3455	0.0222	0.4142	0.4073	0.4319
Q_1, P_3	0.0003	0.3314	0.4346	0.4421	0.0459	0.3577	0.4379	0.3846
Q_1, P_2	0.0011	0.3413	0.4624	0.4411	0.0450	0.3589	0.3998	0.3993
Q_1, P_1	0.0008	0.5314	0.3692	0.3738	0.0380	0.4600	0.3806	0.3787
P_3, P_2	0	0.3916	0.7497	0.7734	0.0214	0.3851	0.5078	0.4970
P_3, P_1	0	0.5290	0.6783	0.4402	0.0273	0.4538	0.5499	0.4131
P_2, P_1	0	0.5316	0.4363	0.6360	0.0233	0.4512	0.3844	0.5373
Q_3, Q_2, Q_1	0.0005	0.3579	0.3607	0.3643	0.0495	0.3408	0.4101	0.3755
Q_3, Q_2, P_3	0.0006	0.2543	0.4702	0.4408	0.0399	0.3191	0.4825	0.4175
Q_3, Q_2, P_2	0.0011	0.2495	0.4581	0.4556	0.0391	0.3138	0.4450	0.4791
Q_3, Q_2, P_1	0.0012	0.4890	0.3875	0.3492	0.0478	0.4325	0.4496	0.4245
Q_3, Q_1, P_3	0.0009	0.3362	0.4084	0.4584	0.0563	0.3640	0.4279	0.3938
Q_3, Q_1, P_2	0.0022	0.3440	0.4337	0.4949	0.0 636	0.3826	0.4339	0.4335
Q_3, Q_1, P_1	0.0019	0.5262	0.3765	0.3940	0.0595	0.4475	0.4229	0.4006
Q_3, P_3, P_2	0.008	0.3839	0.7467	0.8053	0.0437	0.3904	0.4922	0.5559
Q_3, P_3, P_1	0.0022	0.5489	0.6571	0.4596	0.0471	0.4269	0.5304	0.4139
Q_3, P_2, P_1	0.0019	0.5223	0.4624	0.6329	0.0504	0.4460	0.4466	0.5191
Q_2, Q_1, P_3	0.0002	0.3203	0.5232	0.4120	0.0562	0.3620	0.4810	0.4114
Q_2, Q_1, P_2	0.0023	0.3024	0.4881	0.4317	0.0428	0.3454	0.4023	0.4029
Q_2, Q_1, P_1	0.0008	0.5207	0.4258	0.3724	0.0459	0.4326	0.4087	0.3783
Q_2, P_3, P_2	0.0017	0.3841	0.7955	0.7593	0.0351	0.3987	0.5740	0.4739
Q_2, P_3, P_1	0.0019	0.5216	0.6775	0.4474	0.0412	0.4504	0.5577	0.4351
Q_2, P_2, P_1	0.0025	0.5575	0.4769	0.5941	0.0345	0.4255	0.3987	0.5321
Q_1, P_3, P_2	0.0062	0.4360	0.7451	0.7617	0.0688	0.3831	0.49.5	0.4642
Q_1, P_3, P_1	0.0009	0.5485	0.6226	0.4808	0.0512	0.4633	0.5087	0.4274
Q_1, P_2, P_1	0.0028	0.5685	0.4849	0.5934	0.0504	0.4601	0.4113	0.4721
P_3, P_2, P_1	0.0073	0.5758	0.7881	0.7807	0.0446	0.4701	0.5600	0.5463
Q_3, Q_2, Q_1, P_3	0.0015	0.3174	0.4989	0.4351	0.0707	0.3560	0.4904	0.4194
Q_3, Q_2, Q_1, P_2	0.0036	0.3175	40.745	0.4772	0.0732	0.3791	0.4317	0.4540
Q_3, Q_2, Q_1, P_1	0.0026	0.5258	0.4188	0.3722	0.0695	0.4349	0.4497	0.4137
Q_3, Q_2, P_3, P_2	0.0020	0.4527	0.8360	0.8130	0.0605	0.4369	0.5723	0.5396
Q_3, Q_2, P_3, P_1	0.0040	0.5363	0.6626	0.4588	0.0631	0.4348	0.5480	0.4328
Q_3, Q_2, P_2, P_1	0.0047	0.5557	0.4903	0.6001	0.0611	0.4370	0.4575	0.5222
Q_3, Q_1, P_3, P_2	0.0077	0.4389	0.7437	0.7849	0.0928	0.4291	0.5124	0.5121
Q_3, Q_1, P_3, P_1	0.0034	0.5738	0.5974	0.5064	0.0730	0.4491	0.5049	0.4180

Input combination	LR				SVM			
	Healthy	Aorta	Iliac 1	Iliac 2	Healthy	Aorta	Iliac 1	Iliac 2
Q_3, Q_1, P_2, P_1	0.0051	0.5775	0.4908	0.6088	0.0753	0.4499	0.4550	0.4823
Q_3, P_3, P_2, P_1	0.0172	0.5931	0.7889	0.8087	0.0820	0.4699	0.5977	0.5829
Q_2, Q_1, P_3, P_2	0.0075	0.4352	0.7895	0.7508	0.0847	0.4178	0.5586	0.4785
Q_2, Q_1, P_3, P_1	0.0023	0.5479	0.6351	0.4803	0.0714	0.46.64	0.5150	0.4425
Q_2, Q_1, P_2, P_1	0.0056	0.5763	0.5133	0.5667	0.0574	0.4596	0.4104	0.4718
Q_2, P_3, P_2, P_1	0.0177	0.5735	0.8184	0.7726	0.0642	0.4950	0.6024	0.5674
Q_1, P_3, P_2, P_1	0.0163	0.6004	0.7807	0.7739	0.0789	0.4713	0.5192	0.4961
Q_3, Q_2, Q_1, P_3, P_2	0.0140	0.4922	0.8197	0.8103	0.1042	0.4311	0.5588	0.5184
Q_3, Q_2, Q_1, P_3, P_1	0.0057	0.5679	0.6061	0.5083	0.0889	0.4473	0.5300	0.4377
Q_3, Q_2, Q_1, P_2, P_1	0.0078	0.5787	0.5183	0.5772	0.0825	0.4463	0.4495	0.4941
Q_3, Q_2, P_3, P_2, P_1	0.0433	0.6279	0.8592	0.8329	0.0980	0.4961	0.6158	0.5846
Q_3, Q_1, P_3, P_2, P_1	0.0276	0.6073	0.7766	0.8036	0.1144	0.4805	0.5729	0.5383
Q_2, Q_1, P_3, P_2, P_1	0.0245	0.6060	0.8141	0.7700	0.1082	0.5002	0.5823	0.5234
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.0563	0.6421	0.8459	0.8248	0.1206	0.4925	0.5839	0.5495

Table E.2: The sensitivities achieved across the multiclass combination search on the stenosis VPD by the LR and SVM classification methods.

Input combination	LR				SVM			
	Healthy	Aorta	Iliac 1	Iliac 2	Healthy	Aorta	Iliac 1	Iliac 2
Q_3	0.9995	0.6787	0.6095	0.7217	0.9889	0.7743	0.5706	0.6809
Q_2	0.9997	0.7456	0.5986	0.6605	0.9915	0.7734	0.6873	0.5632
Q_1	1.0	0.6559	0.6670	0.6793	0.9891	0.7310	0.6254	0.6625
P_3	1.0	0.7742	0.5995	0.6518	0.9877	0.7568	0.5965	0.6836
P_2	0.9997	0.7699	0.6643	0.5963	0.9881	0.7430	0.6632	0.6230
P_1	0.9995	0.6838	0.6860	0.6622	0.9863	0.7858	0.6609	0.6002
Q_3, Q_2	0.9992	0.7273	0.6174	0.6651	0.9779	0.7580	0.6392	0.6620
Q_3, Q_1	0.9983	0.6677	0.6740	0.6754	0.9673	0.7229	0.6181	0.7179
Q_3, P_3	0.9997	0.7541	0.6091	0.6628	0.9738	0.7400	0.6111	0.6997
Q_3, P_2	0.9995	0.7624	0.6886	0.5855	0.9734	0.7175	0.6971	0.6600
Q_3, P_1	0.9980	0.7011	0.6445	0.6984	0.9724	0.8047	0.6219	0.6571
Q_2, Q_1	0.9991	0.6951	0.6463	0.6718	0.9768	0.7435	0.6494	0.6527
Q_2, P_3	0.9995	0.7823	0.5790	0.6738	0.9777	0.7391	0.6155	0.7149
Q_2, P_2	0.9998	0.7696	0.6529	0.6021	0.9808	0.7585	0.6821	0.6035
Q_2, P_1	0.9994	0.7121	0.6773	0.6472	0.9783	0.8135	0.6639	0.5992
Q_1, P_3	0.9997	0.7597	0.6072	0.6754	0.9559	0.7583	0.6095	0.7225
Q_1, P_2	0.9985	0.7483	0.6808	0.6217	0.9621	0.7503	0.7032	0.6278
Q_1, P_1	0.9986	0.7684	0.6477	0.6391	0.9618	0.8383	0.6272	0.6232
P_3, P_2	0.9998	0.7476	0.7223	0.7155	0.9799	0.6828	0.7130	0.7080
P_3, P_1	0.9989	0.8353	0.4980	0.7970	0.9754	0.8386	0.5356	0.7402
P_2, P_1	0.9994	0.8279	0.7906	0.5025	0.9762	0.8426	0.7492	0.5107
Q_3, Q_2, Q_1	0.9980	0.6910	0.6562	0.6709	0.9563	0.7374	0.6469	0.6972
Q_3, Q_2, P_3	0.9994	0.7662	0.5962	0.6718	0.9645	0.7281	0.6381	0.7234
Q_3, Q_2, P_2	0.9992	0.7686	0.6677	0.5978	0.9655	0.7208	0.7275	0.6442
Q_3, Q_2, P_1	0.9991	0.7248	0.6477	0.6735	0.9622	0.8007	0.6504	0.6613
Q_3, Q_1, P_3	0.9986	0.7468	0.6337	0.6620	0.9488	0.7456	0.6255	0.7308
Q_3, Q_1, P_2	0.9971	0.7512	0.7226	0.5841	0.9446	0.7402	0.7320	0.6495
Q_3, Q_1, P_1	0.9965	0.7724	0.6422	0.6472	0.9513	0.8293	0.6174	0.6720
Q_3, P_3, P_2	0.9991	0.7453	0.7422	0.7033	0.9651	0.6688	0.7662	0.7012
Q_3, P_3, P_1	0.9974	0.8339	0.5052	0.7971	0.9613	0.8262	0.5601	0.7395
Q_3, P_2, P_1	0.9989	0.8284	0.7756	0.5213	0.9604	0.8198	0.7328	0.5838
Q_2, Q_1, P_3	0.9983	0.7623	0.5655	0.7250	0.9492	0.7542	0.6149	0.7464
Q_2, Q_1, P_2	0.9975	0.7564	0.6651	0.6262	0.9616	0.7532	0.7112	0.6143
Q_2, Q_1, P_1	0.9985	0.7775	0.6316	0.6554	0.9611	0.8300	0.6364	0.6277
Q_2, P_3, P_2	1.0	0.7464	0.7043	0.7407	0.9673	0.6795	0.6953	0.7564
Q_2, P_3, P_1	0.9983	0.8398	0.5021	0.7897	0.9676	0.8238	0.5807	0.7284
Q_2, P_2, P_1	0.9991	0.8309	0.7762	0.5205	0.9688	0.8354	0.7475	0.5274
Q_1, P_3, P_2	0.9962	0.7543	0.7301	0.7124	0.9453	0.7501	0.6858	0.7065
Q_1, P_3, P_1	0.9977	0.8383	0.5281	0.7661	0.9541	0.8339	0.5789	0.7252
Q_1, P_2, P_1	0.9971	0.8295	0.7628	0.5402	0.9550	0.8318	0.7249	0.5688
P_3, P_2, P_1	0.9960	0.8522	0.6898	0.6952	0.9676	0.8270	0.6803	0.6546
Q_3, Q_2, Q_1, P_3	0.9977	0.7603	0.5836	0.7093	0.9361	0.7515	0.6166	0.7664
Q_3, Q_2, Q_1, P_2	0.9965	0.7529	0.7020	0.6035	0.9381	0.7413	0.7484	0.6447
Q_3, Q_2, Q_1, P_1	0.9968	0.7720	0.6304	0.6639	0.9475	0.8238	0.6379	0.6708
Q_3, Q_2, P_3, P_2	0.9985	0.7164	0.7572	0.7517	0.9515	0.6475	0.7812	0.7481
Q_3, Q_2, P_3, P_1	0.9980	0.8358	0.5182	0.7813	0.9536	0.8141	0.5971	0.7376
Q_3, Q_2, P_2, P_1	0.9982	0.8301	0.7579	0.5451	0.9544	0.8106	0.7378	0.5991
Q_3, Q_1, P_3, P_2	0.9965	0.7503	0.7289	0.7234	0.9261	0.7236	0.7461	0.7220
Q_3, Q_1, P_3, P_1	0.9959	0.8420	0.5439	0.7543	0.9367	0.8266	0.5824	0.7472

Input combination	LR				SVM			
	Healthy	Aorta	Iliac 1	Iliac 2	Healthy	Aorta	Iliac 1	Iliac 2
Q_3, Q_1, P_2, P_1	0.9948	0.8256	0.7719	0.5440	0.9398	0.8185	0.7265	0.6137
Q_3, P_3, P_2, P_1	0.9931	0.8508	0.6987	0.7048	0.9444	0.8027	0.7093	0.7016
Q_2, Q_1, P_3, P_2	0.9954	0.7523	0.7310	0.7214	0.9349	0.7282	0.6956	0.7579
Q_2, Q_1, P_3, P_1	0.9975	0.8416	0.5245	0.7697	0.9464	0.8250	0.6017	0.7331
Q_2, Q_1, P_2, P_1	0.9958	0.8351	0.7470	0.5548	0.9558	0.8310	0.7276	0.5701
Q_2, P_3, P_2, P_1	0.9945	0.8572	0.7004	0.6908	0.9538	0.7996	0.7114	0.6889
Q_1, P_3, P_2, P_1	0.9923	0.8513	0.6969	0.6985	0.9420	0.8245	0.6733	0.6818
Q_3, Q_2, Q_1, P_3, P_2	0.9951	0.7256	0.7579	0.7550	0.9189	0.7142	0.7453	0.7543
Q_3, Q_2, Q_1, P_3, P_1	0.9959	0.8454	0.5430	0.7542	0.9295	0.8231	0.6055	0.7502
Q_3, Q_2, Q_1, P_2, P_1	0.9939	0.8355	0.7486	0.5590	0.9308	0.8167	0.7351	0.6172
Q_3, Q_2, P_3, P_2, P_1	0.9908	0.8475	0.7273	0.7226	0.9364	0.7797	0.7454	0.7122
Q_3, Q_1, P_3, P_2, P_1	0.9885	0.8525	0.6996	0.7104	0.9184	0.8049	0.7127	0.7189
Q_2, Q_1, P_3, P_2, P_1	0.9908	0.8605	0.7108	0.6884	0.9305	0.8061	0.6991	0.7231
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.9863	0.8513	0.7295	0.7255	0.9159	0.7977	0.7256	0.7254

Table E.3: The specificities achieved across the multiclass combination search on the stenosis VPD by the LR and SVM classification methods.

Appendix F

Mesh independence study

To allow for fair comparison of the two implementations, a mesh independence study is completed for each. For both the DCG and SDC implementations the following four steps are carried out:

1. The physics based model is solved for the healthy aortic bifurcation network using a very fine temporal and spatial discretisation to obtain a “ground truth” solution. The SDC implementation being used assumes that vessel properties, and so consequentially pressure and flow-rate profiles, vary linearly across the length of each element. A very fine spatial discretisation of 1×10^{-3} mm is, therefore, applied when obtaining the ground truth solutions for the SDC implementation. As the DCG implementation is capable of applying non-linear profiles to vessel properties, and so consequentially pressure and flow-rate profiles, across the length of each element a coarser spatial discretisation of 4.3×10^{-1} mm for the aorta and 4.25×10^{-1} mm for the two common iliacs is applied. The SDC implementation ground truth solution is obtained using a temporal discretisation of 1×10^{-4} s, while the DCG implementation ground truth solution is obtained using a temporal discretisation of 1×10^{-6} s.
2. The target element size, *i.e* the targeted spatial discretisation, is repeatedly increased. For each target element size the physics based model is solved, and the discrepancies between the ground truth solution and the current solution is computed. The discrepancies between the ground truth and current solutions are computed in the aorta and the right common iliac. Computation of the discrepancies in the left common iliac is omitted as the right and left common iliacs, and there respective terminal boundaries, are identical and so these two vessels have identical pressure and flow-rate profiles. The discrepancies between the ground truth and current solution is quantified using a modified version of the six error metrics outlined in [58]. These metrics are the discrepancy between the mean, systolic, and diastolic pressure and flow-

rate values, and are computed using the equations:

$$E_{P,avg} = \frac{\text{mean}(\mathbf{P}_C) - \text{mean}(\mathbf{P}_G)}{\text{mean}(\mathbf{P}_G)}, \quad E_{Q,avg} = \frac{\text{mean}(\mathbf{Q}_C) - \text{mean}(\mathbf{Q}_G)}{\max(\mathbf{Q}_G)},$$

(F.1)

$$E_{P,sys} = \frac{\max(\mathbf{P}_C) - \max(\mathbf{P}_G)}{\max(\mathbf{P}_G)}, \quad E_{Q,sys} = \frac{\max(\mathbf{Q}_C) - \max(\mathbf{Q}_G)}{\max(\mathbf{Q}_G)},$$

(F.2)

$$E_{P,dias} = \frac{\min(\mathbf{P}_C) - \min(\mathbf{P}_G)}{\min(\mathbf{P}_G)}, \quad E_{Q,dias} = \frac{\min(\mathbf{Q}_C) - \min(\mathbf{Q}_G)}{\max(\mathbf{Q}_G)},$$

(F.3)

where $E_{P,avg}$, $E_{Q,avg}$, $E_{P,sys}$, $E_{Q,sys}$, $E_{P,dias}$, and $E_{Q,dias}$ represent the mean pressure and flow-rate discrepancy; the systolic pressure and flow-rate discrepancy; and the diastolic pressure and flow-rate discrepancy respectively. In the above equations \mathbf{P}_C , \mathbf{Q}_C , \mathbf{P}_G , and \mathbf{Q}_G represent the current pressure and flow-rate profile, and the ground truth pressure and flow-rate profile respectively. The error metrics used have been edited from those in [58] as the mean absolute error between ground truth and current profiles at each discrete time point can not be computed when the temporal discretisation is varied.

3. The largest target element size that can be used while still producing results with a negligible discrepancy to the ground truth solution is found. The spatial discretisation is fixed using this maximum usable target element size. The aneurysm and stenosed VPs are then solved using this maximum usable target element size, however with a mesh refinement applied to the target element size of the diseased vessel. This mesh refinement is given as a factor by which the target element size is divided. Initially the mesh refinement is set to be very high, *i.e.* there is a very small target element size for diseased vessels. As the SDC implementation is only capable of producing linear element property profiles, as outlined in step one, a spatial mesh refinement factor of 100 is

applied to the stenosed and aneurysm vessels. As the DCG implementation is capable of non-linear vessel property profiles, a spatial mesh refinement factor of 20 is applied to the stenosed and aneurysm vessels. The solutions obtained when applying these large mesh refinement factors are taken as the ground truth stenosed and aneurysm VP solutions. A similar process to step 2 is then carried out, with the mesh refinement being gradually decreased and the discrepancies at each diseased vessel target element size recorded.

4. Once a maximum target element size and minimum mesh refinement has been found that produce results with negligible discrepancy to the ground truth solutions, the temporal discretisation is optimised. A similar process to step 2 is repeated on the healthy, stenosed, and aneurysm VPs however with the time step, *i.e.* the temporal discretisation, being increased at each iteration. The largest possible time step that produces consistent results with the ground truth solution for all VPs is found.

The results produced when carrying out step one and two of the mesh independence study are shown for the SDC implementation in Figure F.1, and for the DCG implementation in Figure F.2. Figure F.1 shows that, as is expected, small target element sizes produce low discrepancies between the current and ground truth solutions. A clear point of destabilisation is seen at a target element size of approximately 0.05mm, beyond which the discrepancies increase. Despite this destabilisation, the discrepancies incurred still remain low. The SDC method being used assumes that vessel properties taper linearly across the length of each element. As the aortic bifurcation network being solved has constant radii and vessel wall mechanical properties along each vessels length, the SDC method is able to accurately capture the vessel property profiles using just a single element. This likely explains why even when using the largest target element size of 4mm, resulting in four elements per vessel due to the fact that the implementation of the SDC method being used creates an even number of elements, the percentage discrepancies all have a magnitude less than 0.004%. When solving VPs with non-linear radius profiles, *i.e.* stenosed and aneurysm VPs, large element sizes will not be able to fully capture the profile of vessel properties and so higher errors are expected. Imposing a maximum allowable discrepancy from the ground truth solutions of $1 \times 10^{-3}\%$ for all errors results in the maximum usable target element

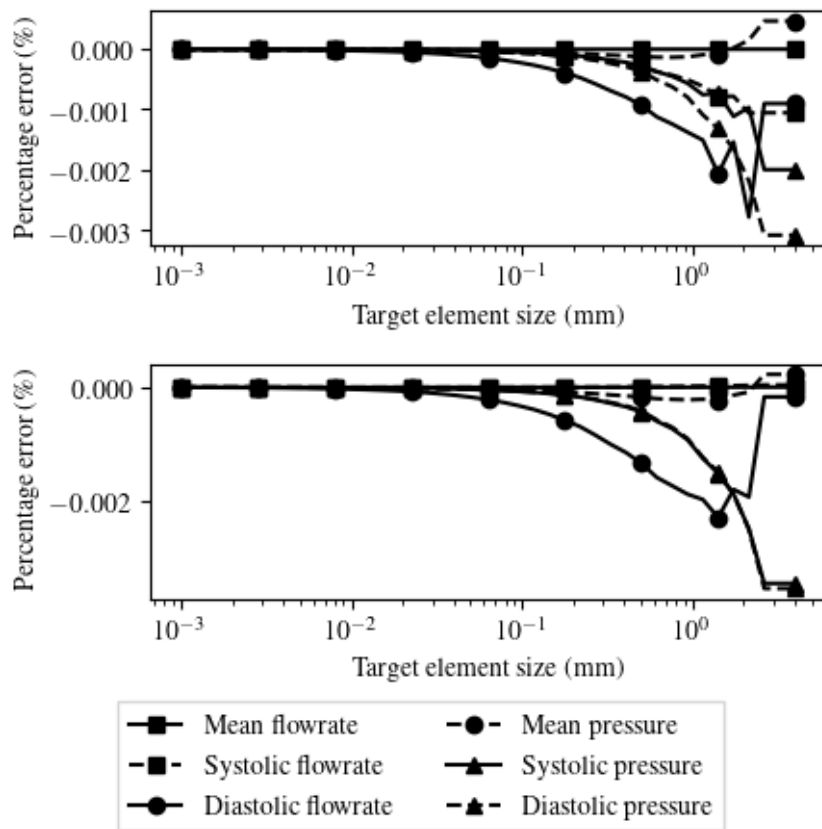


Figure F.1: The results of the second step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the spatial discretisation applied to the healthy VP is varied, are shown for the SDC implementation. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

size being 0.93mm for the SDC method.

Initially looking at Figure F.2 it appears as if the pressure and flow-rate profiles produced when using the DCG implementation are more sensitive to the spatial discretisation used. Unlike in the case of the SDC implementation, there is not a clear point of destabilisation of the discrepancies produced. Instead oscillatory behaviour can be seen across the full range of target element sizes. Looking at the ranges of discrepancies produced, however, it is seen that while the discrepancies do oscillate, their magnitudes are very low. Even when using a target element size of 4.3mm for the aorta and 4.25mm for the two common iliacs, resulting in 2 elements per vessel, the magnitude of all percentage discrepancies remain less than $5 \times 10^{-6}\%$.

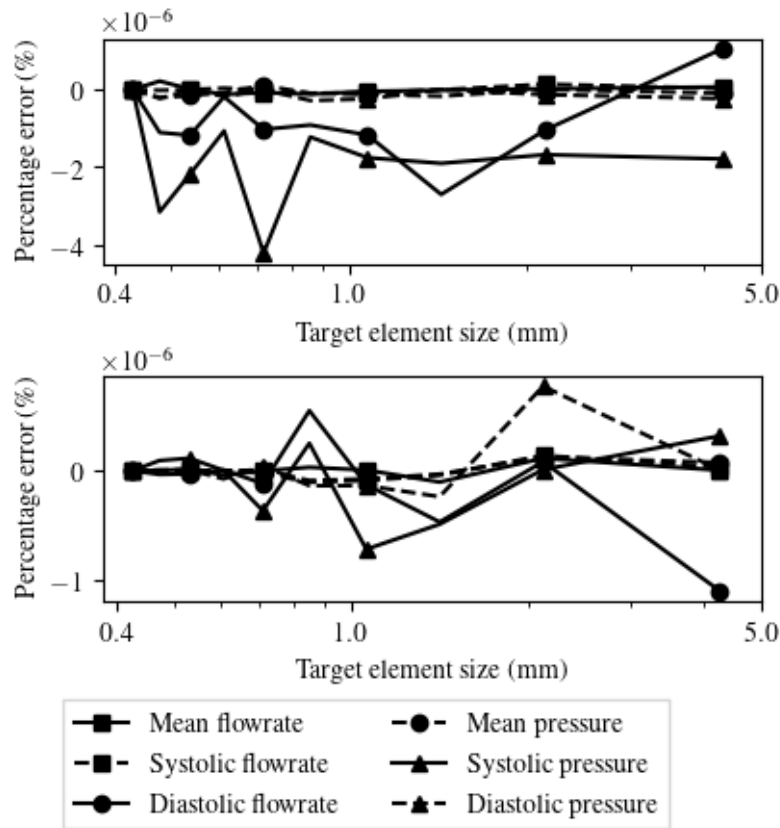


Figure F.2: The results of the second step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the spatial discretisation applied to the healthy VP is varied, are shown for the DCG implementation. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

The maximum usable target element size is, therefore, equal to 4.3mm for the aorta and 4.25mm for the two common iliacs when using the DCG implementation.

Using a target element size of 0.93mm and a temporal discretisation of 1×10^{-4} s, ground truth stenosis and aneurysm solutions are obtained using the SDC implementation by applying a mesh refinement factor of 100 to diseased vessels. The results of the iterative process of decreasing the mesh refinement factor; solving the governing system of equations to compute the pressure and flow-rate profiles using the SDC implementation; and calculating the discrepancies between the ground truth and current solutions are shown for the stenosed VP in Figure F.3 and for the aneurysm VP in Figure F.4. Figures F.3 and F.4 show that, in a similar manner

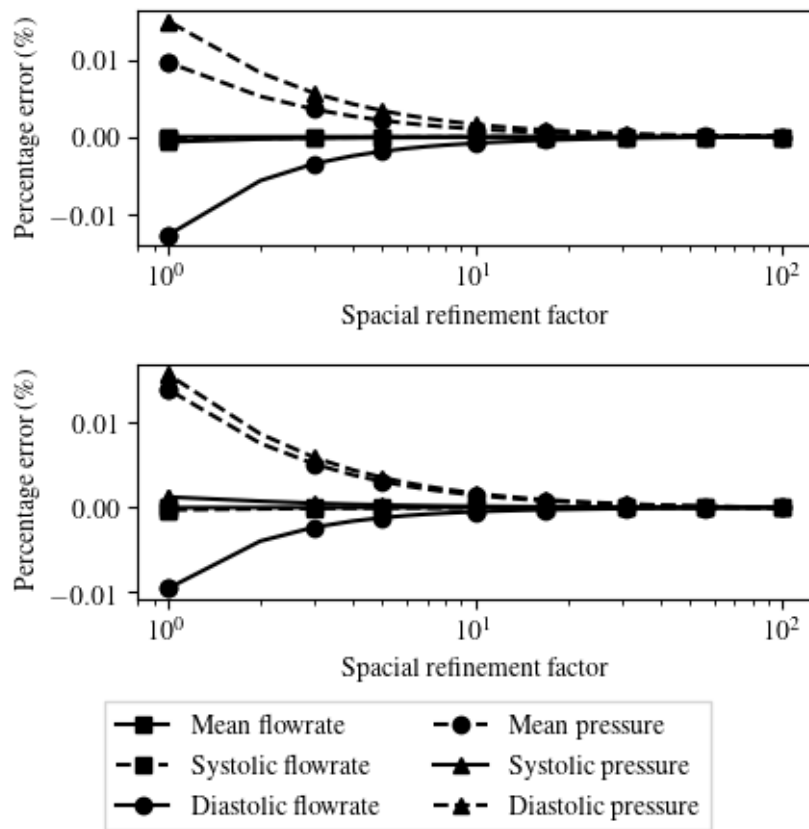


Figure F.3: The results of the third step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

to the case of the obtainment of the maximum usable element size, there are low discrepancies between the ground truth solutions and the current solutions when applying a large mesh refinement factor, *i.e.* a small target element size for diseased vessels. In the case of both diseased VPs a clear point of destabilisation is seen at a mesh refinement factor of approximately 20. The minimum mesh refinement factor, that can be applied to the target element size of diseased vessels, that results in discrepancies of less than $1 \times 10^{-3}\%$ from the ground truth stenosis and aneurysm solutions, when using the the SDC method is 17. This results in a target element size for the diseased vessels of 0.055mm when using the SDC implementation.

Ground truth stenosis and aneurysm solutions are obtained using the DCG

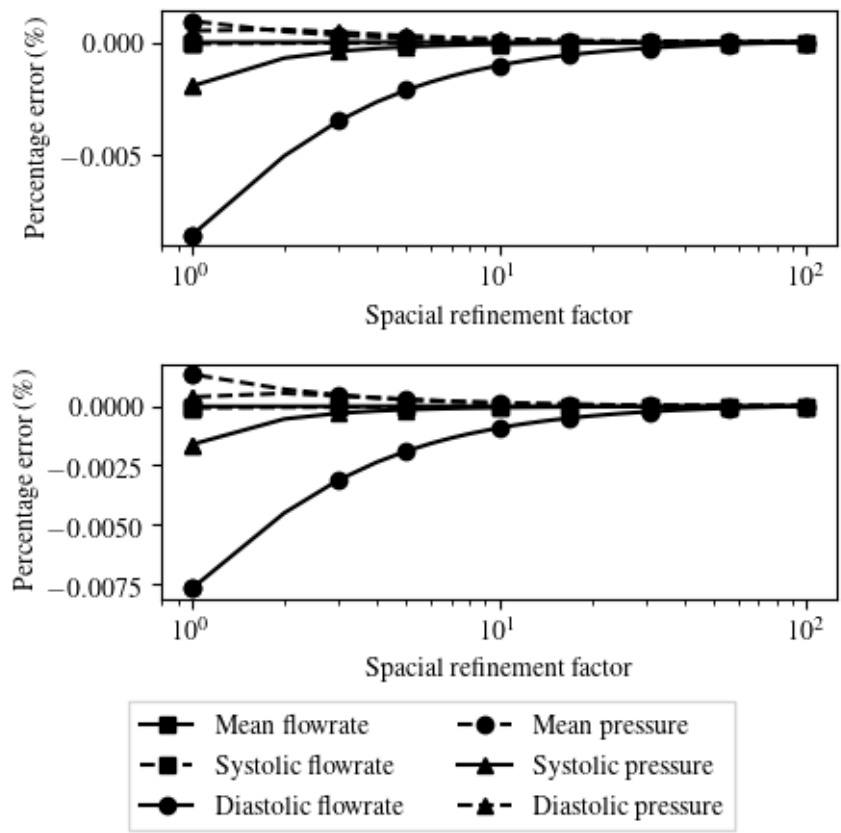


Figure F.4: The results of the third step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the aneurysms VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

implementation by applying a target element size of 4.3mm to the aorta and 4.25mm to the two common iliacs, a temporal discretisation of 1×10^{-6} s, and a mesh refinement factor for diseased vessels of 20. The results of the process of obtaining the minimum usable mesh refinement factor for diseased vessels when solving the stenosed and aneurysm VPs using the DCG implementation are shown in Figures F.5 and F.6 respectively. The pattern of behaviour seen in the discrepancies produced by the DCG implementation when varying the mesh refinement is very similar to that seen in the case of the SDC implementation. When using a large mesh refinement factor, in the case of the DCG implementation the maximum mesh refinement used is equal to 20, there is very little discrepancy between the results produced and the

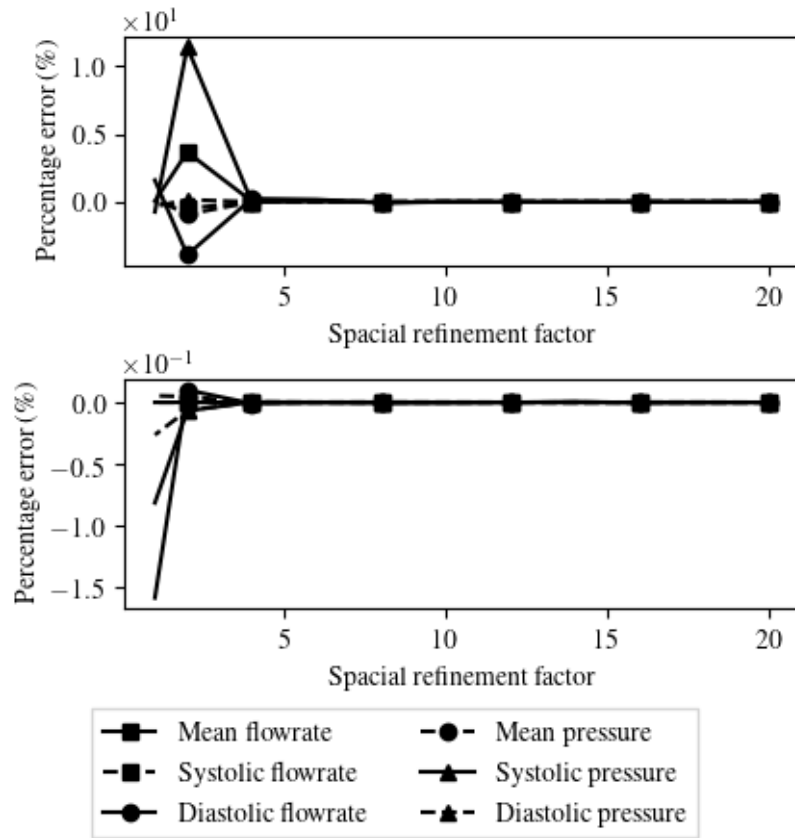


Figure F.5: The results of the third step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

ground truth solutions. A clear visual point of destabilisation of the pressure and flow-rate profiles produced occurs at a mesh refinement factor of approximately 5. This point of destabilisation is significantly lower than that seen when using the SDC method. This is likely due to the fact that the DCG method is capable of applying non-linear profiles to the properties of arterial vessels. The minimum mesh refinement factor, that can be applied to the target element size of diseased vessels, that results in discrepancies of less than $1 \times 10^{-3}\%$ from the ground truth stenosis and aneurysm solutions is 12. This results in a target element size of 0.36mm when either a stenosis or aneurysm is included within the aorta.

The results of the final step of the mesh independence study, *i.e.* the obtainment

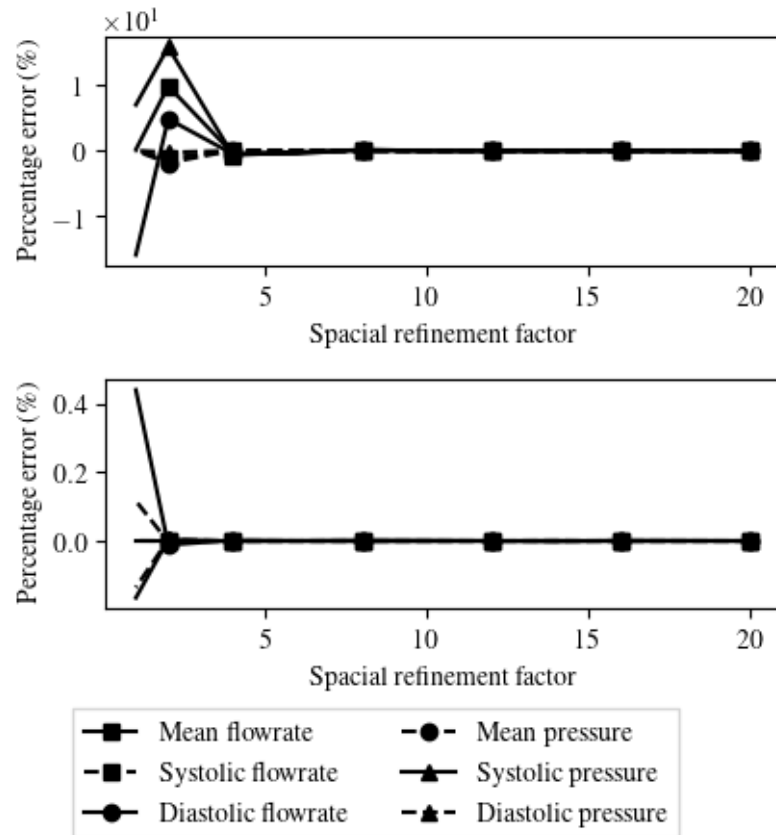


Figure F.6: The results of the third step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the spatial discretisation applied to diseased vessels is varied, are shown for the aneurysm VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

of the maximum usable temporal discretisation, are shown when using the SDC implementation for the healthy VP within Figure F.7, the stenosed VP within Figure F.8, and the aneurysm VP within Figure F.9. In a very similar manner to what is seen in the previous steps of the mesh independence study, the discrepancies produced from the ground truth solutions can be seen to be very low and stable when using a very fine temporal discretisation. A clear point of visual destabilisation, beyond which the discrepancies produced both increase in magnitude and begin to oscillate, is seen at a temporal discretisation of approximately 1×10^{-3} s for both vessels and all three VPS. The maximum usable temporal discretisation, that produces discrepancies of less than $1 \times 10^{-3}\%$ within all 3 vessels of the 3 VPs, is

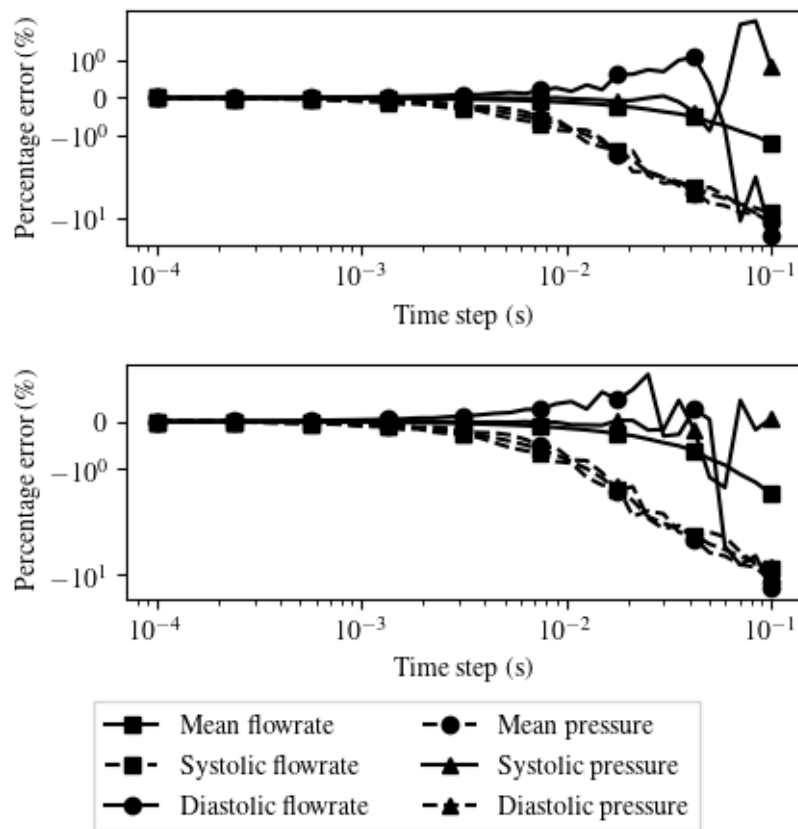


Figure F.7: The results of the fourth step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the healthy VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

found to be 1.2×10^{-4} s for the SDC implementation.

The results of the fourth step of the mesh independence study when using the DCG implementation are shown for the healthy VP in Figure F.10, the stenosed VP in Figure F.11, and the aneurysm VP in Figure F.12. The maximum temporal discretisation trialled when completed the fourth step of the mesh independence study using the DCG implementation is found to be limited to 1.33×10^{-4} s. It is found that the DCG implementation being used is unable to converge on static solutions for all three VPs when using a temporal discretisation greater than this limit. Figures F.10, F.11, and F.12 show that, unlike in the case of the SDC implementation, there is not a point of clear destabilisation beyond which the errors produced

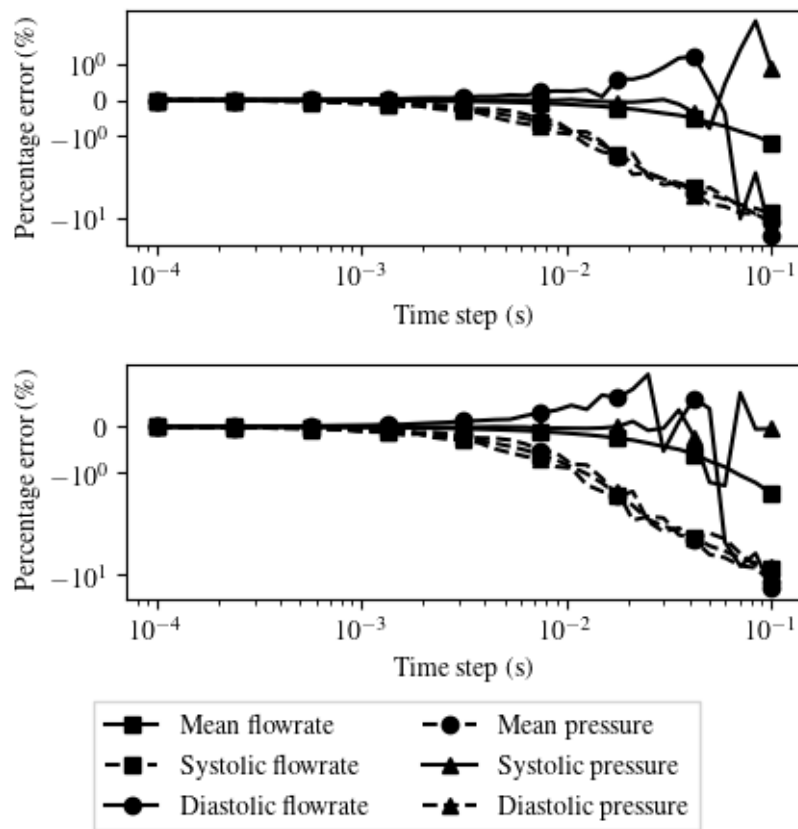


Figure F.8: The results of the fourth step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

increase. Instead the DCG implementation appears to be sensitive to the temporal discretisation used, with oscillatory behaviour occurring throughout the full range of different temporal discretisation applied. It is found that, due to the sensitivity of the temporal discretisation on the pressure and flow-rate profiles produced, any deviation from the temporal discretisation used to obtain ground truth solutions resulted in discrepancies of more than $1 \times 10^{-3}\%$.

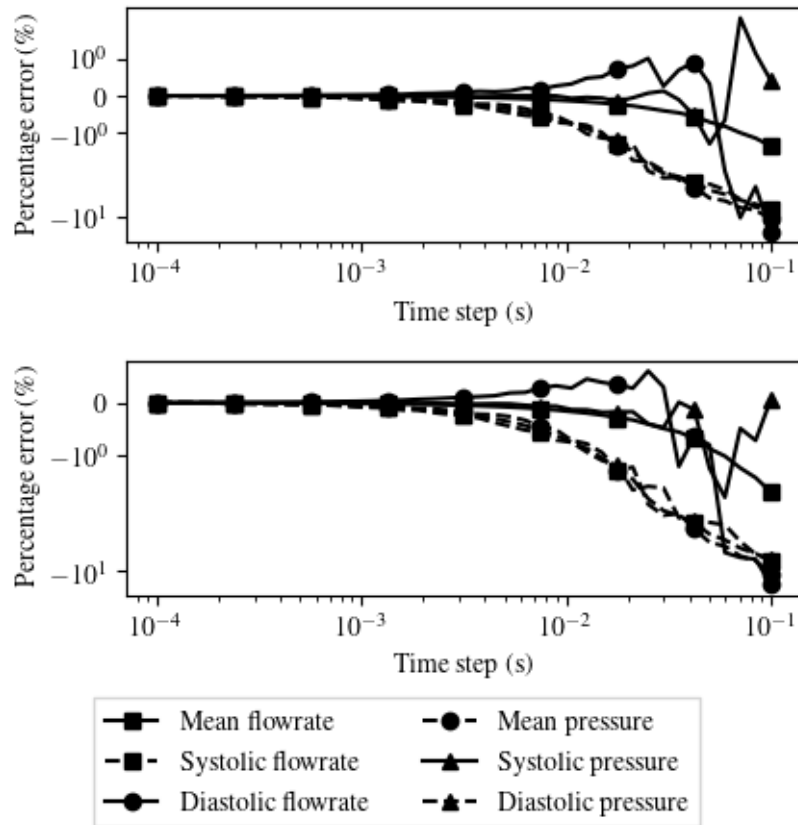


Figure F.9: The results of the fourth step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the aneurysm VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

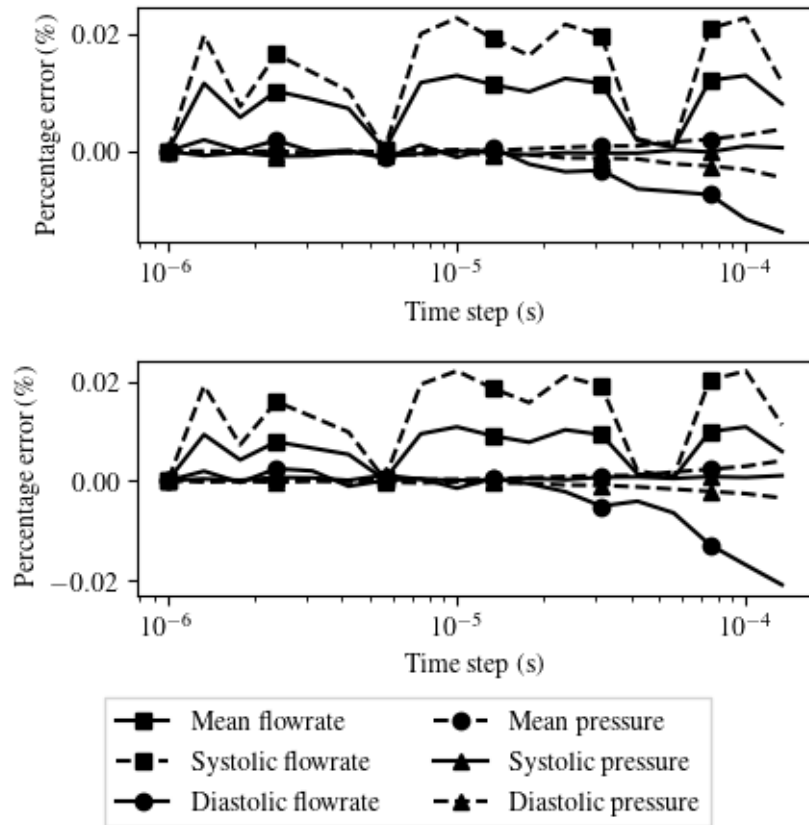


Figure F.10: The results of the fourth step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the healthy VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

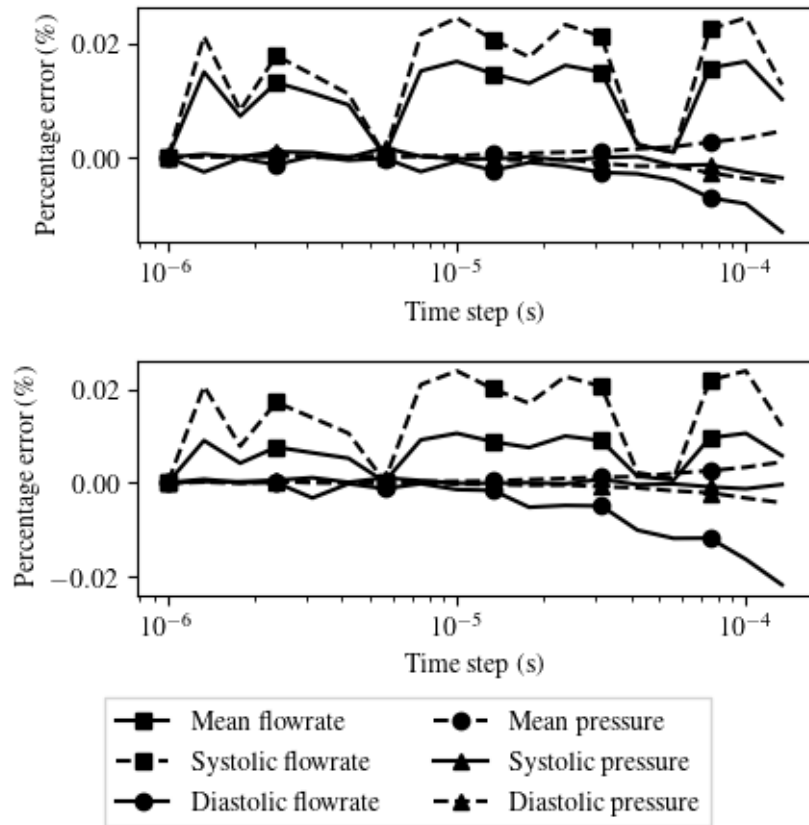


Figure F.11: The results of the fourth step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the stenosed VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

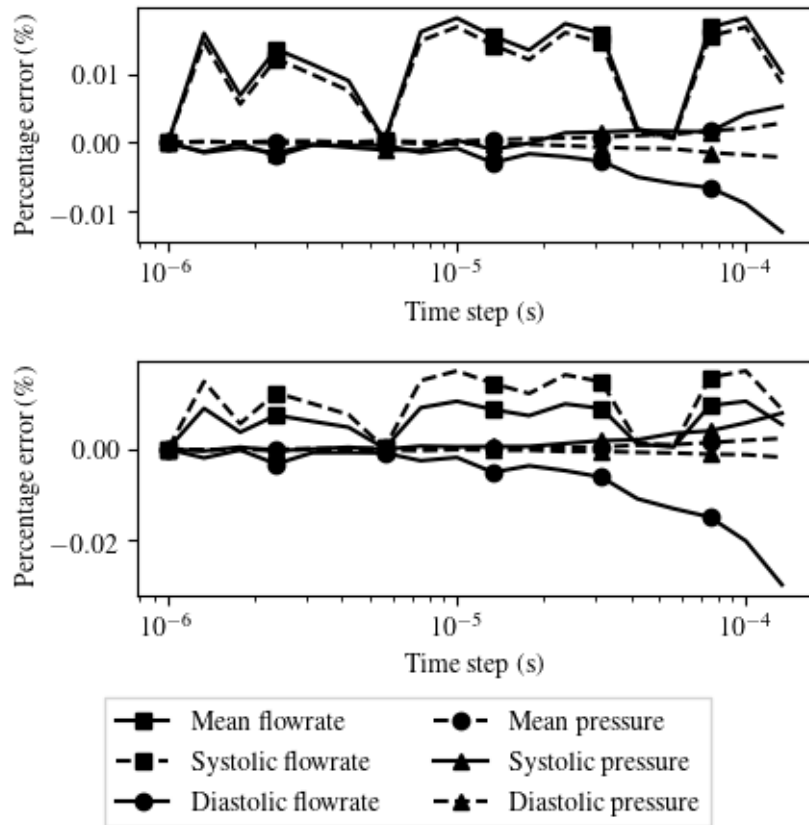


Figure F.12: The results of the fourth step of the mesh independence study, *i.e.* the pressure and flow-rate discrepancies produced when the temporal discretisation is varied, are shown for the aneurysm VP. Mean, systolic, and diastolic pressure and flow-rate discrepancies in the centre of the aorta and the right common iliac are shown in the top and bottom plots respectively.

Appendix G

Arterial vessel length

Each arterial vessel requires specification of its length. To individually assign a length to every vessel requires 71 parameters, accounting for a large proportion of the 269 remaining dimensions. It is thus proposed to reduce the dimensionality of the network by applying a singular scaling term to the lengths of all the vessels. Generating the lengths of vessels in this way removes 70 dimensions from the parameterisation, however will likely restrict the variability seen in pressure and flow-rate profiles in VPs. It is important to ensure that any efforts to reduce the complexity of VPs does not come at the cost of significant restriction or reduction to the physiological realism of the VPD.

To assess the importance of vessel length on the variability of pressure and flow-rate profiles two sets of VPs are created using the reference network outlined in Section 4.4.3. For all VPs, all parameters are set to their reference values, except for the length of each vessel. For the first set of VPs the length of each vessel is randomised by applying an individual independent scaling term to the reference length of each. The scaling term applied to the reference length of each vessel is sampled from a normal distribution with a mean of 1 and a standard deviation of 0.2. For the second set of VPs a single scaling term is applied to the reference length of all vessels within the arterial network. As with the first case, this scaling term is sampled from a normal distribution with mean of 1 and standard deviation of 0.2. For each of the two cases outlined above 30,000 VPs are sampled, and the pressure and flow-rate profiles associated with each computed.

Pressure or flow-rate profiles are taken from each VP at all non-invasive measurement locations—highlighted within Section 4.4.2—and the average, maximum, and minimum of each profile recorded. The average, maximum, and minimum pressure is also recorded at the inlet of the aorta, to compare the affect of vessel length on pressure profiles at a location with known flow-rate. The mean and standard deviation of the average, maximum and minimum pressure at each appropriate examination location is shown for patients with a constant length scaling term in Figure G.1, and individual scaling terms in Figure G.2. The mean and the standard deviation of the flow-rate measurements at the appropriate examination

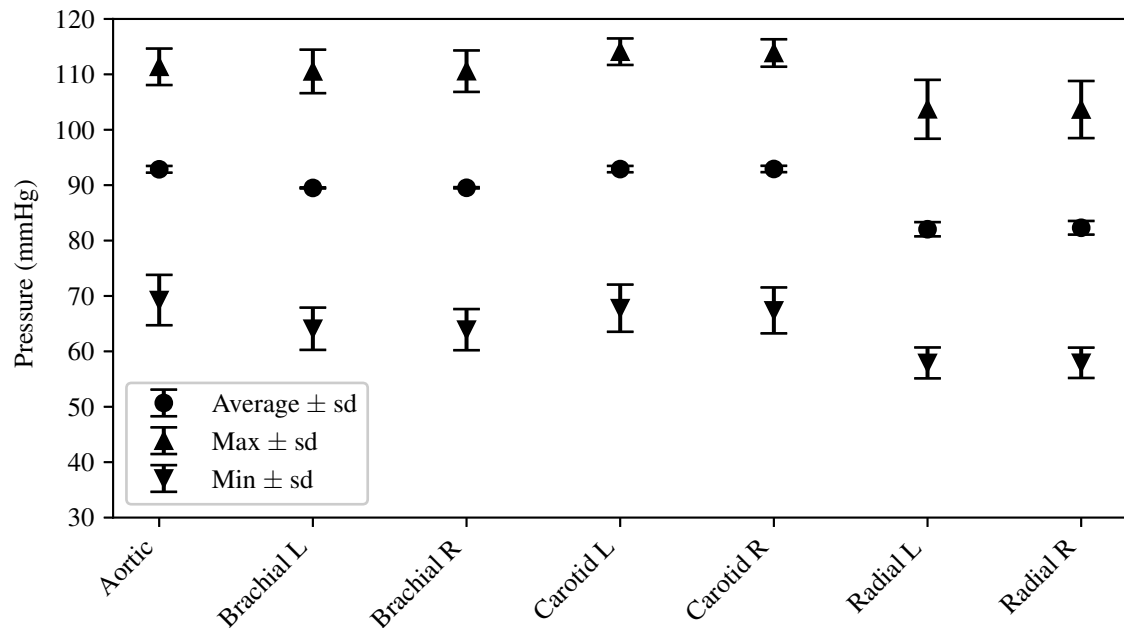


Figure G.1: The mean and standard deviation of the average, maximum, and minimum pressure at all appropriate examination locations for the 30,000 VPs created using a constant length scaling term.

locations are shown for patients produced using a constant length scaling term in Figure G.3, and individual scaling terms in Figure G.4. It is expected that the mean of the average, maximum, and minimum pressure and flow-rate measurements will be constant for both methods of creation of VPs. As the mean of the distribution used to sample vessel length scaling terms is the same for both methods, the mean arterial network—and so consequently the mean pressure and flow-rate profiles—produced using each of the two methods should be identical.

Figures G.1 and G.2 show that, as is expected, the mean of all pressure measurements are relatively consistent when using each of the two methods for generating VPs. The standard deviation of all pressure measurements is larger when applying a constant length scaling term to all vessels within the arterial network rather than individual scaling terms applied to each vessel. This increase in standard deviation is most noticeable when comparing the maximum and minimum pressure values. The standard deviation of the average pressure at each location is relatively low for both methods of VP creation. Figures G.3 and G.4 show that,

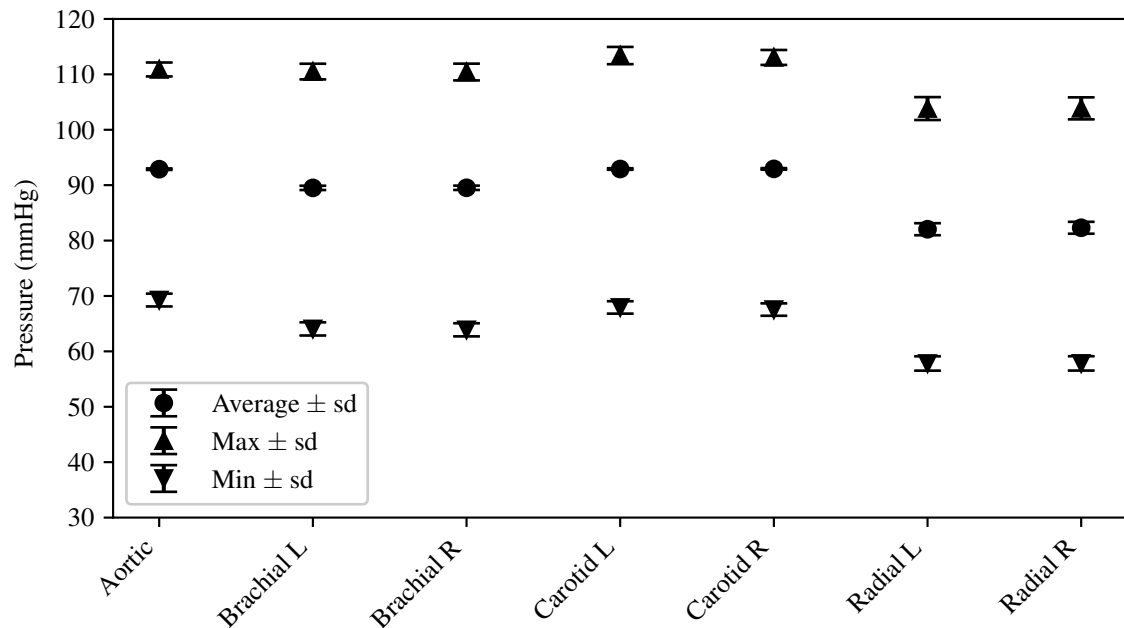


Figure G.2: The mean and standard deviation of the average, maximum, and minimum pressure at all appropriate examination locations for the 30,000 VPs created using individual length scaling terms applied to each vessel within the network.

as with mean pressure measurements and as is expected, the mean of the flow-rate measurements are relatively consistent when using each of the two sampling methods. Unexpectedly, the standard deviations of the flow-rate measurements are higher when using a constant length scaling term, rather than individual scaling terms.

From the current analysis it appears as if applying a constant scaling term to the length of all vessels within the network increases the variability seen within pressure and flow-rate profiles produced, rather than restricts it. To fully understand the affect of applying a singular scaling factor to the length of all arterial vessels, however, it is important to look at not only the mean and standard deviation of pressure and flow-rate measurements but also the correlation between different measurements. The Pearson correlation coefficient [15] between the average, maximum, and minimum pressure or flow-rate measurement at every location is computed and plotted. Due to the high number of individual correlations computed,

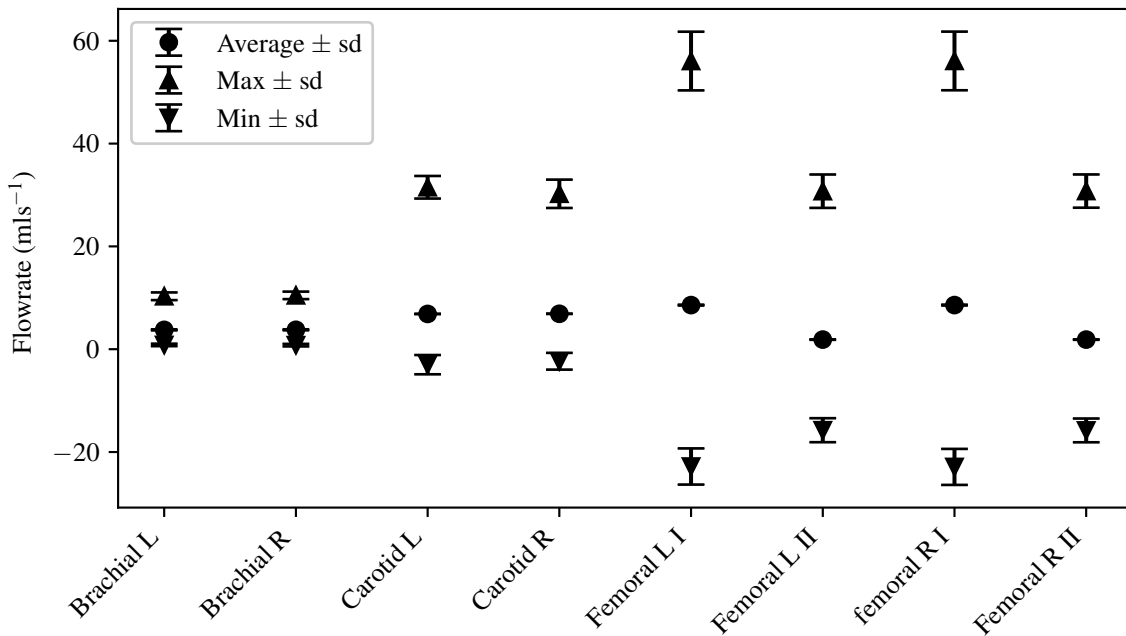


Figure G.3: The mean and standard deviation of the average, maximum, and minimum flow-rate at all appropriate examination locations for the 30,000 VPs created using a constant length scaling term.

it is not possible to show all in a single figure. Instead the pressure and flow-rate measurements are split into a grid, and a series of sub-figures produced. These sub-figures are then arranged, using the positioning shown in Figure 35, to produce a singular large corner plot of the correlation between all measurements.

The correlations between the measurements of average pressure and flow-rate at each location are shown in Figure G.6 for VPs created with a constant vessel length scaling term, and Figure G.7 for VPs created with individual vessel length scaling terms. The correlations between the maximum and average; minimum and average; maximum and maximum; minimum and maximum; and minimum and minimum pressure and flow-rate measurements at each location are shown in Figures G.8, G.10, G.12, G.14, and G.16 for VPs created using a constant length scaling term respectively, and Figures G.9, G.11, G.13, G.15, and G.17 for VPs created using an individual scaling terms respectively. It is expected that the correlation between measurements taken from VPs created using a constant scaling term applied to the length of all vessels within the network will be consistently larger in magnitude

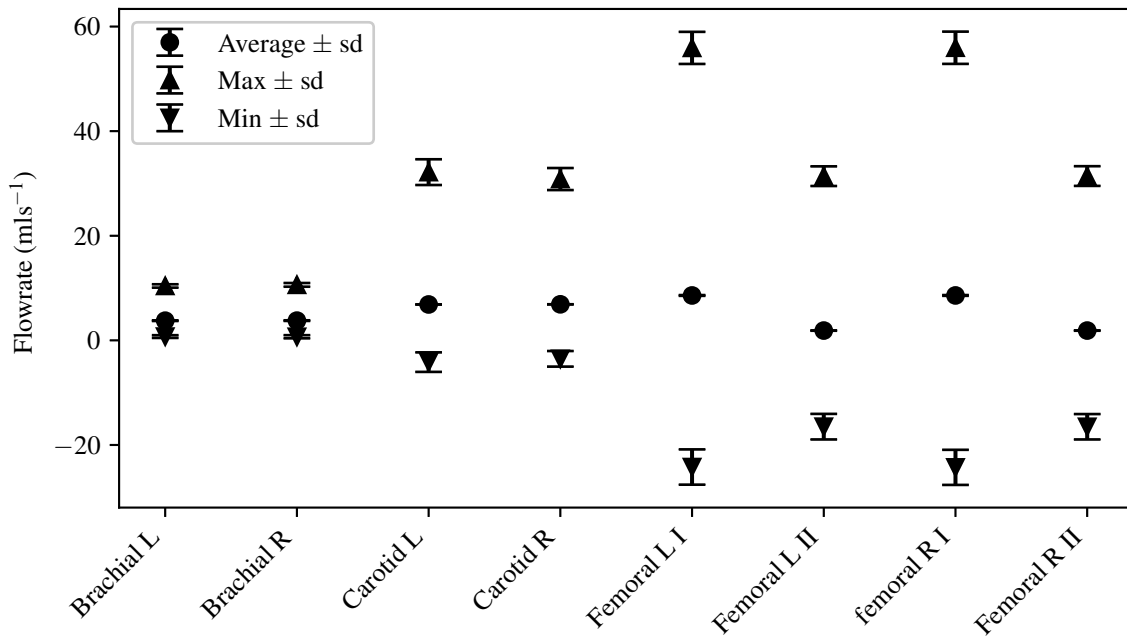


Figure G.4: The mean and standard deviation of the average, maximum, and minimum flow-rate at all appropriate examination locations for the 30,000 VPs created using individual length scaling terms applied to each vessel within the network.

than the correlation recorded between measurements taken from VPs created using individual scaling terms. The differences between the correlation recorded using each of the two methods is expected to be significant in this analysis, as the two most extreme possible cases are being compared. Even if a single scaling term is not used, it is unlikely that when creating VPs the length of each vessel will be independent of one another. Comparing independent scaling terms to a constant scaling term will magnify any differences between the two methods.

Looking at all figures listed above it can be seen that, as is expected, VPs created using a constant vessel length scaling term produce higher magnitude correlations than VPs created using individual independent length scaling terms. Certain behaviours and patterns that are seen within the correlations between measurements taken from VPs created using a constant vessel length scaling term, however, are preserved, to a lesser magnitude, within the correlations between measurements taken from VPs created using individual vessel length scaling terms. Within Figure

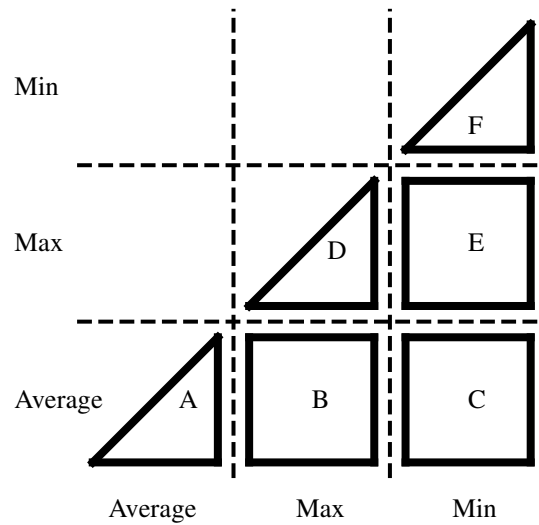


Figure G.5: Due to the high number of individual correlations, it is not possible to show all in a singular plot. Instead the pressure and flow-rate measurements have been broken down into a grid, and a series of subplots created. The location of each subplot within the larger correlation corner plot.

G.6 it is seen that there is strong negative correlations between:

1. The average flow-rate in the brachial arteries;
2. and the average pressure in the brachial arteries, and the average pressure and flow-rate in the carotid arteries.

There is a rectangle of further strong negative correlation between:

1. The average flow-rate in the four femoral artery segments;
2. and the average pressure and flow-rate in the two common carotid arteries, and the average pressure in the two brachial arteries.

A further correlation pattern seen within Figure G.6 is that there is strong negative correlation between:

1. The average pressure in the radial arteries;
2. and the average pressure and flow-rate in the two common carotid arteries, and the average pressure in the brachial arteries.

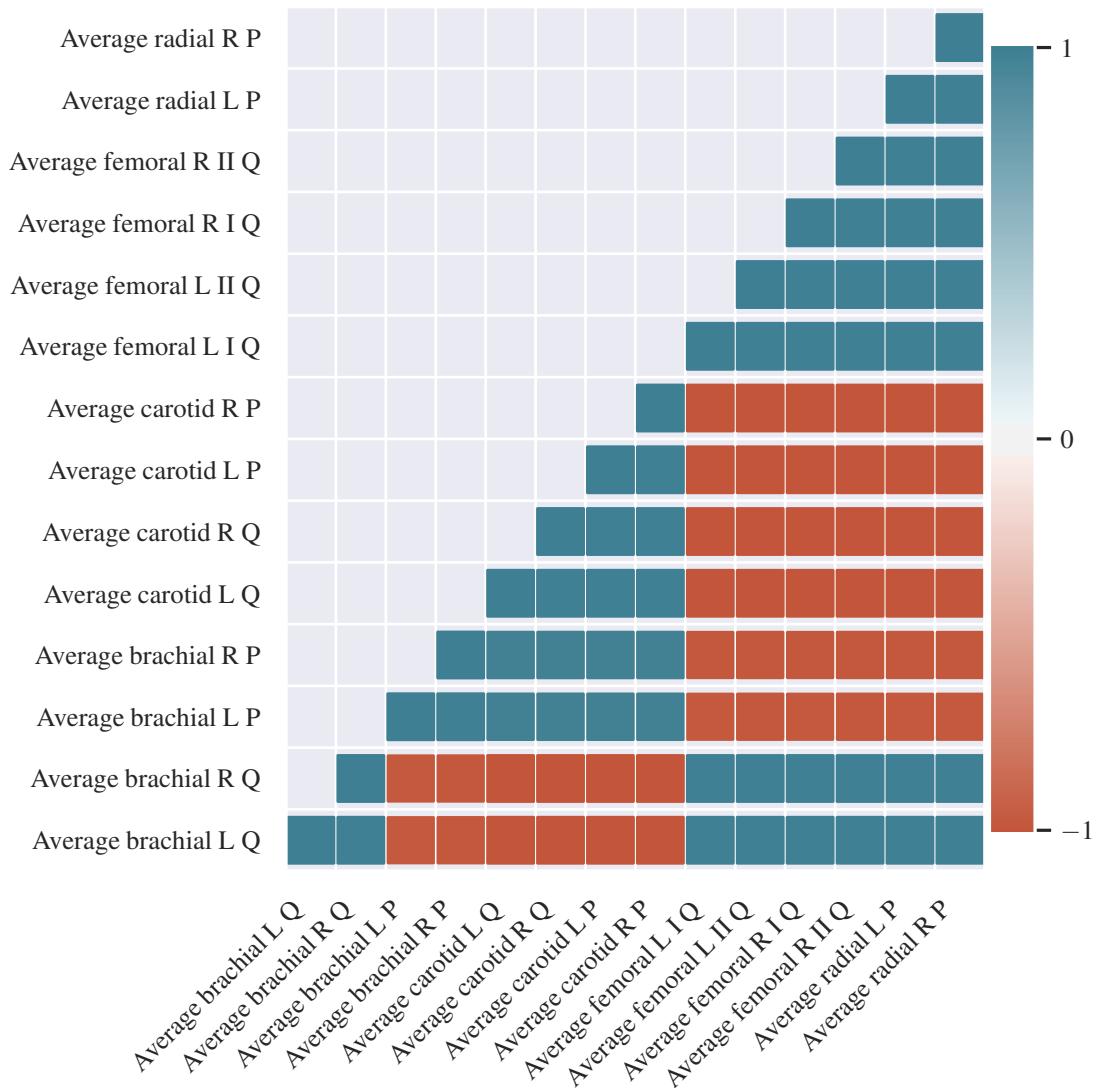


Figure G.6: The Pearson correlation coefficients between the average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.

The strong negative correlation between average brachial and average radial pressures is unexpected, as the radial artery bifurcates off of the brachial artery. These areas of low correlation can also be seen when computing the correlation between measurements taken from VPs created using individual length scaling terms, shown in Figure G.7, however to a lesser magnitude. Further patterns that are mirrored between the two methods of producing VPs are:

- The strong negative correlations shown in Figures G.8 and G.9 between:
 1. The average common carotid pressure and flow-rate, and the average brachial pressure;
 2. and the maximum of all pressure and flow-rate measurements, excluding the maximum common carotid flow-rate.

When independent scaling terms are applied to the length of each vessel, however, negative correlation can be seen between the average and maximum of flow-rate within each common carotid artery that is not present when using a constant scaling term.

- When creating VPs using a single length scaling term strong negative correlations can be seen between:
 1. The maximum common carotid flow-rate;
 2. and the maximum of all other pressure and flow-rate profiles.

This pattern is shown in Figure G.12. While these correlations are not present when using independent length scaling terms applied to each vessel individually, shown in Figure G.13, there is a lack of positive correlation at these locations.

- Within Figure G.14, strong positive correlations can be seen between:
 1. The maximum flow-rate in the right and left common carotid arteries;
 2. and the minimum pressure in the brachial arteries, the minimum pressure in the common carotid arteries, and the minimum pressure in the radial arteries.

As with the correlation patterns listed above, these correlation patterns can not be directly seen when using individual vessel length scaling terms. Looking at Figure G.15, however, it can be seen that there is a lack of negative correlation at these locations.

- High correlation can be seen within Figure G.16 between:
 1. The minimum brachial pressure;
 2. and the minimum common carotid pressure, and minimum radial pressure.

These areas of high correlation can be directly seen when generating patients using individual vessel length scaling terms, shown in Figure G.17.

While it can be seen from this analysis that using a constant length scaling term does increase the correlation between measurements, it is important to again remember that this analysis is being carried out in the most extreme case possible. It is highly likely, due to the indirect affect of vessel length on pressure and flow-rate profiles, that the high correlations between measurements seen when using a constant length scaling term will be mitigated by the variability introduced through the randomisation of all other parameters describing a VP's arterial network. This analysis has shown that behaviours and patterns seen in pressure and flow-rate profiles when allowing maximum freedom to the length of vessels is not lost when applying a singular vessel length scaling term. It is therefore decided that, to reduce the dimensionality associated with the generation of VPs, the length of all vessels within a VP's arterial network is scaled by applying a singular term to the reference length of all vessels.

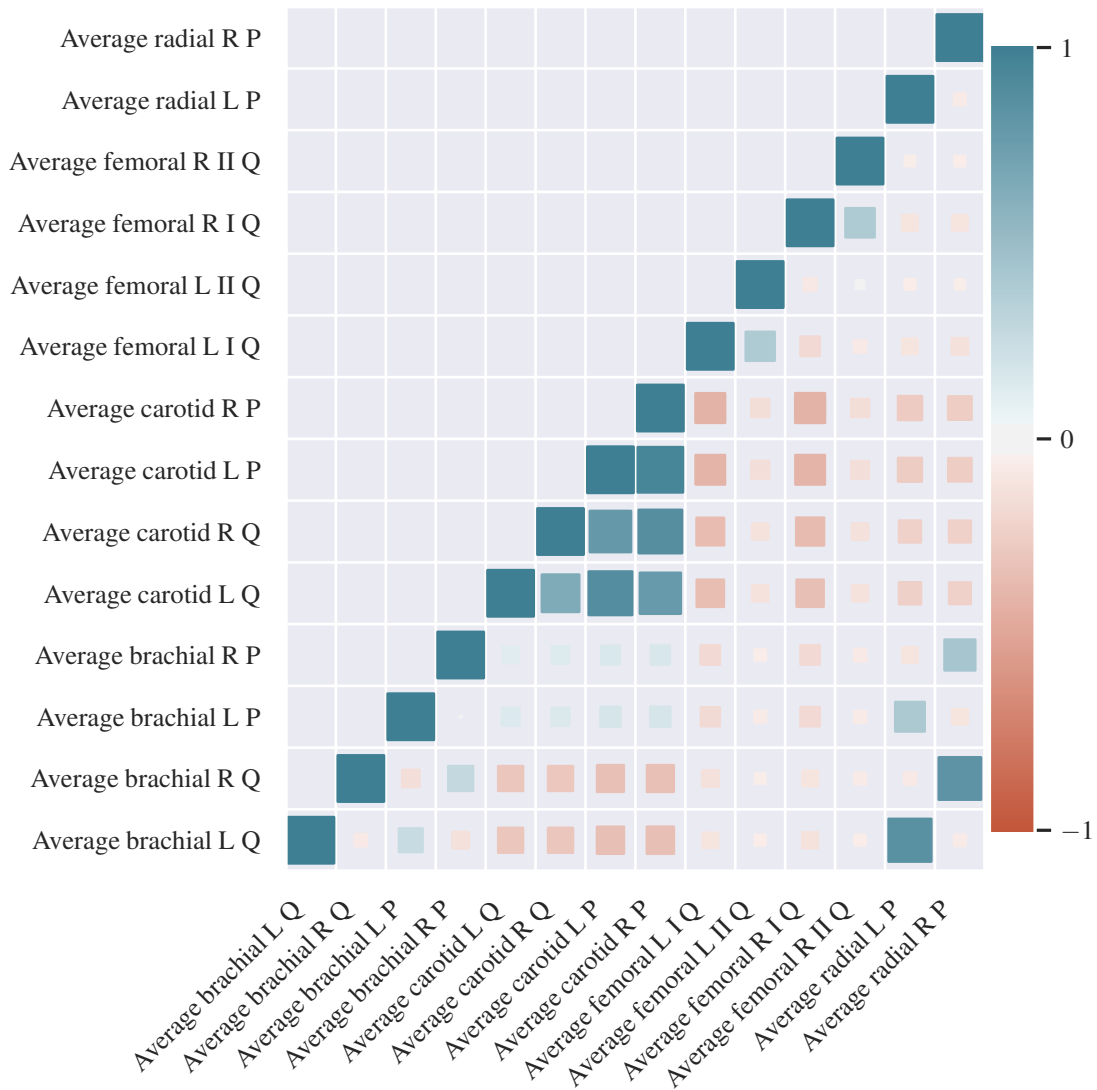


Figure G.7: The Pearson correlation coefficients between the average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using an individual scaling factors applied to the length of each vessel within the network.

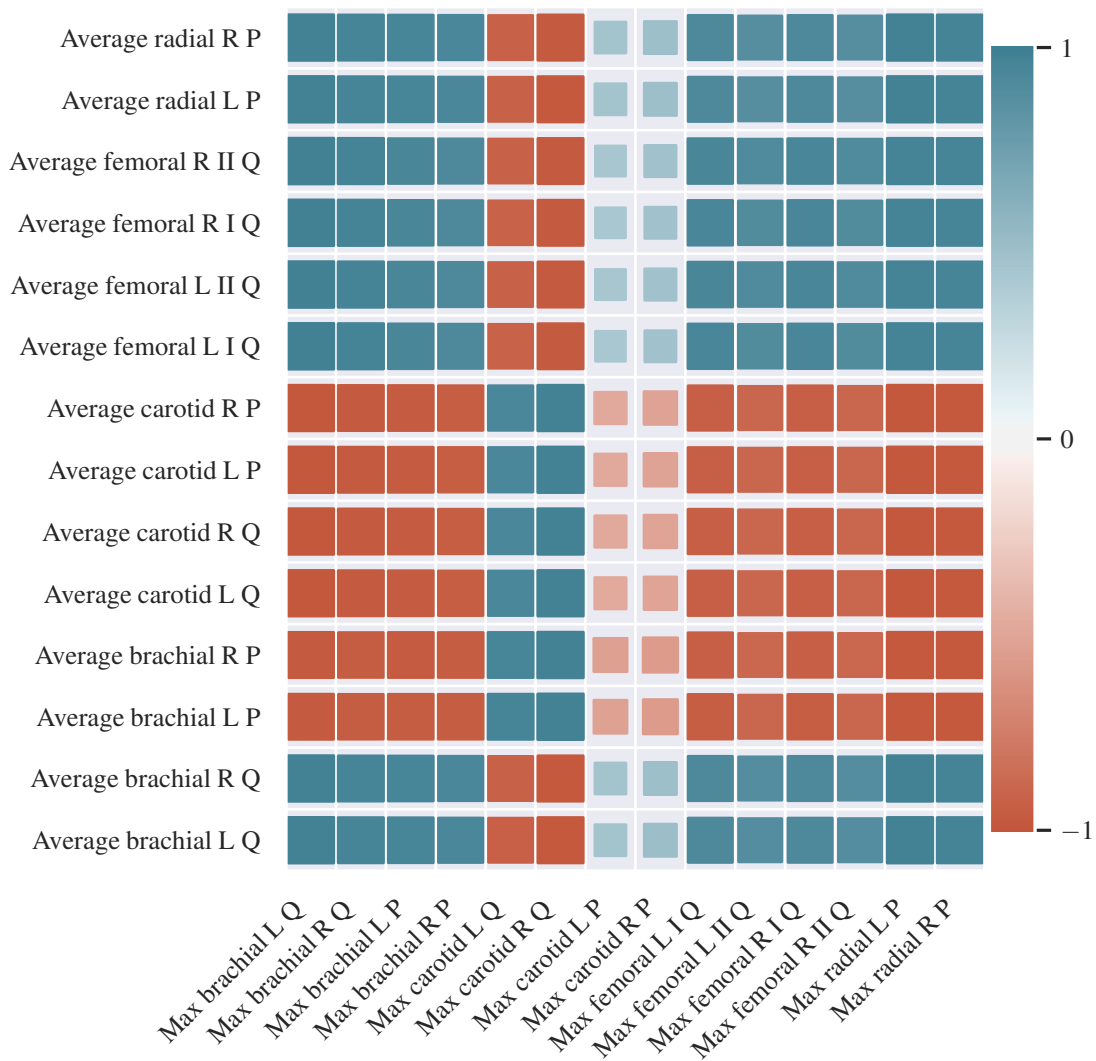


Figure G.8: The Pearson correlation coefficients between the maximum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.

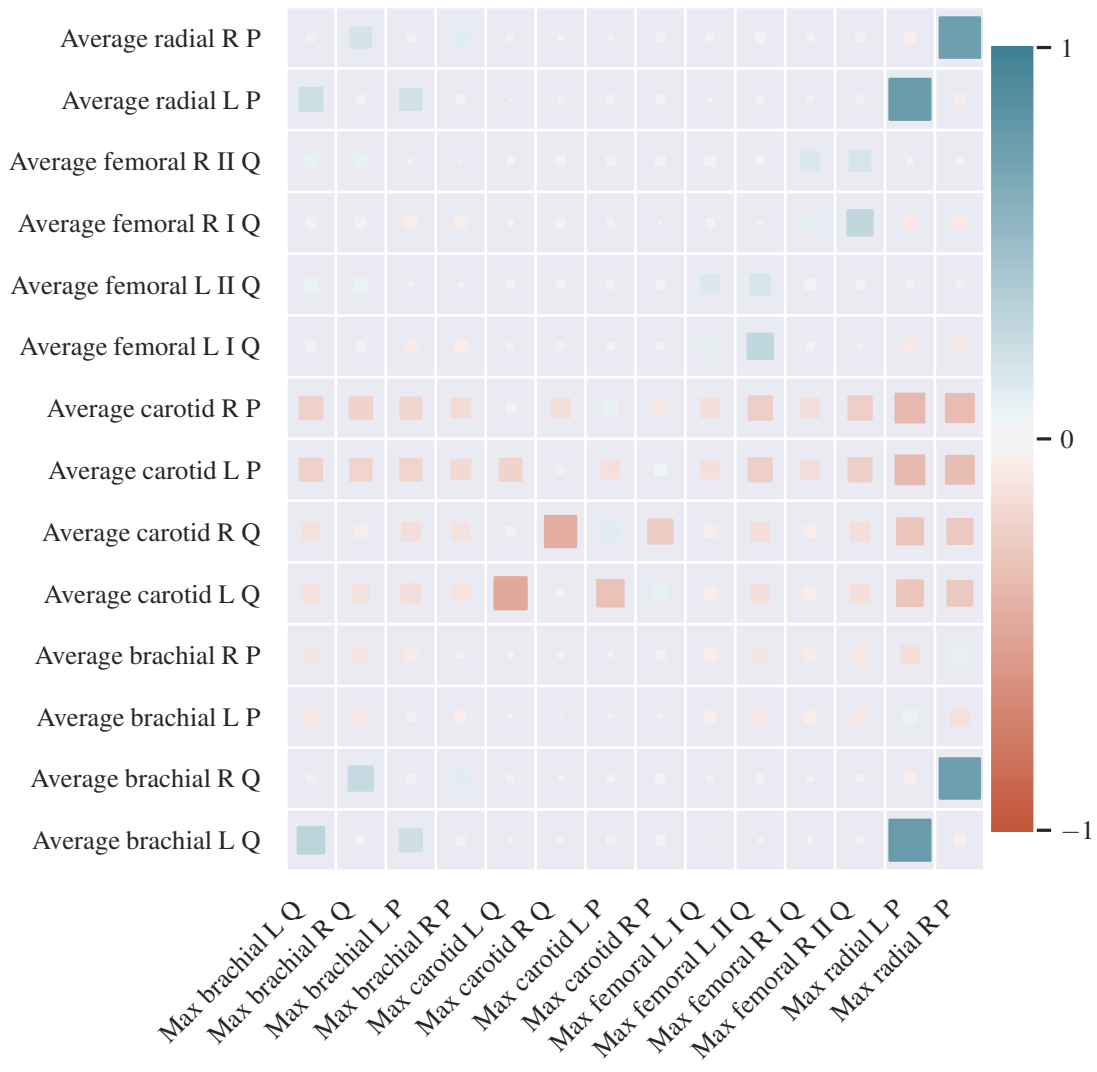


Figure G.9: The Pearson correlation coefficients between the maximum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.

Sub-figure C.1: Min-Average correlation using a constant scaling factor.

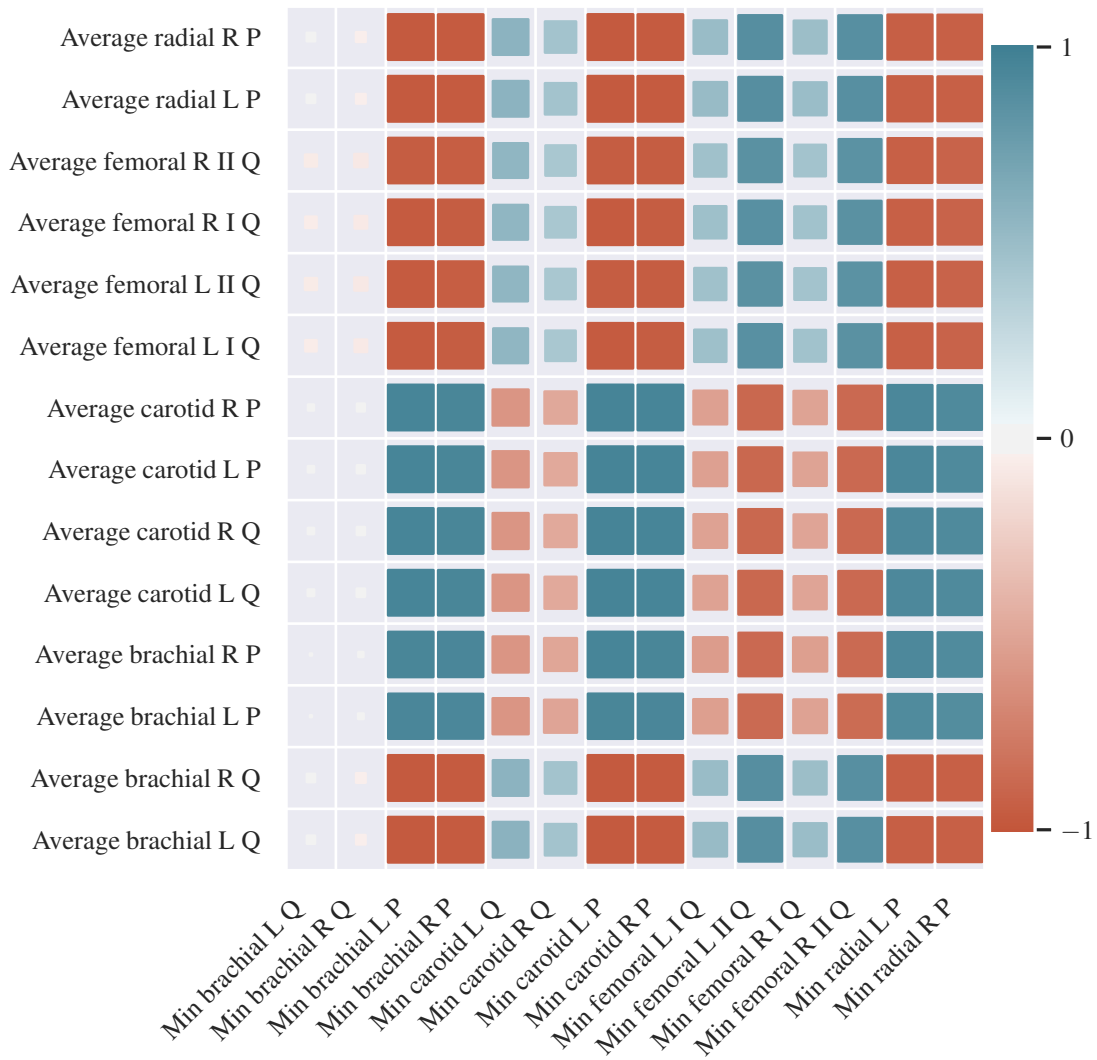


Figure G.10: The Pearson correlation coefficients between the minimum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.

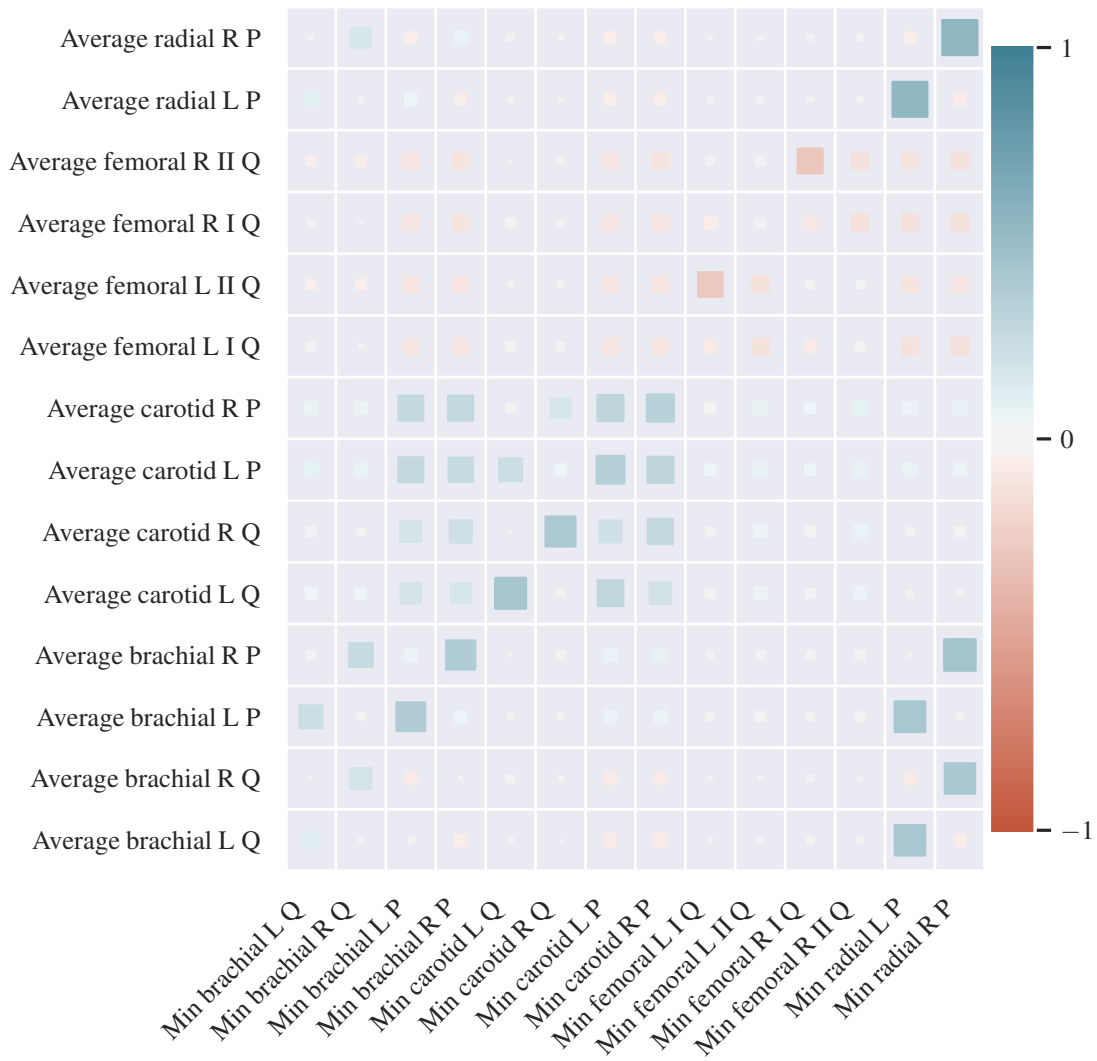


Figure G.11: The Pearson correlation coefficients between the minimum and average of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.

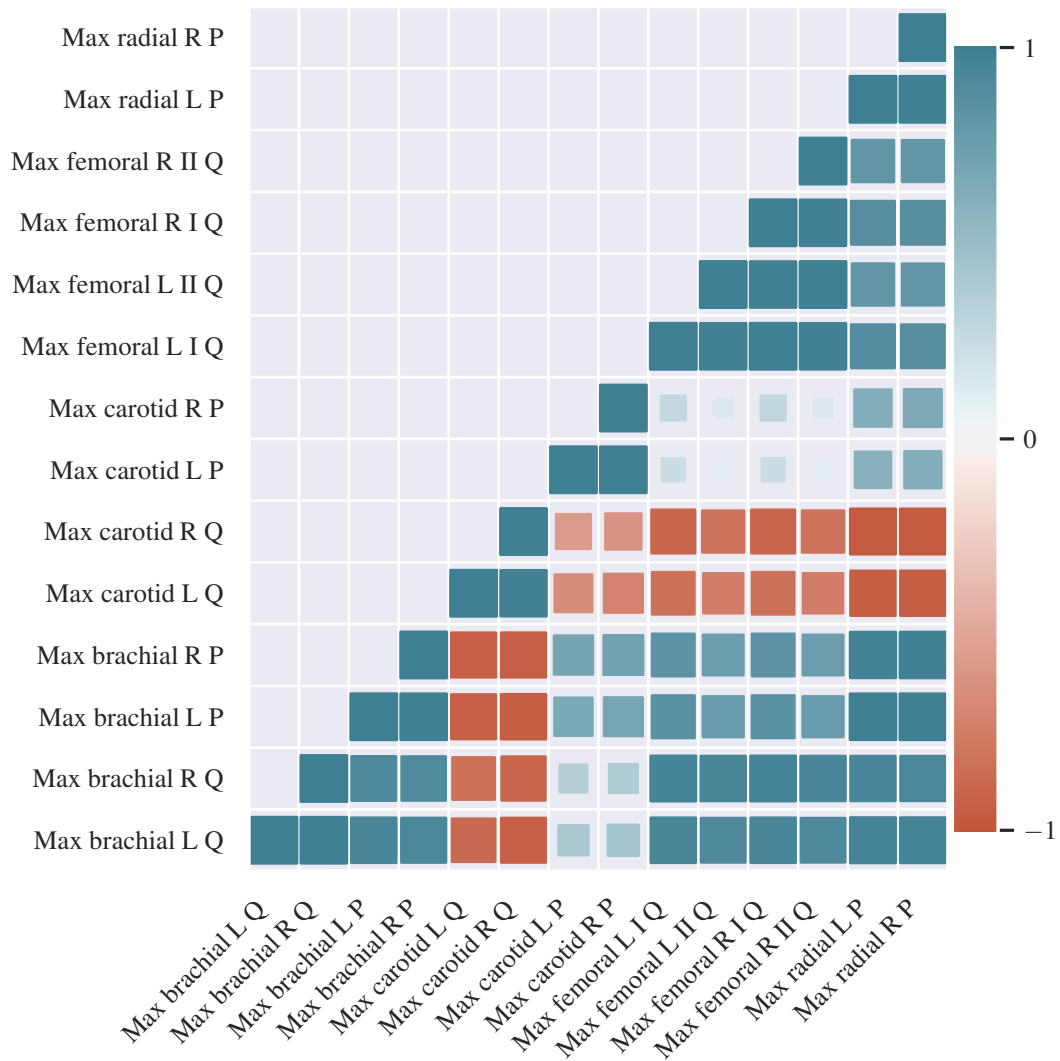


Figure G.12: The Pearson correlation coefficients between the maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.

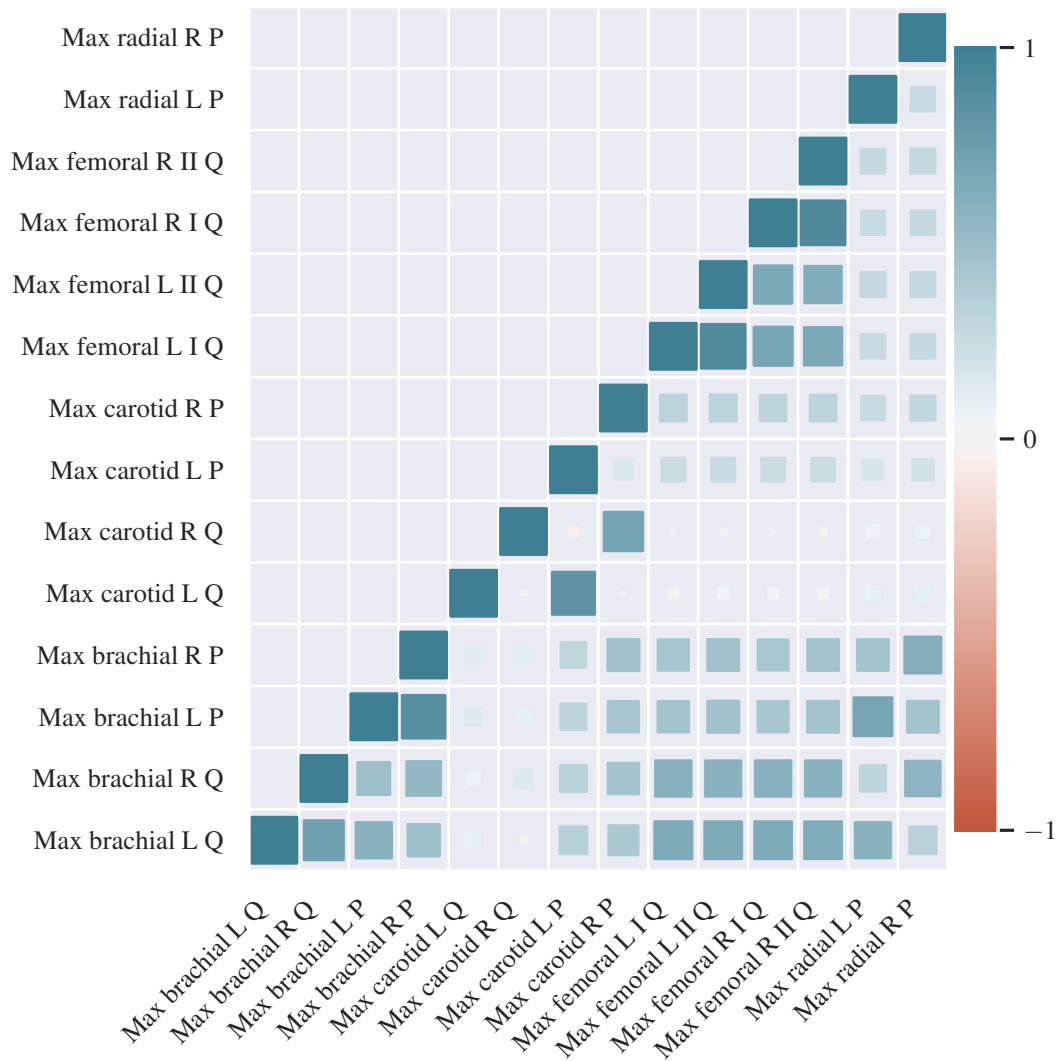


Figure G.13: The Pearson correlation coefficients between the maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network

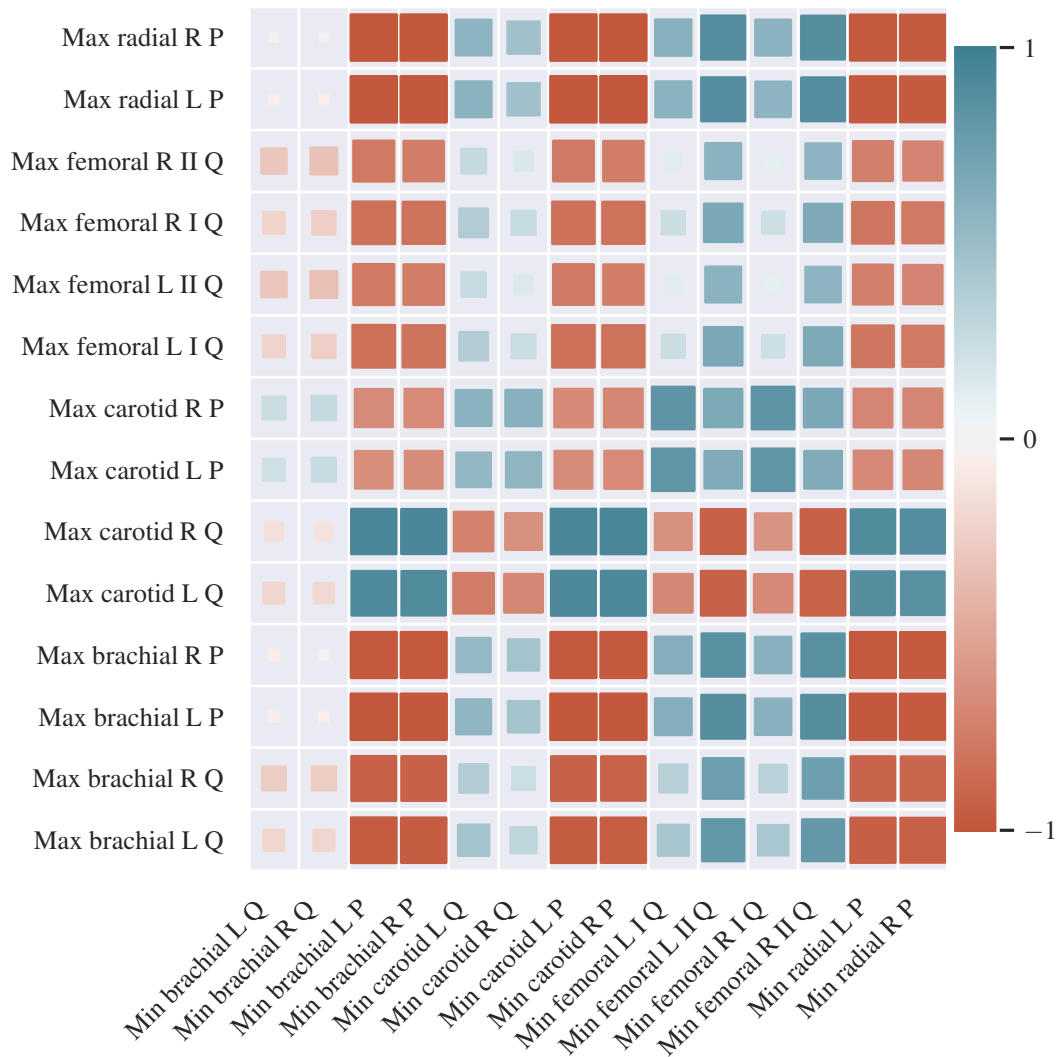


Figure G.14: The Pearson correlation coefficients between the minimum and maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.

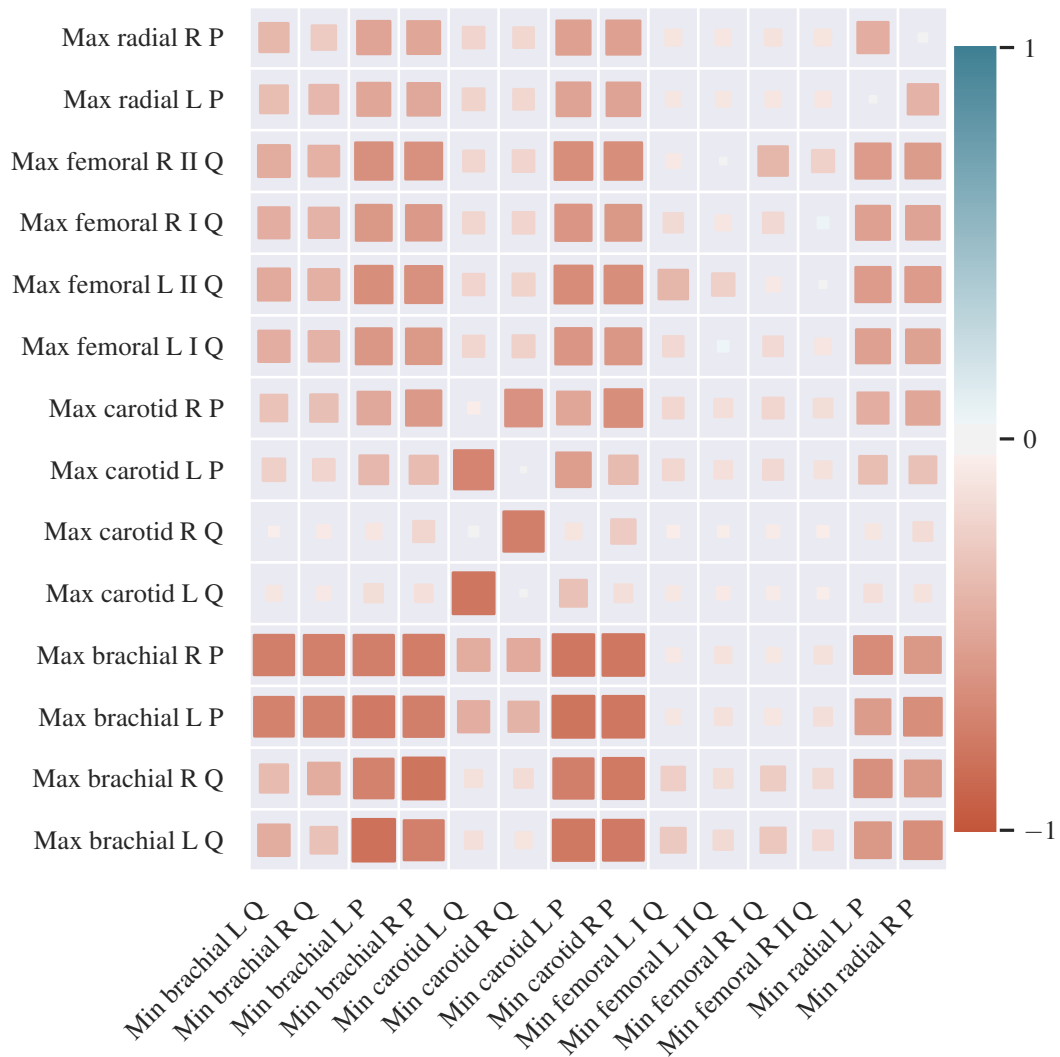


Figure G.15: The Pearson correlation coefficients between the minimum and maximum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.

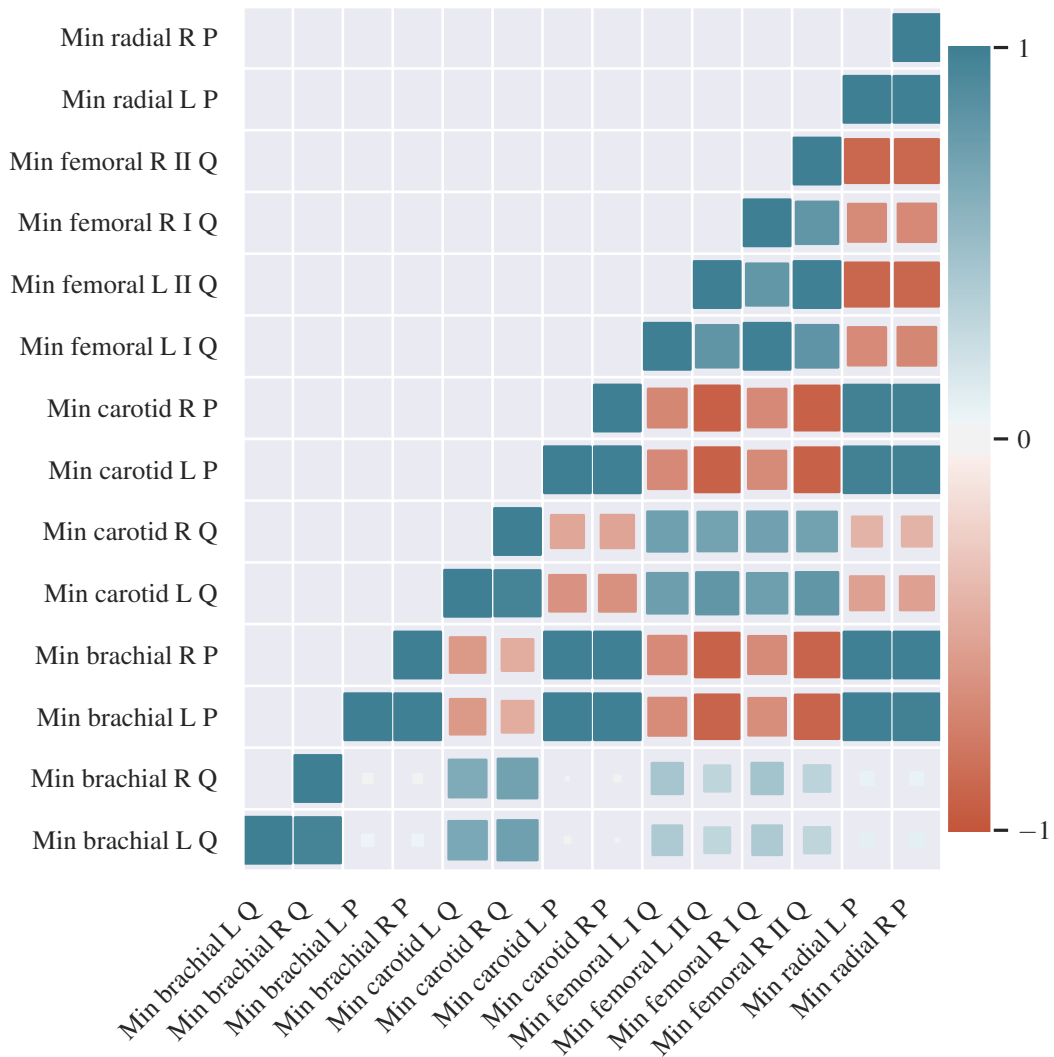


Figure G.16: The Pearson correlation coefficients between the minimum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a single constant scaling factor applied to the length of all vessels within the network.

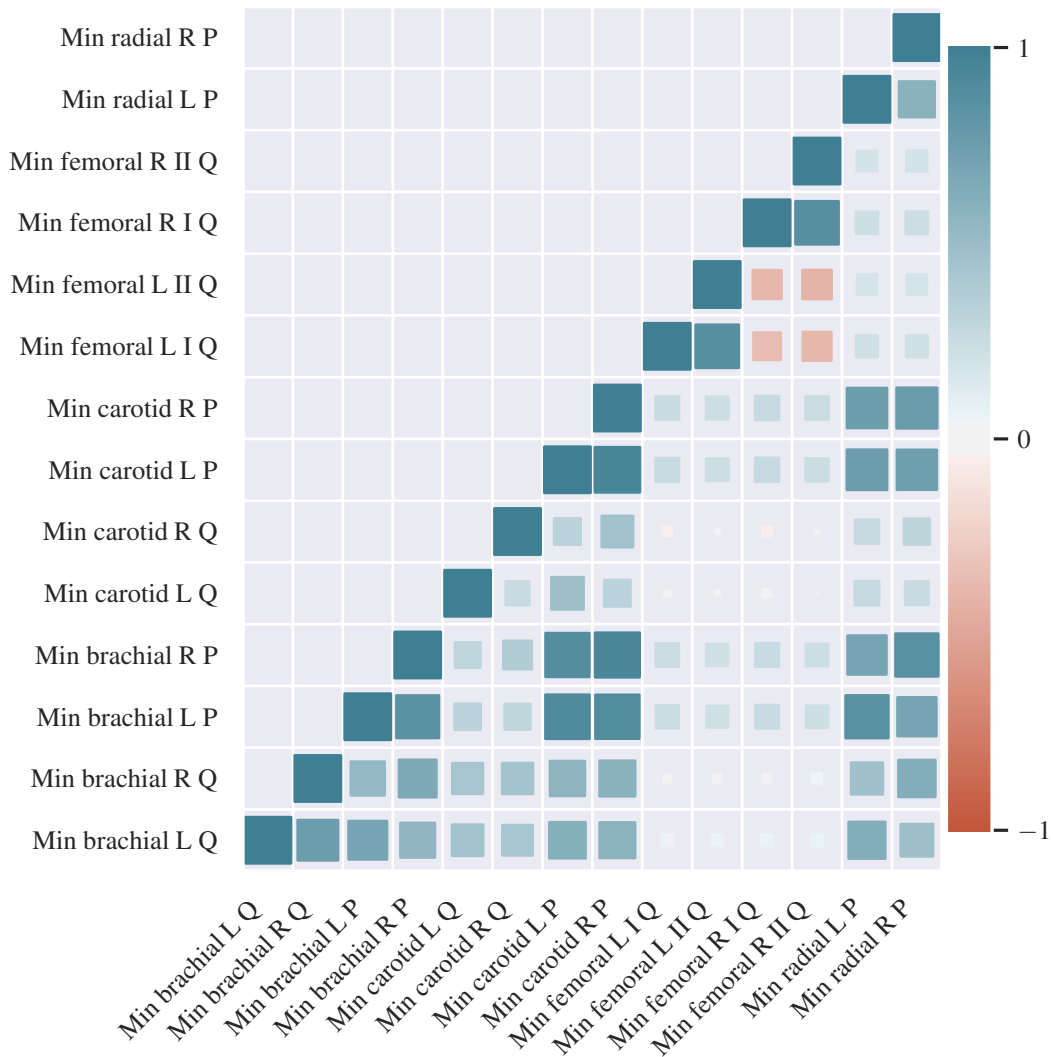


Figure G.17: The Pearson correlation coefficients between the minimum of pressure and flow-rate profiles at all examination locations are shown above. These correlations are computed using measurements taken from patients that have been generated using a individual scaling factors applied to the length of each vessel within the network.

Appendix H

MCMC trace plots

The trace plots of all parameters within each VPs arterial network are split into the following figures:

- The scaling terms applied to the reference parameters of the vessel wall mechanical property profiles of vessels with varying properties along their lengths are shown in Figures H.1, H.2, and H.3.
- The scaling terms applied to the reference parameters of the vessel radius property profiles of vessels with varying properties along their lengths are shown in Figures H.4, H.5, and H.6.
- The scaling terms applied to the reference daughter-parent ratios of the vessel wall mechanical properties and the radii of all vessels bifurcating off of the aorta with constant properties are shown in Figures H.7 and H.11 respectively.
- The scaling terms applied to the reference daughter-parent ratios of the vessel wall mechanical properties and the radii are shown for vessels with constant properties in the right upper extremities in Figures H.8 and H.12 respectively, and left upper extremities in Figures H.9 and H.13 respectively.
- The scaling terms applied to the reference daughter-parent ratios of the vessel wall mechanical properties and the radii of vessels with constant properties within the lower extremities are shown in Figures H.10 and H.14 respectively.
- The scaling terms applied to the reference parameters of the Windkessel models in the aortic region, right upper extremities, left upper extremities, and the legs are shown in Figures H.15, H.16, H.17, and H.18 respectively.

To aid in visualisation of the results shown in the above listed figures, all trace plots are thinned out by plotting only each 100th iteration of the chain. This is done to filter out high frequency noise, clarifying the low frequency behaviour of the chain.

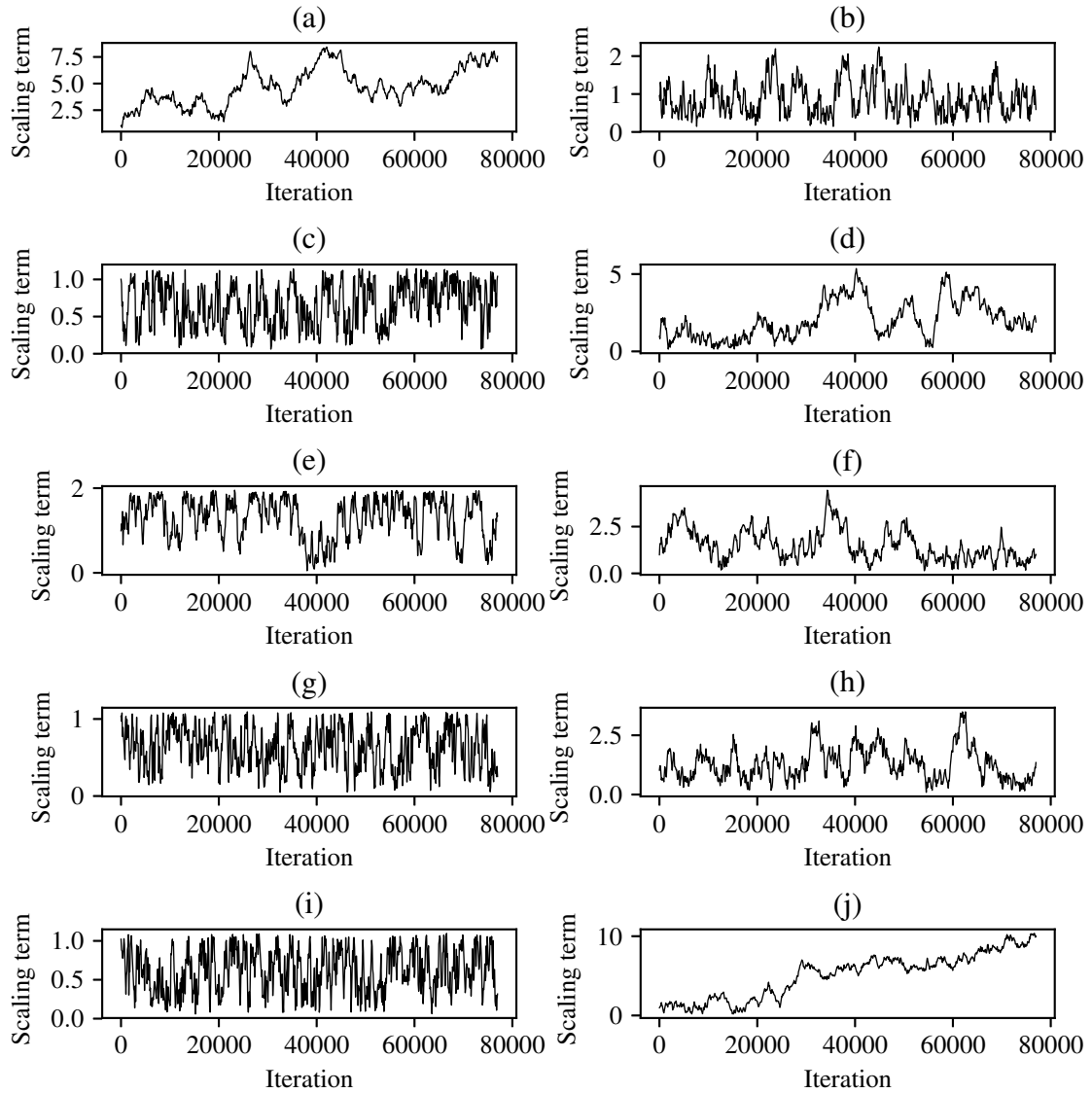


Figure H.1: The parameter scaling terms at each 100th iteration of the Markov chain applied to the: aorta chain initialising value (a), aorta chain decay term (b), right arm chain daughter/parent ratio (c), right arm chain decay term (d), left arm chain daughter/parent ratio (e), left arm chain decay term (f), right leg chain daughter/parent ratio (g), right leg chain decay term (h), left leg chain daughter/parent ratio (i), and the left leg chain decay term (j) for the β properties of the network.

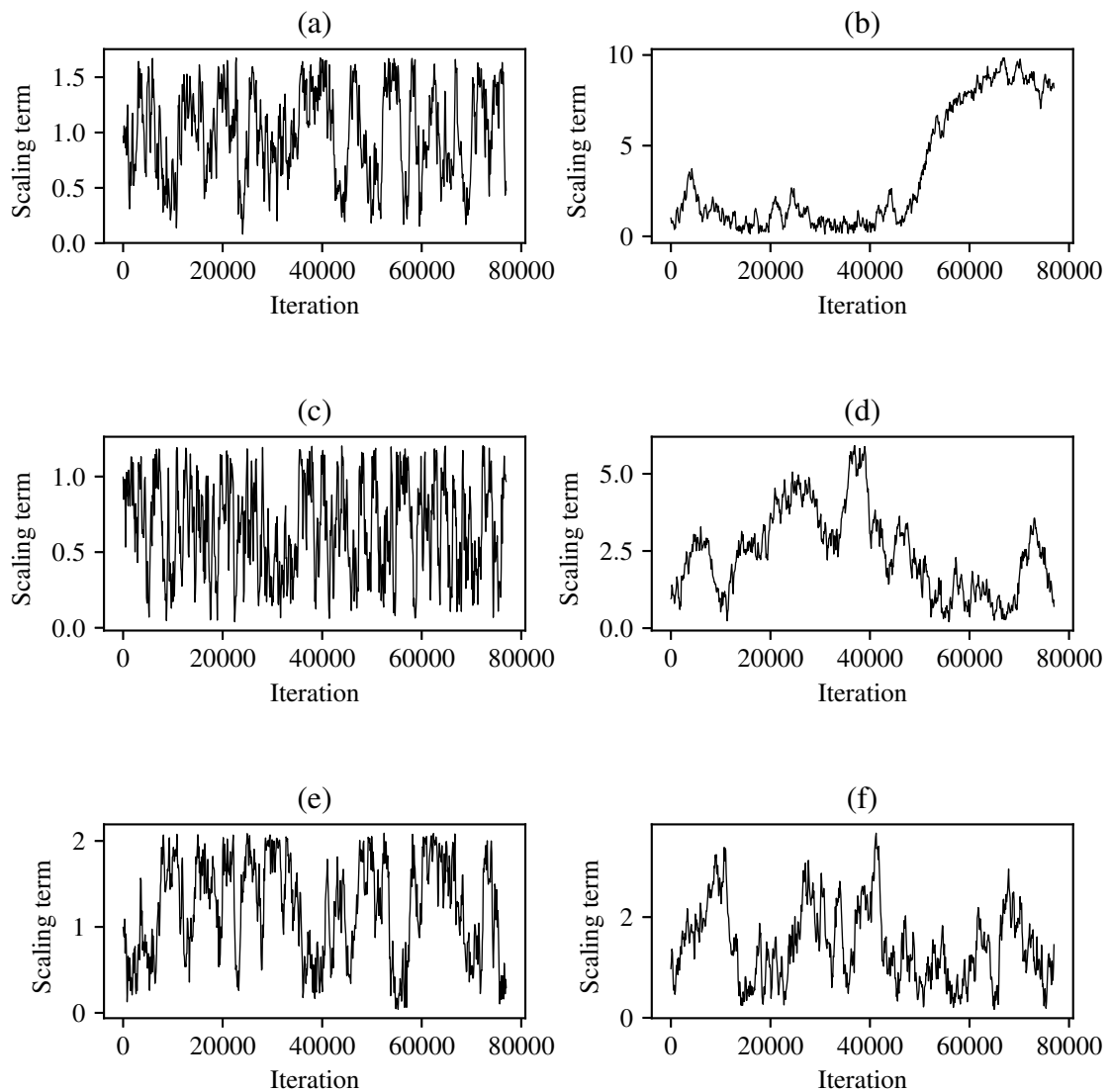


Figure H.2: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the: brachiocephalic trunk daughter/parent ratio (a), brachiocephalic trunk decay term (b), right common carotid daughter/parent ratio (c), right common carotid decay term (d), left common carotid daughter/parent ratio (e), and the left common carotid decay term (f) for the β properties of the network.

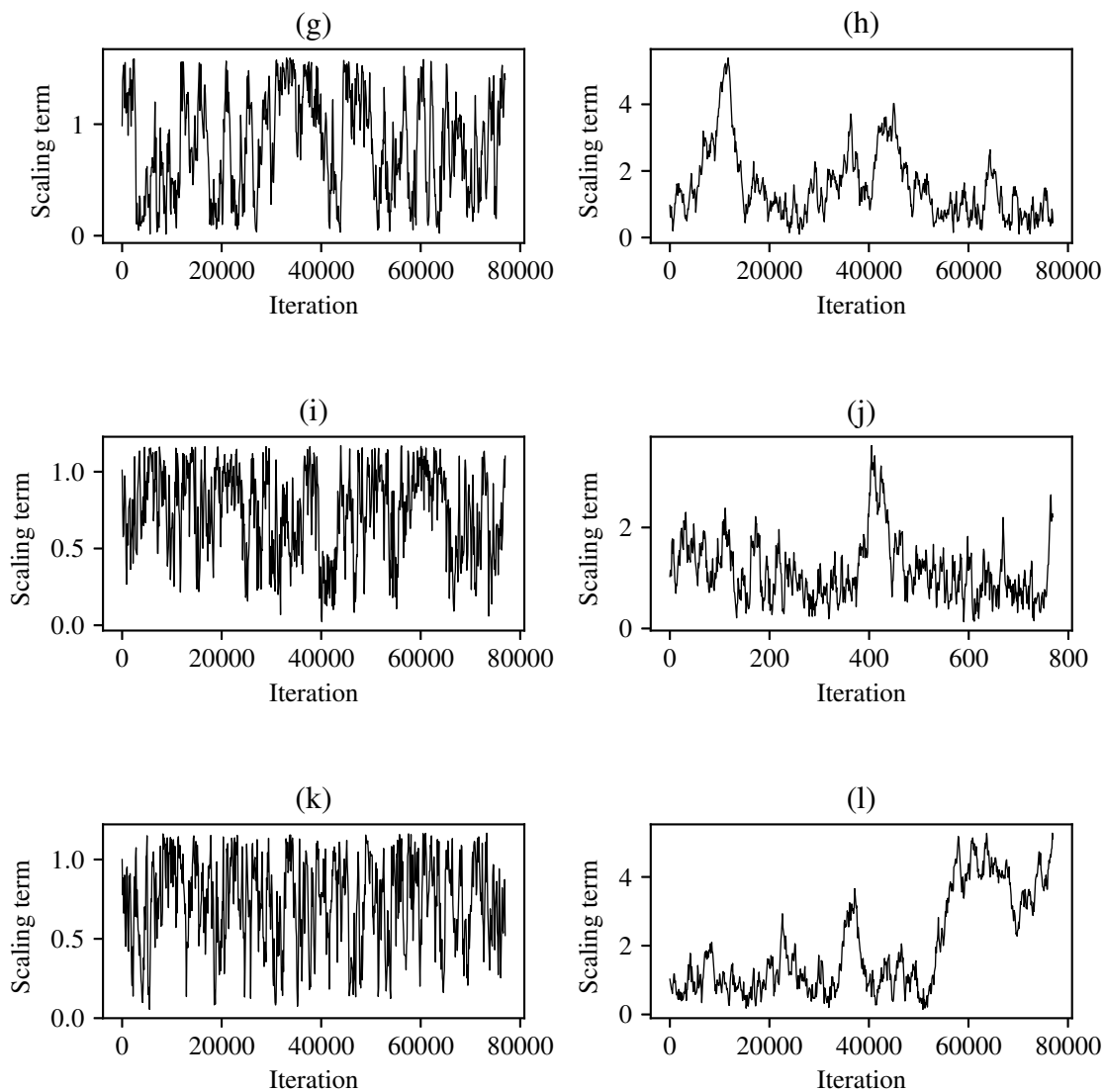


Figure H.3: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the: celiac trunk daughter/parent ratio (g), celiac trunk decay term (h), right common iliac daughter/parent ratio (i), right common iliac decay term (j), left common iliac daughter/parent ratio (k), and the left common iliac decay term (l) for the β properties of the network.

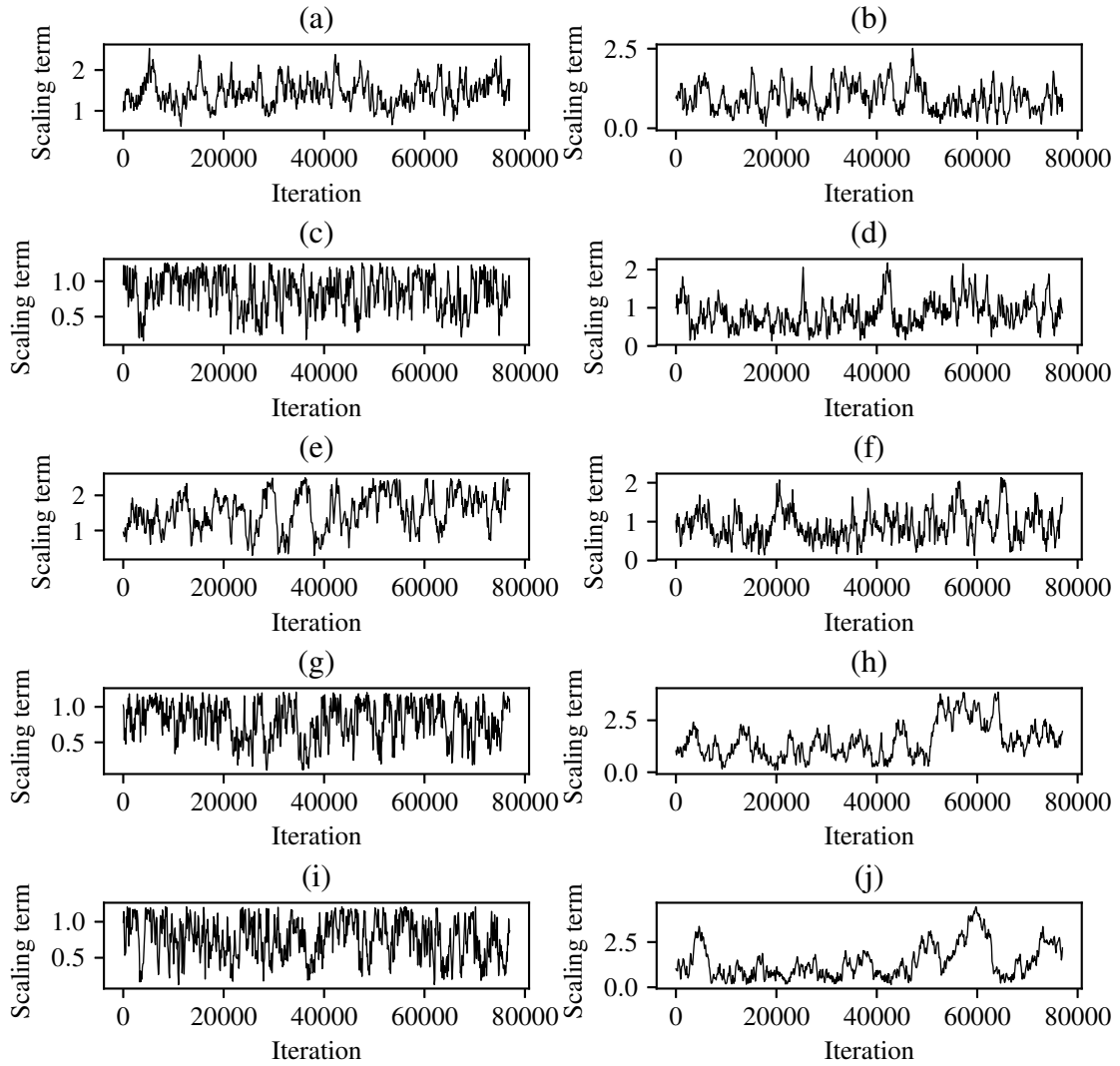


Figure H.4: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the: aorta chain initialising value (a), aorta chain decay term (b), right arm chain daughter/parent ratio (c), right arm chain decay term (d), left arm chain daughter/parent ratio (e), left arm chain decay term (f), right leg chain daughter/parent ratio (g), right leg chain decay term (h), left leg chain daughter/parent ratio (i), and the left leg chain decay term (j) for the r_0 properties of the network.

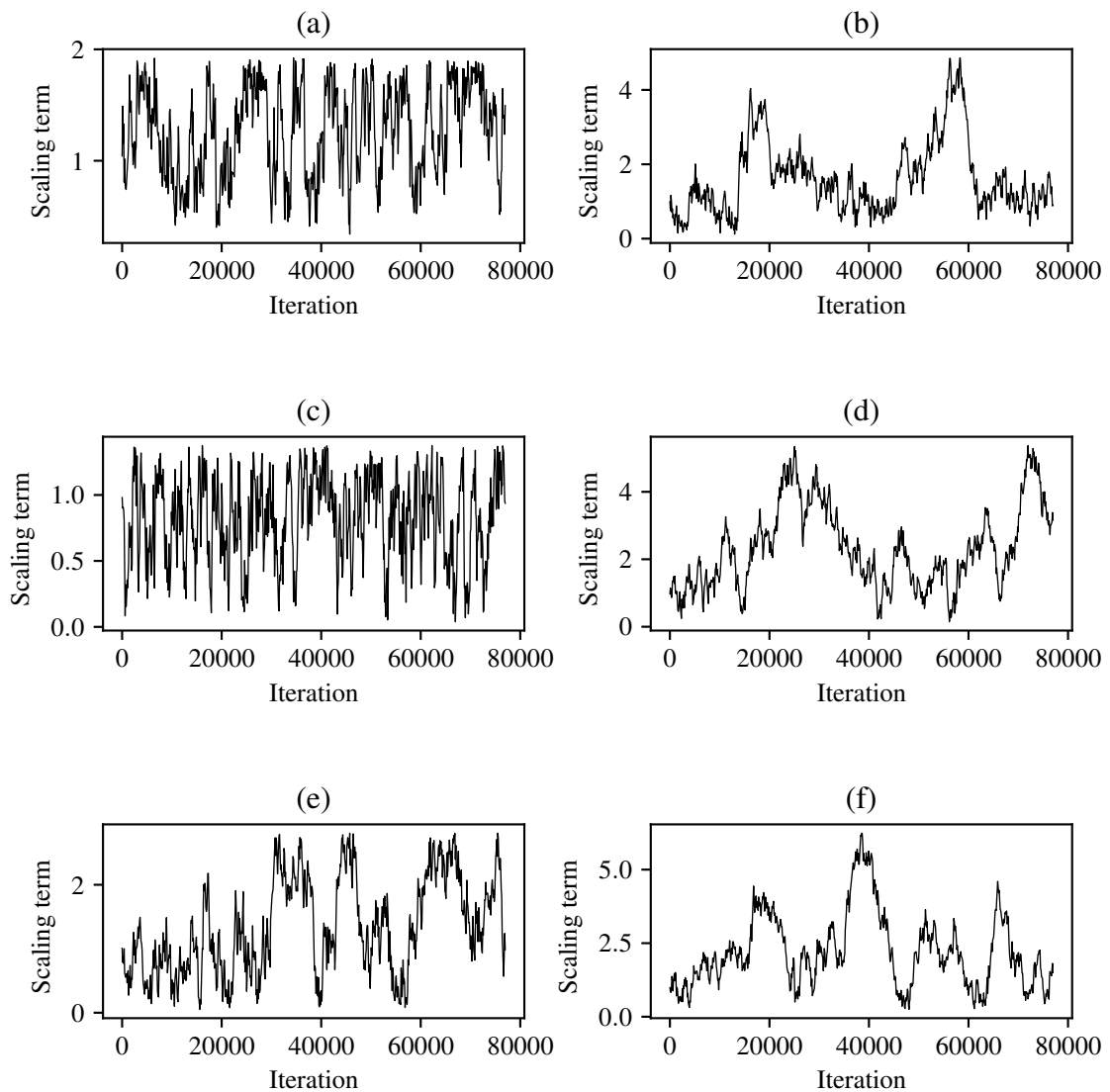


Figure H.5: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the: brachiocephalic trunk daughter/parent ratio (a), brachiocephalic trunk decay term (b), right common carotid daughter/parent ratio (c), right common carotid decay term (d), left common carotid daughter/parent ratio (e), the left common carotid decay term (f) for the r_0 properties of the network.

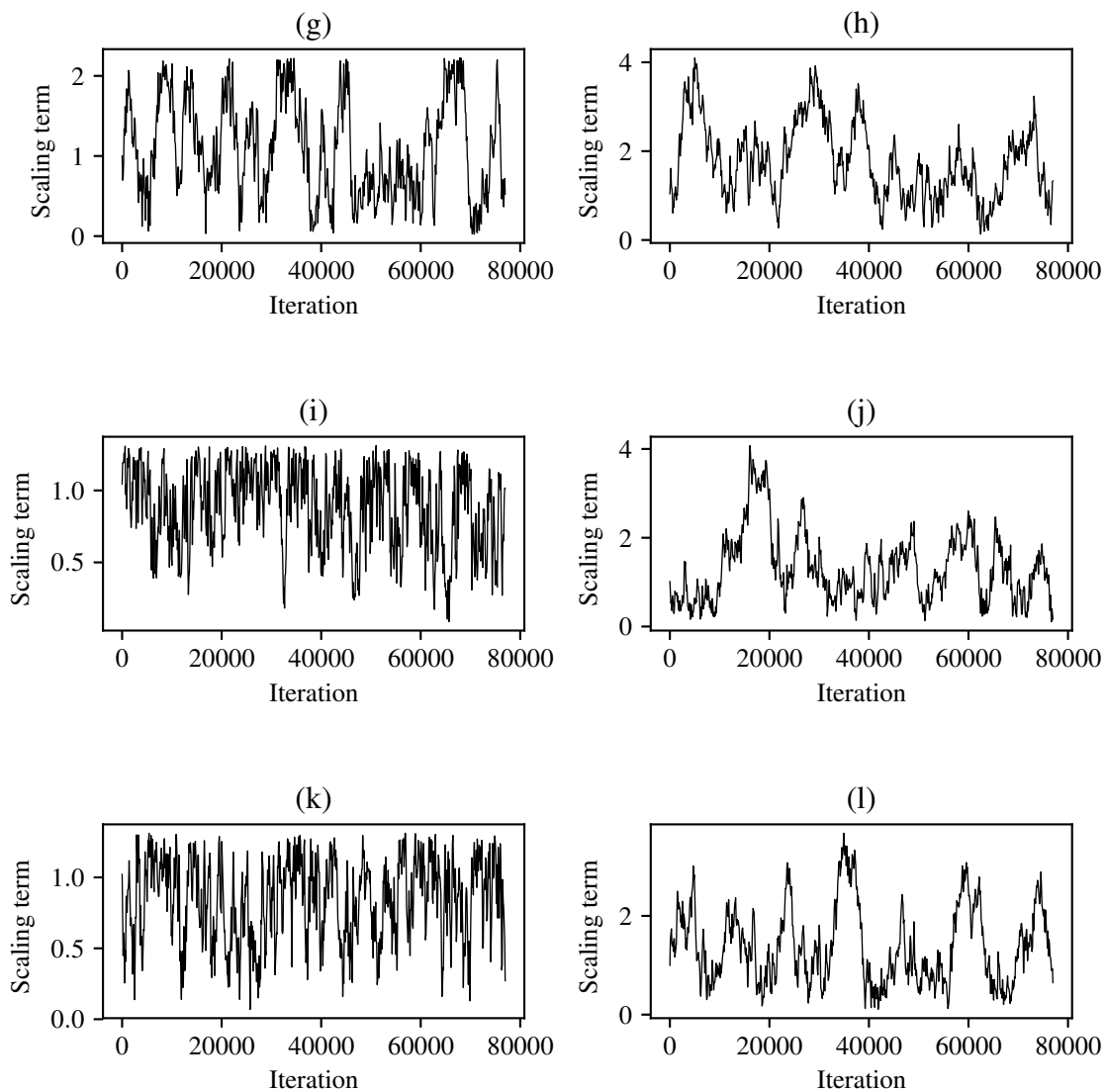


Figure H.6: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the: celiac trunk daughter/parent ratio (g), celiac trunk decay term (h), right common iliac daughter/parent ratio (i), right common iliac decay term (j), left common iliac daughter/parent ratio (k), and the left common iliac decay term (l) for the r_0 properties of the network.

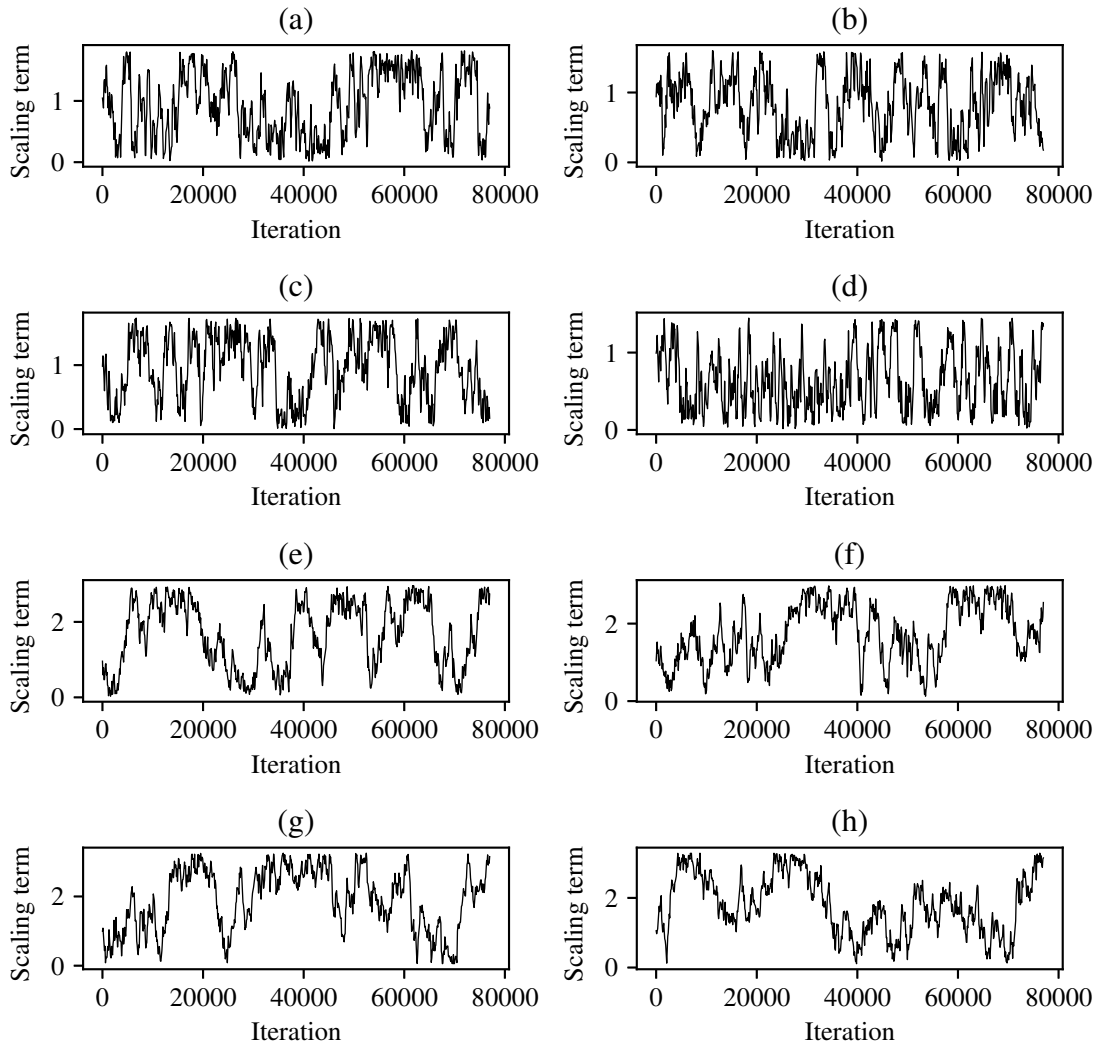


Figure H.7: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: inferior mesenteric (a), right renal (b), left renal (c), superior mesenteric (d), second left posterior intercostal (e), second right posterior intercostal (f), first left posterior intercostal (g), first right posterior intercostal (h) for the β properties of the network.

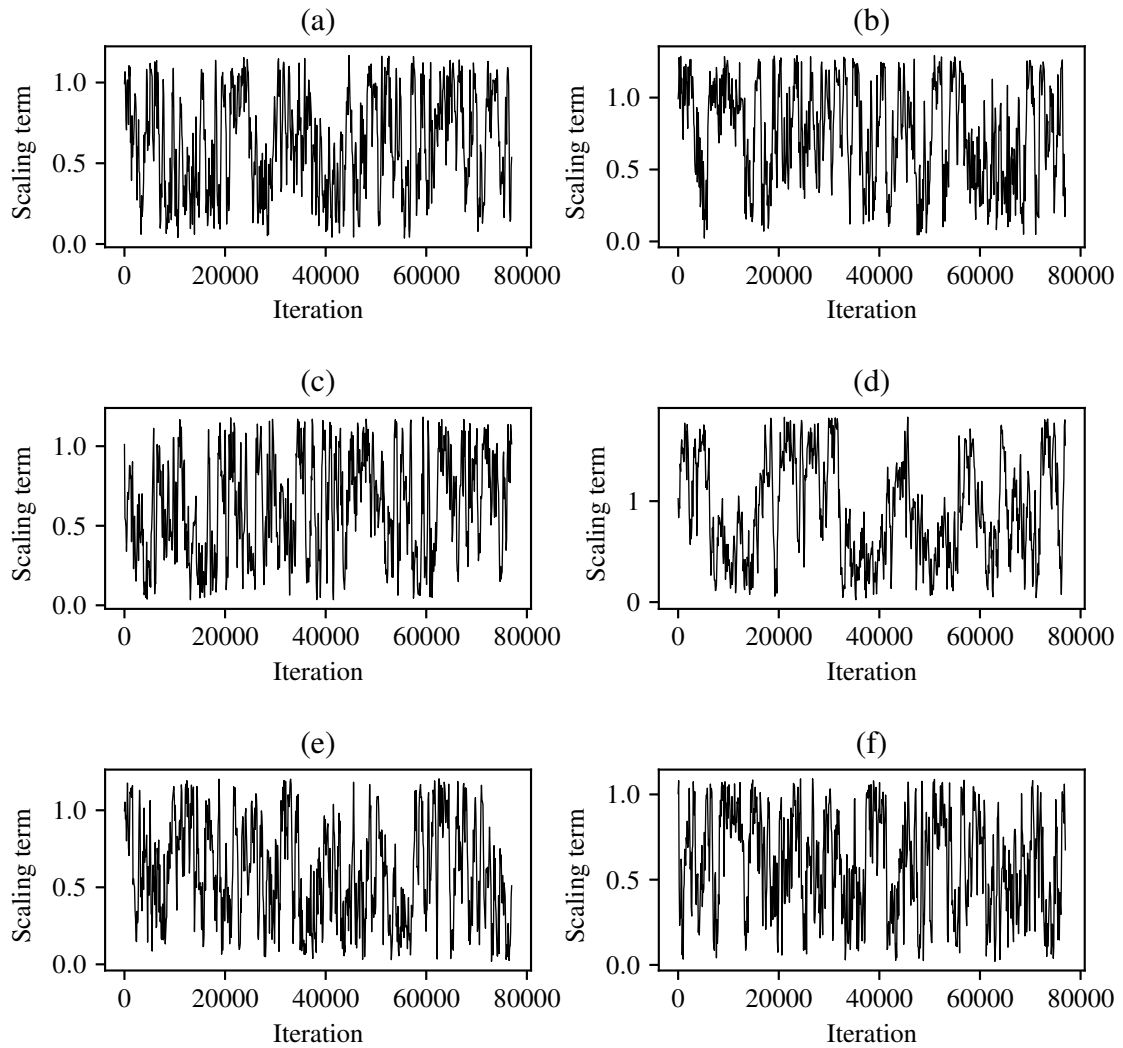


Figure H.8: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: first right ulnar (a), right common interosseous (b), right radial (c), right vertebral (d), right external carotid (e), right internal carotid (f) for the β properties of the network.

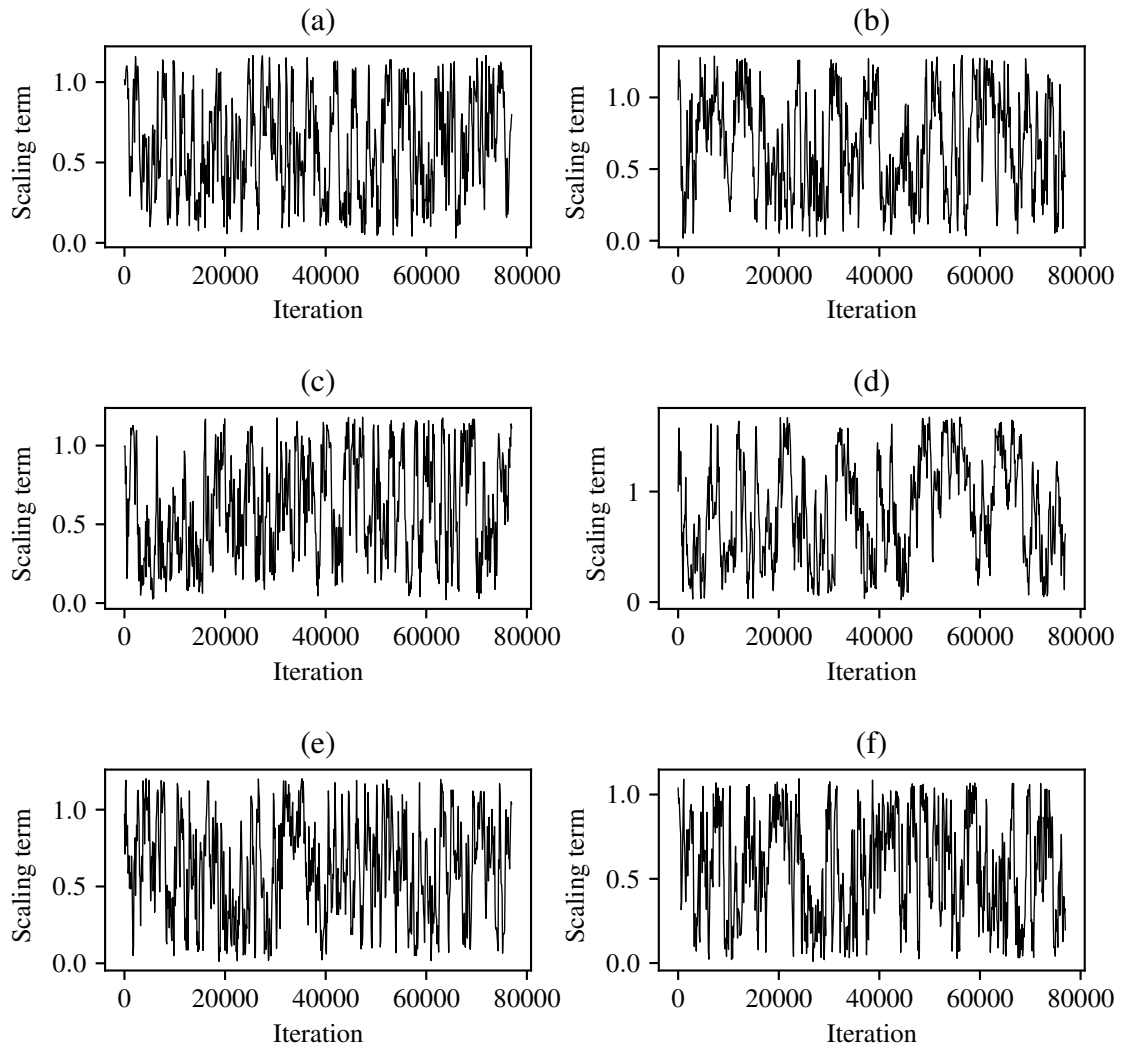


Figure H.9: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: first left ulnar (a), left common interosseous (b), left radial (c), left vertebral (d), left external carotid (e), left internal carotid (f) for the β properties of the network.

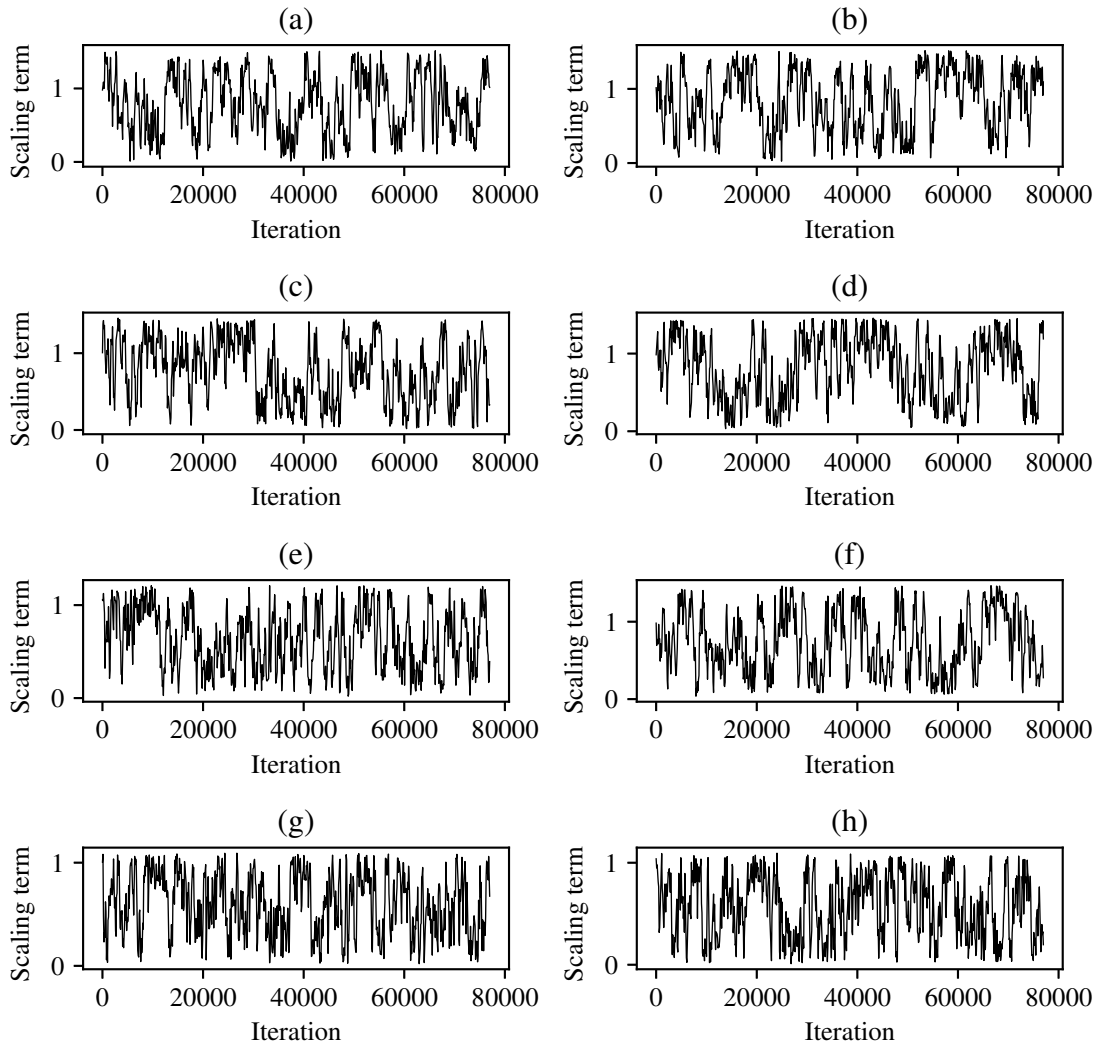


Figure H.10: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: right anterior tibial (a), left anterior tibial (b), right posterior tibial (c), left posterior tibial (d), right profunda femoris (e), left profunda femoris (f), right internal carotid (g), left internal carotid (h) for the β properties of the network.

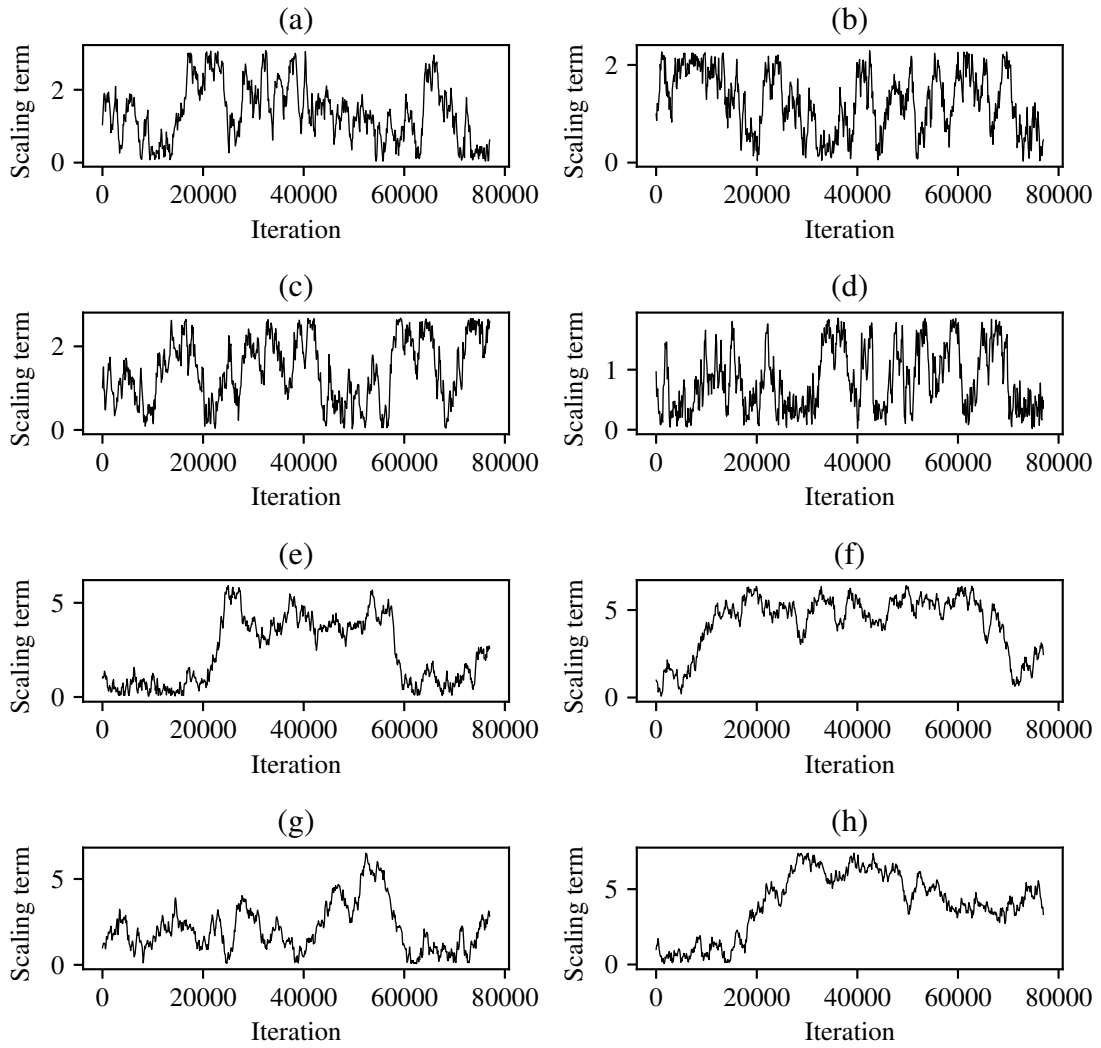


Figure H.11: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: inferior mesenteric (a), right renal (b), left renal (c), superior mesenteric (d), second left posterior intercostal (e), second right posterior intercostal (f), first left posterior intercostal (g), first right posterior intercostal (h) for the r_0 properties of the network.

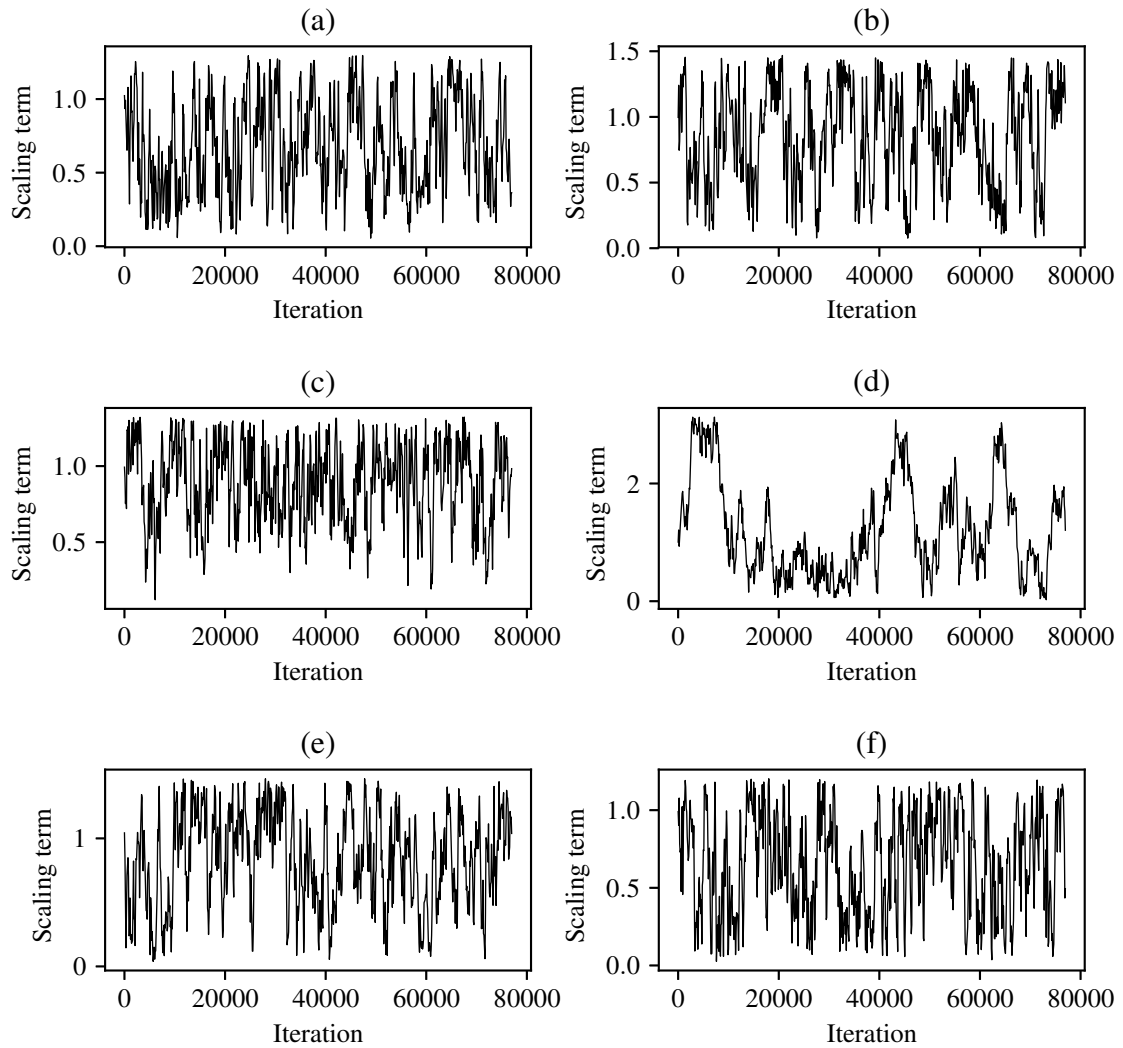


Figure H.12: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: first right ulnar (a), right common interosseous (b), right radial (c), right vertebral (d), right external carotid (e), right internal carotid (f) for the r_0 properties of the network.

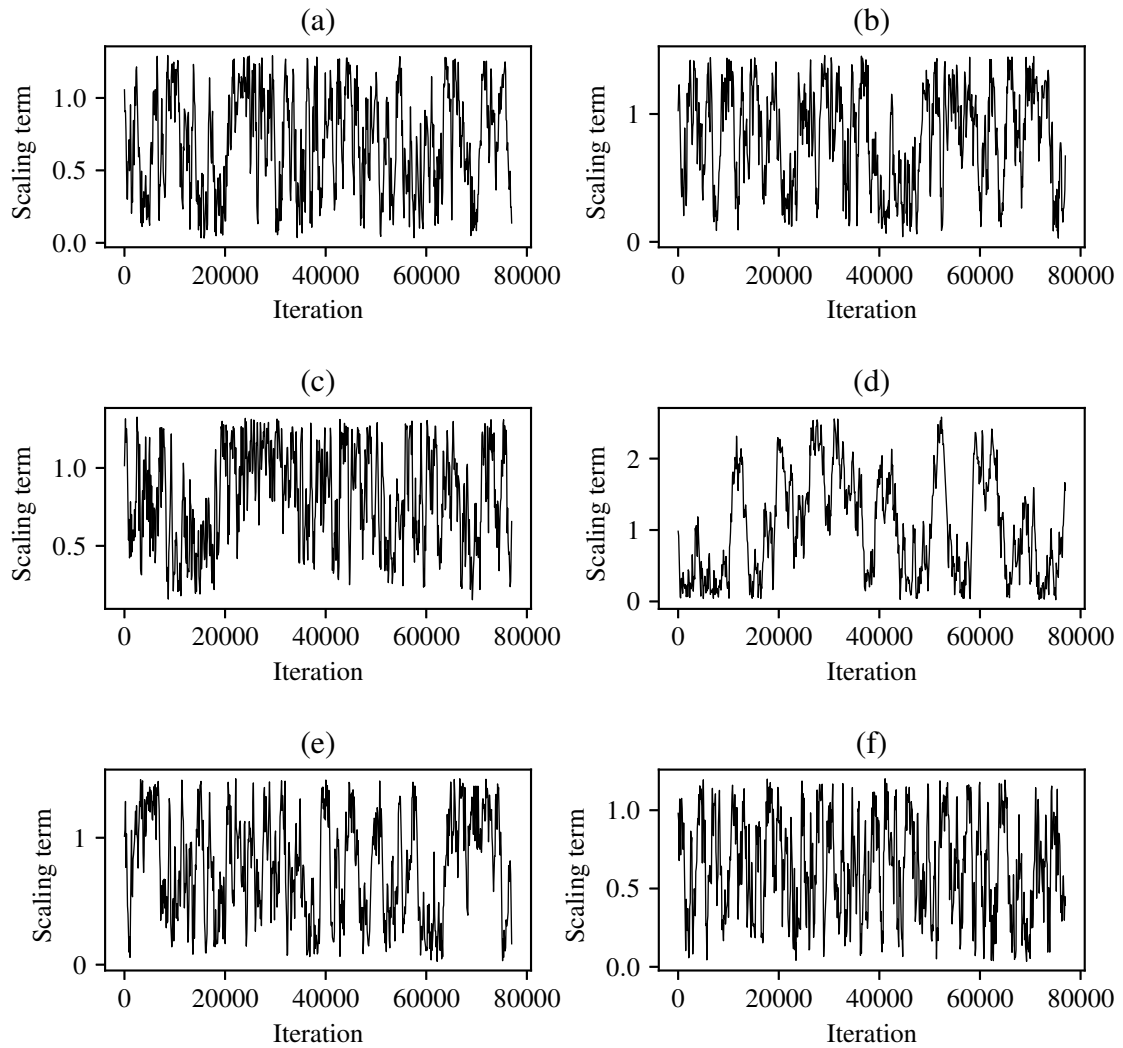


Figure H.13: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: first left ulnar (a), left common interosseous (b), left radial (c), left vertebral (d), left external carotid (e), left internal carotid (f) for the r_0 properties of the network.

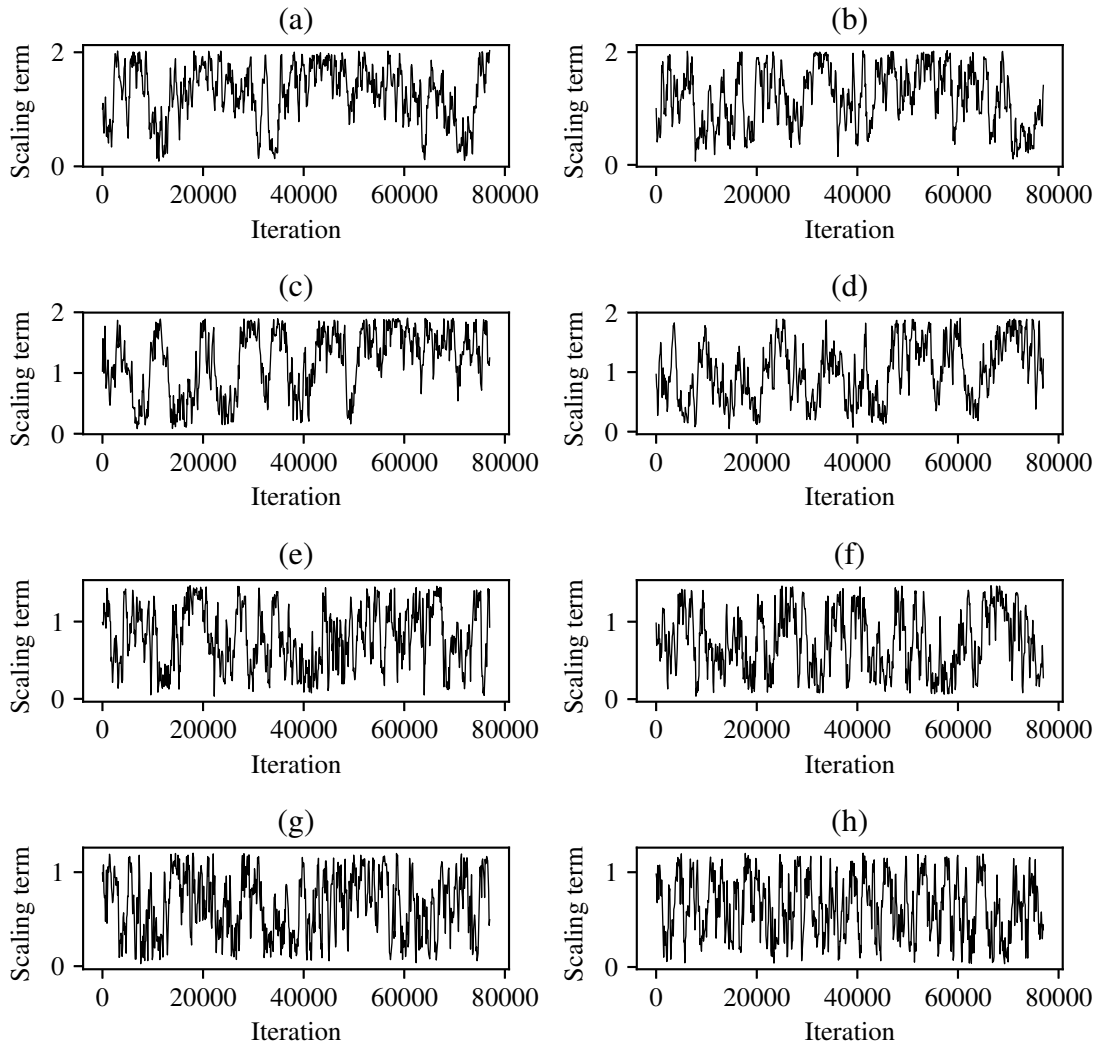


Figure H.14: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the daughter/parent ratio of the: right anterior tibial (a), left anterior tibial (b), right posterior tibial (c), left posterior tibial (d), right profunda femoris (e), left profunda femoris (f), right internal carotid (g), left internal carotid (h) for the r_0 properties of the network.

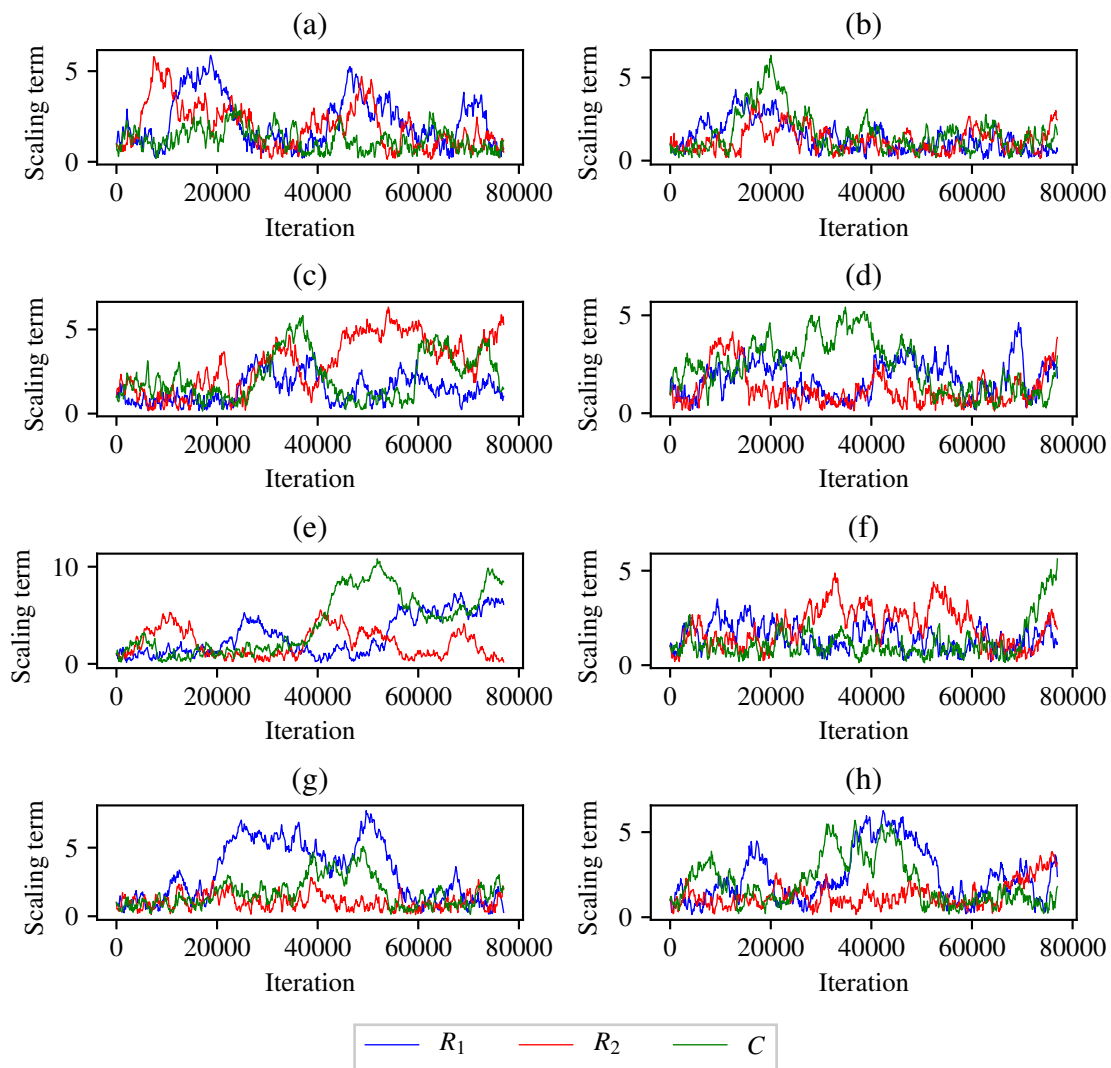


Figure H.15: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: inferior mesenteric (a), right renal (b), left renal (c), superior mesenteric (d), second left posterior intercostal (e), second right posterior intercostal (f), first left posterior intercostal (g), and the first right posterior intercostal (h).

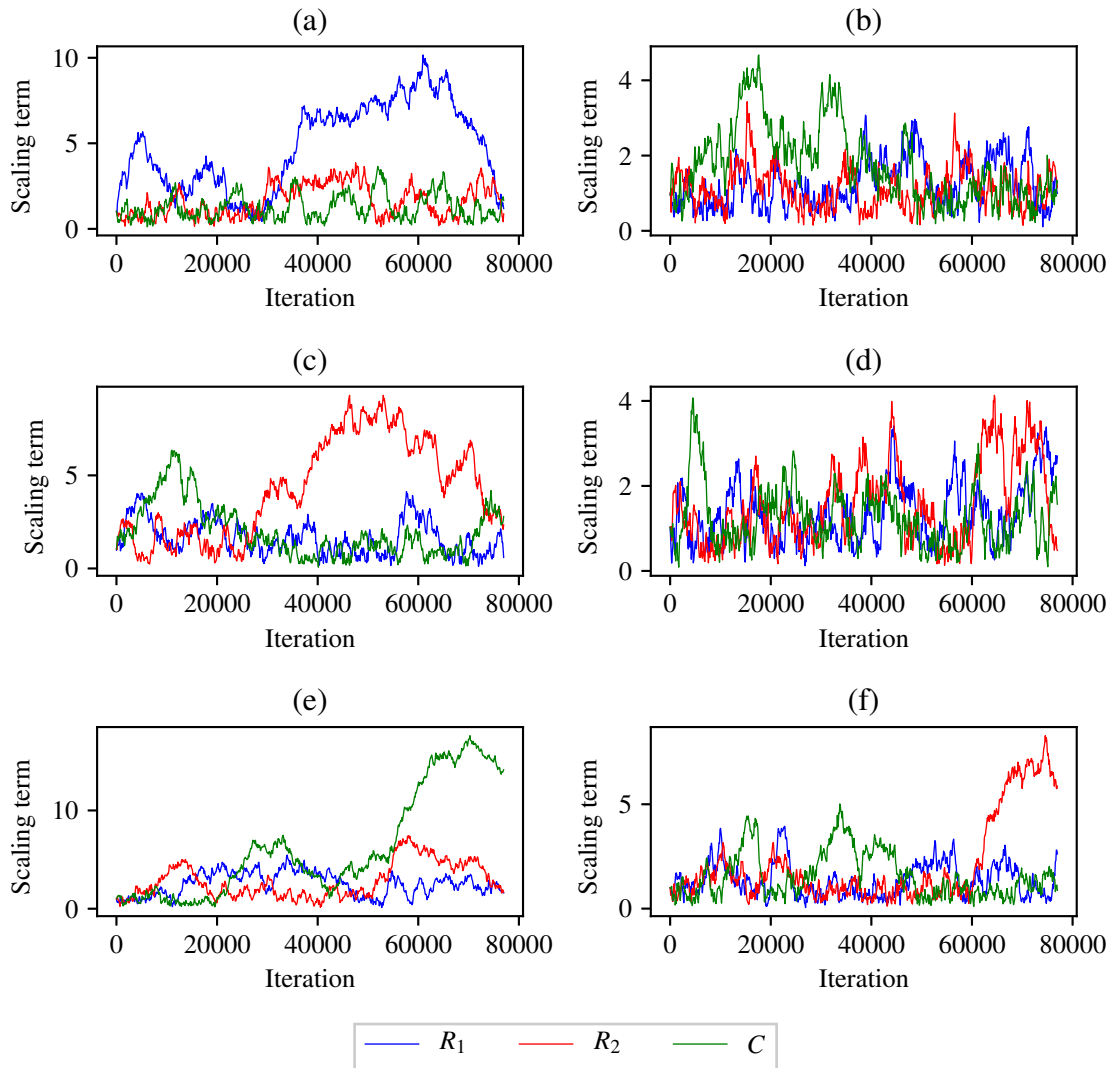


Figure H.16: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: second right ulnar (a), right common interosseous (b), right radial (c), right vertebral (d), right external carotid (e), and the right internal carotid (f).

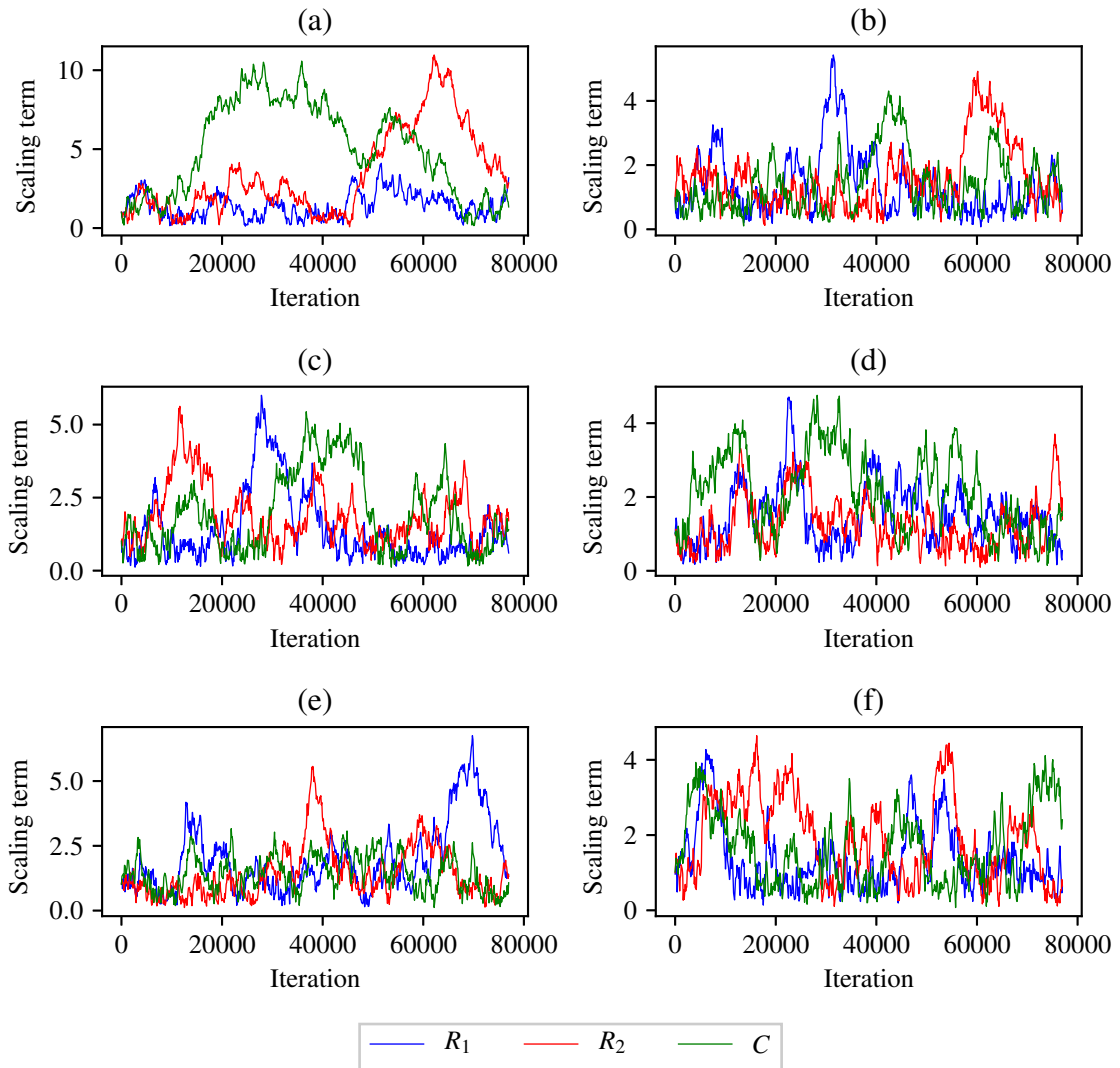


Figure H.17: The parameter scaling terms at each 100th iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: second left ulnar (a), left common interosseous (b), left radial (c), left vertebral (d), left external carotid (e), and the left internal carotid (f).

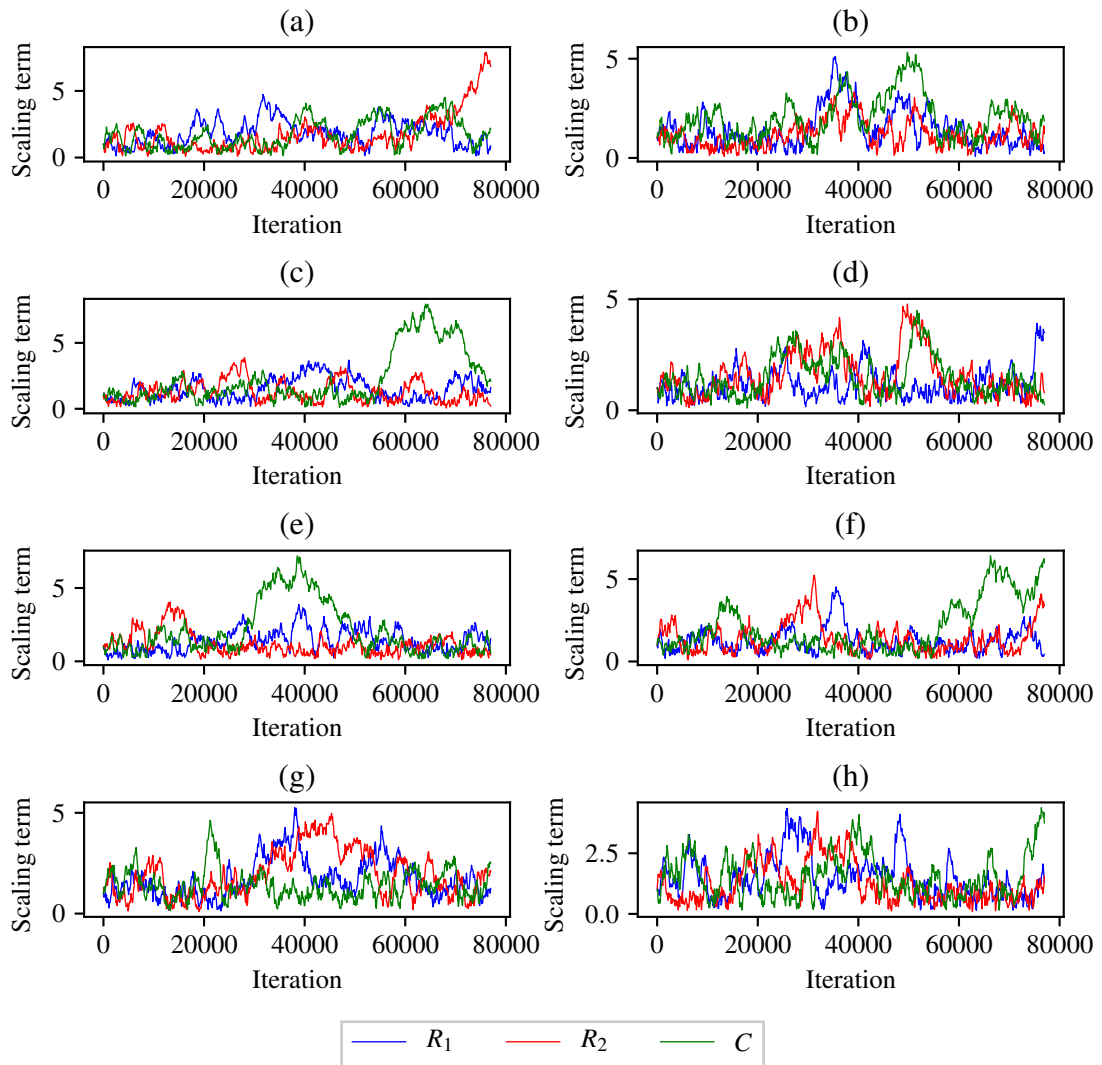


Figure H.18: The parameter scaling terms at each 100^{th} iteration of the Markov chain applied to the Windkessel model parameters at the terminal boundary of the: right anterior tibial (a), left anterior tibial (b), right posterior tibial (c), left posterior tibial (d), right profunda femoris (e), left profunda femoris (f), right internal iliac (g), left internal iliac (h).

Appendix I

Pressure and flow-rate profiles from random VPs

Within the subsequent three figures (Figures I.1, I.2, and I.3) pressure and flow-rate profiles are shown from 15 VPs randomly sampled from the VPD. The pressure or flowrate profile at each location measured within the reference network is also included, as well as the literature based measurements and associated error incorporated into the posterior distribution.

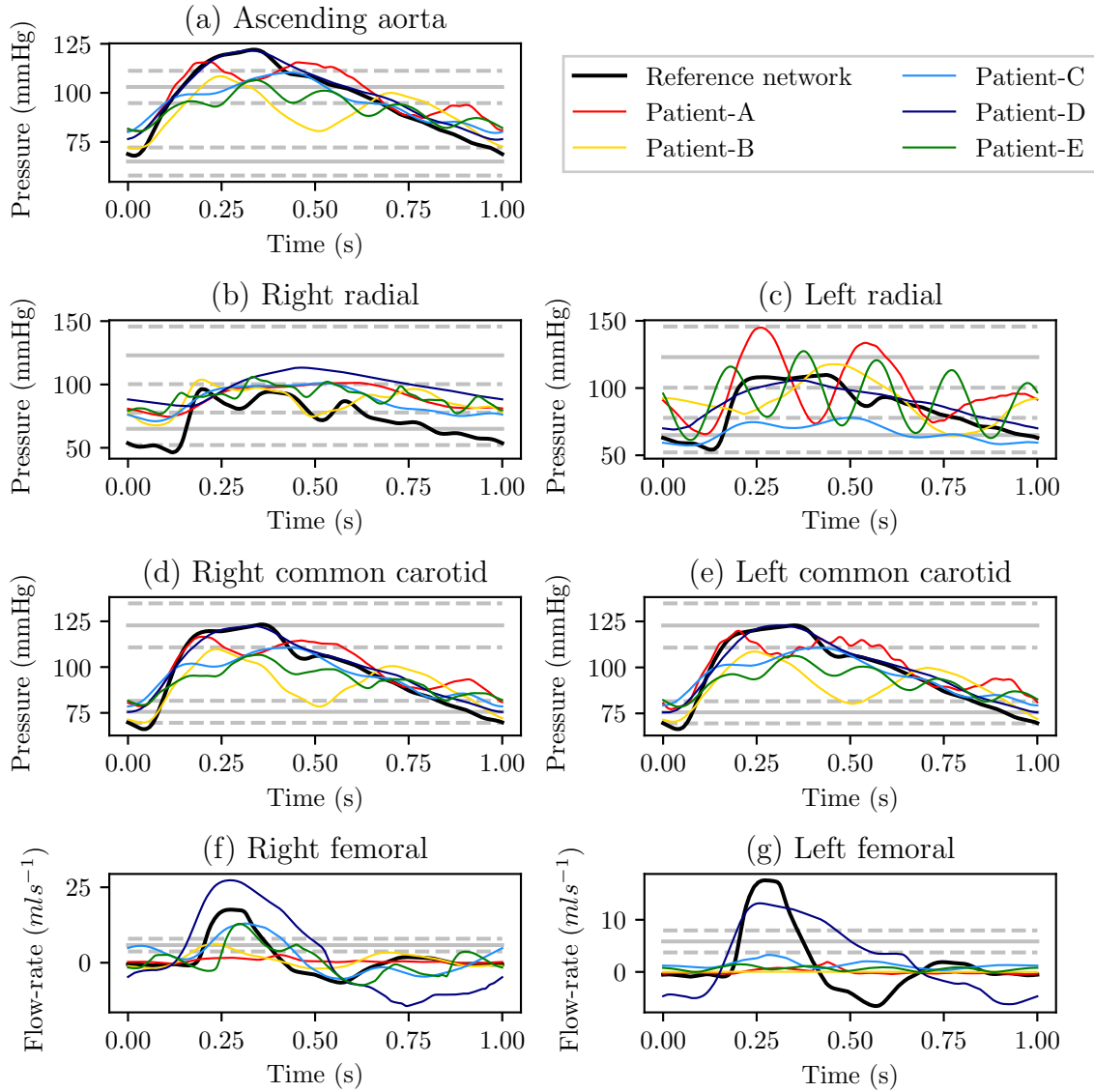


Figure I.1: In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively.

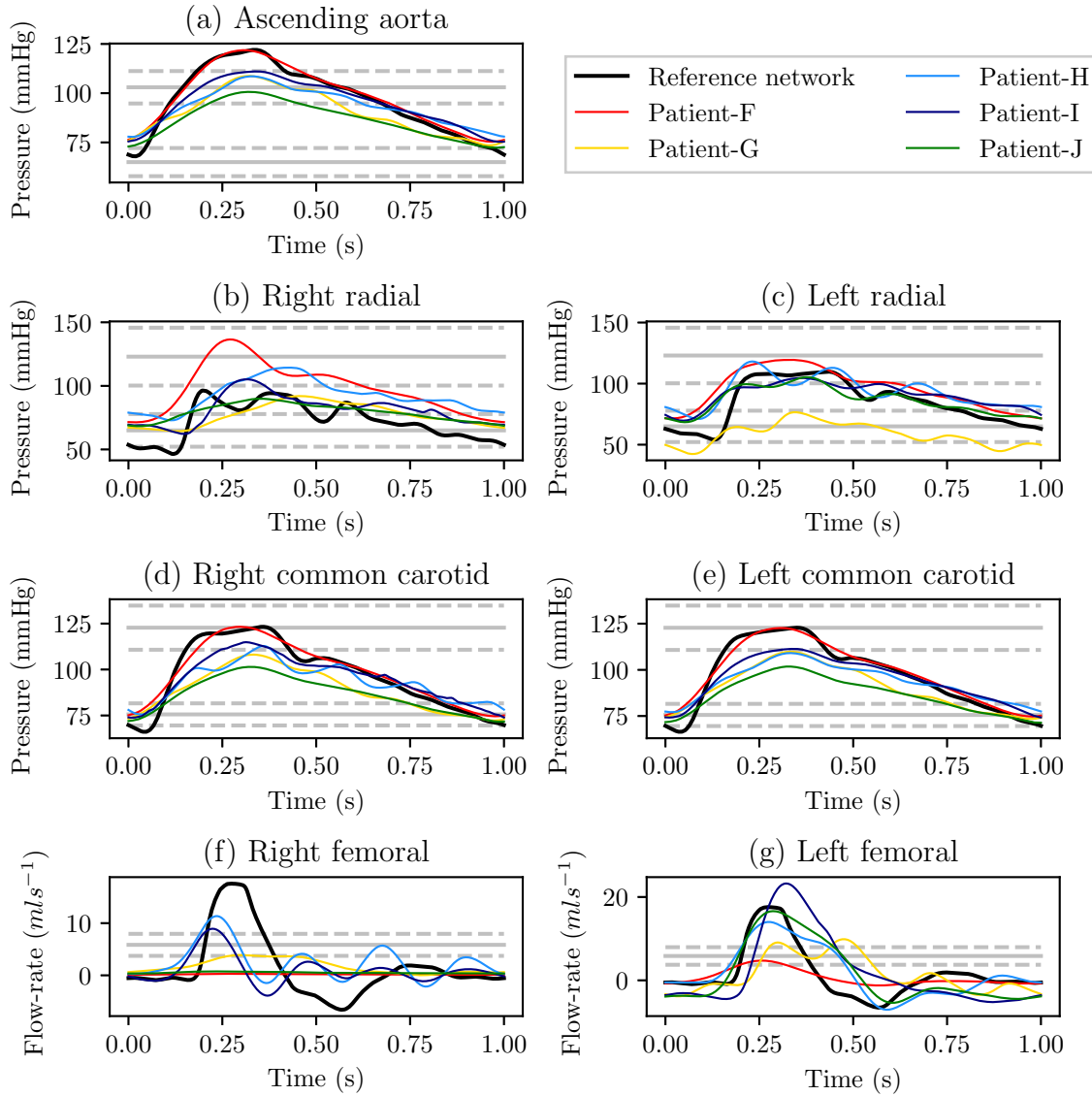


Figure I.2: In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right common carotid artery (d), pressure profiles in the left common carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively.

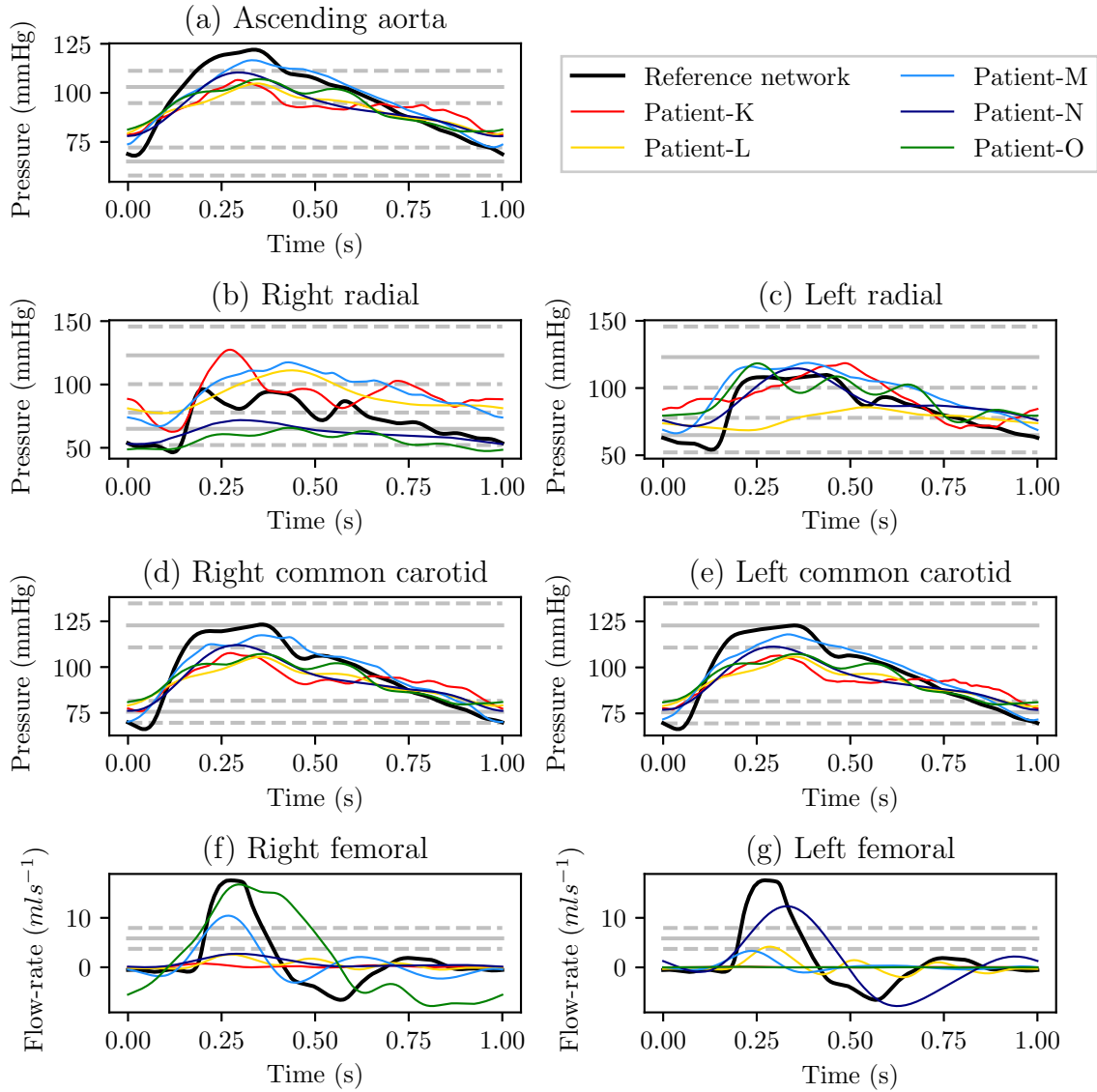


Figure I.3: In the above figure the subplots show the: pressure profiles in the ascending aorta (a), pressure profiles in the right radial artery (b), pressure profiles in the left radial artery (c), pressure profiles in the right carotid artery (d), pressure profiles in the left carotid artery (e), flow-rate profiles in the right second femoral artery (f), and flow-rate profiles in the left second femoral artery (g). In each figure the profiles taken from the reference network are shown in black; and the literature reported measurements and associated error are shown by the solid and dashed grey lines respectively.

Appendix J

Hyperparameter contour plots

J.1 CAS

The contour plots of the F_1 score achieved when using all six measurements and different combinations of hyper-parameters are shown when employing the RF, GB, and MLP methods in Figures J.1, J.2, and J.3, respectively. It is seen from the three aforementioned figures that there are not clear sharp peaks and troughs in the F_1 score achieved, and thus accuracy appears to be relatively insensitive to the combination of hyper-parameters used.

J.2 SAS

The contour plots of the F_1 score achieved when using all six measurements and different combinations of hyper-parameters are shown when employing the RF, GB, and MLP methods in Figures J.4, J.5, and J.6, respectively. As with CAS, the accuracy appears to be relatively insensitive to the combination of hyper-parameters used.

J.3 PAD

The contour plots of the F_1 score achieved when using all six measurements and different combinations of hyper-parameters are shown when employing the RF, GB, and MLP methods in Figures J.7, J.8, and J.9, respectively. As with CAS and SAS the accuracy appears to be relatively insensitive to the combination of hyper-parameters used.

J.4 AAA

The contour plots of the F_1 score achieved when using all six measurements and different combinations of hyper-parameters are shown when employing the RF,

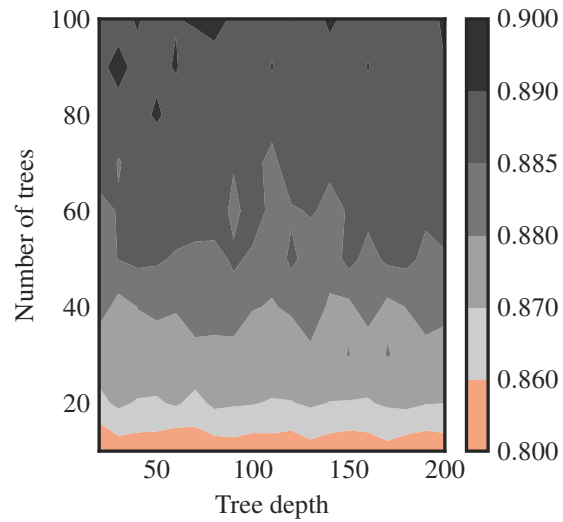


Figure J.1: The contour plot of the F_1 scores achieved for CAS classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.

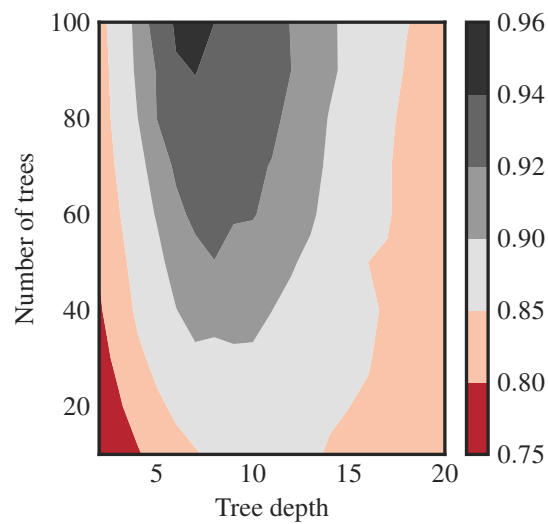


Figure J.2: The contour plot of the F_1 scores achieved for CAS classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.

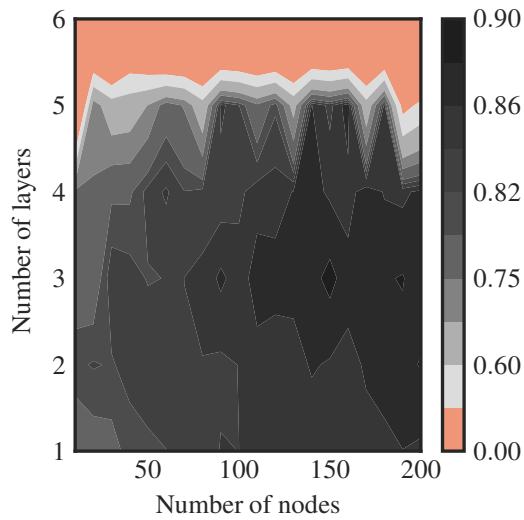


Figure J.3: The contour plot of the F_1 scores achieved for CAS classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.

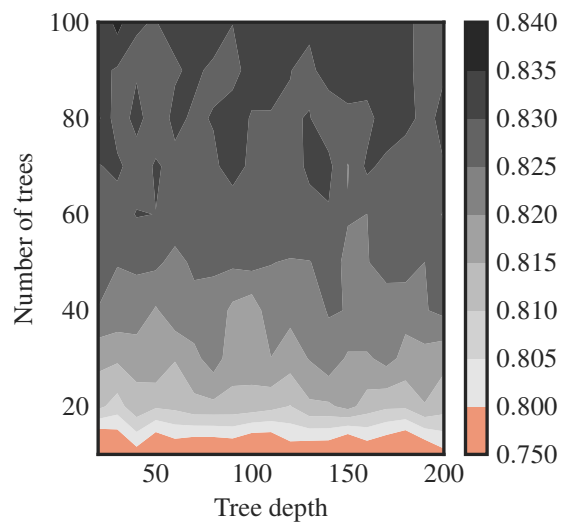


Figure J.4: The contour plot of the F_1 scores achieved for SAS classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.

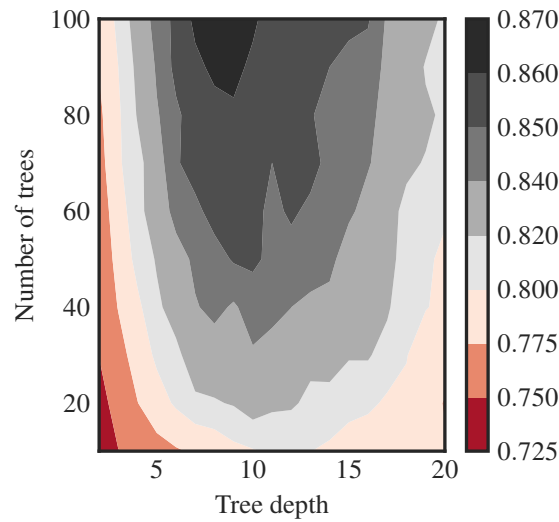


Figure J.5: The contour plot of the F_1 scores achieved for SAS classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.

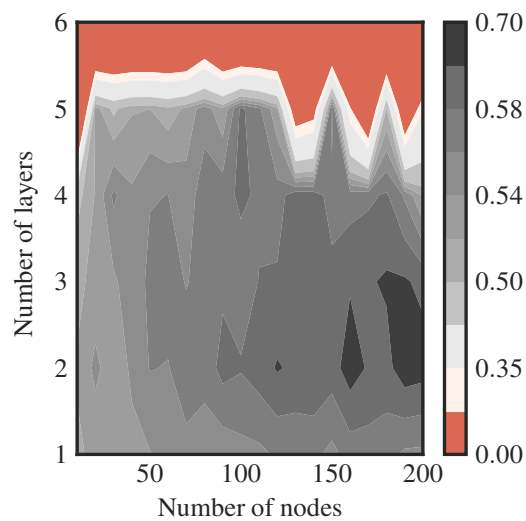


Figure J.6: The contour plot of the F_1 scores achieved for SAS classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.

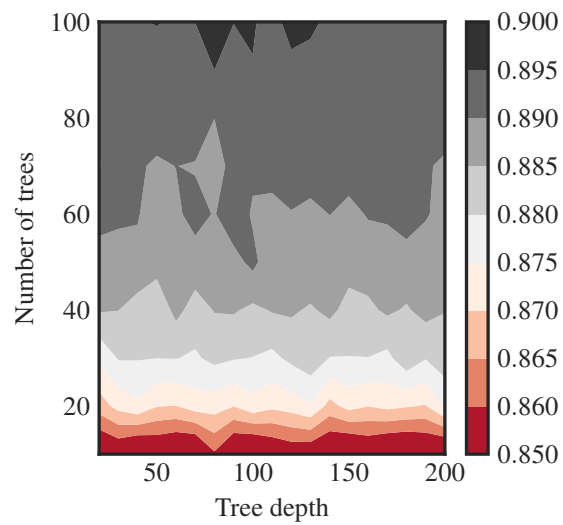


Figure J.7: The contour plot of the F_1 scores achieved for PAD classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.

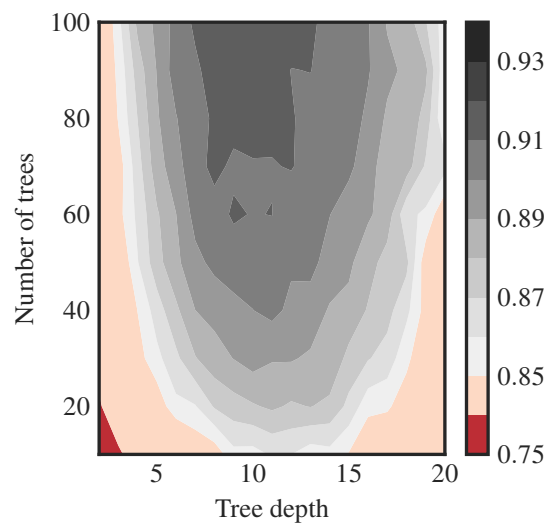


Figure J.8: The contour plot of the F_1 scores achieved for PAD classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.

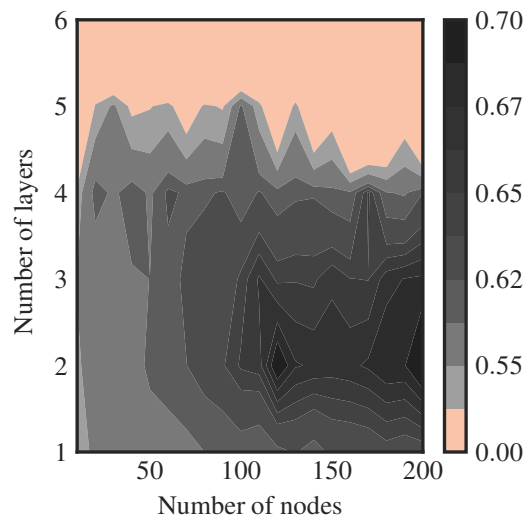


Figure J.9: The contour plot of the F_1 scores achieved for PAD classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.

GB, and MLP methods in Figures J.10, J.11, and J.12, respectively. As with the previous three forms of disease, the accuracy appears to be relatively insensitive to the combination of hyper-parameters used.

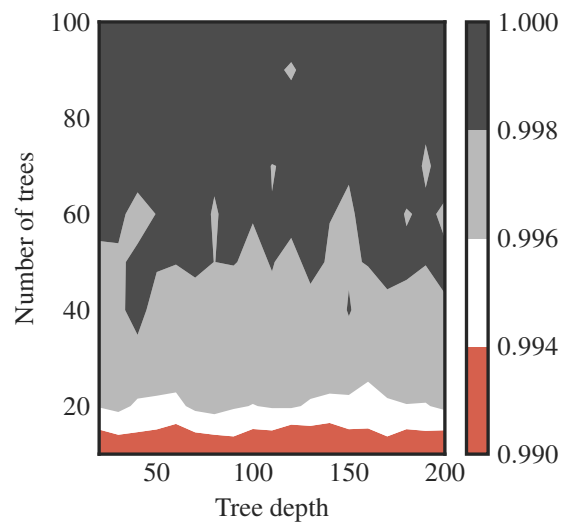


Figure J.10: The contour plot of the F_1 scores achieved for AAA classification when employing the RF method, using all six measurements, and different combinations of hyper-parameters.

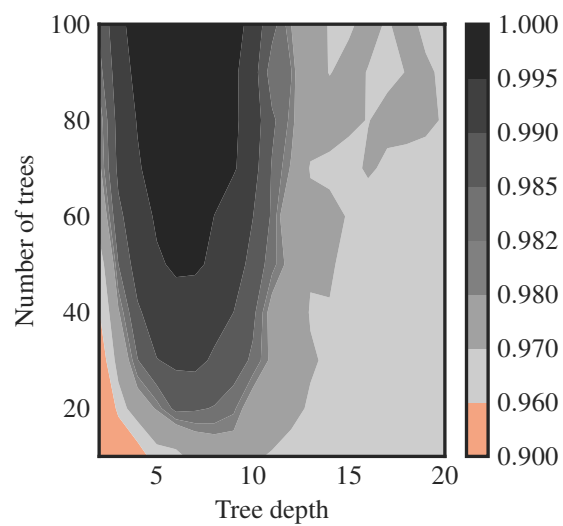


Figure J.11: The contour plot of the F_1 scores achieved for AAA classification when employing the GB method, using all six measurements, and different combinations of hyper-parameters.

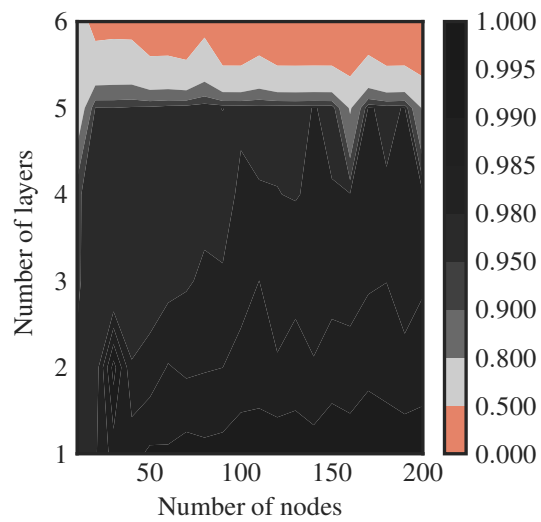


Figure J.12: The contour plot of the F_1 scores achieved for AAA classification when employing the MLP method, using all six measurements, and different combinations of hyper-parameters.

Appendix K

CAS combination search results

The F_1 scores, sensitivities, and specificities achieved for CAS classification when using each of the six ML methods are shown in Table K.1, K.2, and K.3 respectively.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.5547	0.5110	0.5157	0.5807	0.4365	0.5606
Q_2	0.5105	0.5080	0.4955	0.6858	0.4410	0.6565
Q_1	0.5676	0.5033	0.5953	0.8809	0.6459	0.8521
P_3	0.4927	0.5023	0.4991	0.5441	0.4805	0.5131
P_2	0.4413	0.5066	0.5260	0.5628	0.3741	0.5412
P_1	0.5473	0.4917	0.5712	0.6681	0.7013	0.7082
Q_3, Q_2	0.5684	0.4955	0.5104	0.6955	0.4915	0.6889
Q_3, Q_1	0.4831	0.5050	0.5544	0.8790	0.6944	0.8629
Q_3, P_3	0.5213	0.4935	0.5124	0.5825	0.4929	0.5659
Q_3, P_2	0.5853	0.5018	0.5142	0.5918	0.4904	0.5849
Q_3, P_1	0.5048	0.5034	0.5576	0.6601	0.6864	0.7105
Q_2, Q_1	0.4600	0.4975	0.5540	0.8913	0.6648	0.8824
Q_2, P_3	0.4804	0.4940	0.5109	0.6833	0.4158	0.6805
Q_2, P_2	0.5290	0.5037	0.5125	0.6836	0.5618	0.6908
Q_2, P_1	0.4434	0.4978	0.5597	0.7204	0.6741	0.7562
Q_1, P_3	0.4470	0.4990	0.5595	0.8732	0.6860	0.8577
Q_1, P_2	0.5341	0.5029	0.5629	0.8774	0.7090	0.8684
Q_1, P_1	0.4927	0.5018	0.6233	0.8837	0.7822	0.8850
P_3, P_2	0.5507	0.5117	0.5263	0.5581	0.5313	0.5431
P_3, P_1	0.5266	0.4963	0.5725	0.6837	0.7384	0.7539
P_2, P_1	0.5089	0.4944	0.6885	0.7938	0.8878	0.8950
Q_3, Q_2, Q_1	0.4299	0.4995	0.5425	0.8907	0.6838	0.8868
Q_3, Q_2, P_3	0.4822	0.4980	0.5058	0.6910	0.5300	0.7072
Q_3, Q_2, P_2	0.5346	0.4975	0.5204	0.6962	0.5211	0.7102
Q_3, Q_2, P_1	0.5267	0.5024	0.5428	0.7229	0.6084	0.7693
Q_3, Q_1, P_3	0.4636	0.5016	0.5317	0.8685	0.6699	0.8660
Q_3, Q_1, P_2	0.5186	0.4960	0.5580	0.8751	0.6469	0.8728
Q_3, Q_1, P_1	0.5257	0.5020	0.5888	0.8843	0.7532	0.8903
Q_3, P_3, P_2	0.4493	0.5032	0.5119	0.5923	0.5418	0.5888
Q_3, P_3, P_1	0.5019	0.4892	0.5527	0.6751	0.7159	0.7602
Q_3, P_2, P_1	0.4312	0.5042	0.6303	0.7564	0.8623	0.8923
Q_2, Q_1, P_3	0.5222	0.5041	0.5300	0.8840	0.6354	0.8776
Q_2, Q_1, P_2	0.5155	0.4957	0.5586	0.8847	0.7001	0.8844
Q_2, Q_1, P_1	0.5251	0.4940	0.6039	0.8941	0.7611	0.8968
Q_2, P_3, P_2	0.4893	0.5041	0.5241	0.6824	0.5335	0.6929
Q_2, P_3, P_1	0.4067	0.4965	0.5421	0.7249	0.7185	0.8064
Q_2, P_2, P_1	0.5479	0.4858	0.6415	0.7740	0.8735	0.9040
Q_1, P_3, P_2	0.4766	0.4969	0.5505	0.8700	0.7048	0.8651
Q_1, P_3, P_1	0.5037	0.4908	0.5975	0.8777	0.7645	0.8956
Q_1, P_2, P_1	0.4997	0.4972	0.6772	0.8872	0.8680	0.9389

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
P_3, P_2, P_1	0.5090	0.4997	0.6451	0.7694	0.8831	0.8936
Q_3, Q_2, Q_1, P_3	0.4569	0.4921	0.5408	0.8835	0.6258	0.8855
Q_3, Q_2, Q_1, P_2	0.4253	0.5022	0.5462	0.8871	0.6655	0.8887
Q_3, Q_2, Q_1, P_1	0.4934	0.5068	0.5783	0.8925	0.7163	0.9004
Q_3, Q_2, P_3, P_2	0.4875	0.5026	0.5234	0.6852	0.5483	0.7145
Q_3, Q_2, P_3, P_1	0.4481	0.4945	0.5399	0.7231	0.6714	0.8125
Q_3, Q_2, P_2, P_1	0.4329	0.5034	0.6043	0.7619	0.8618	0.9025
Q_3, Q_1, P_3, P_2	0.4934	0.4972	0.5400	0.8717	0.6299	0.8761
Q_3, Q_1, P_3, P_1	0.5365	0.5011	0.5802	0.8789	0.7197	0.8978
Q_3, Q_1, P_2, P_1	0.5068	0.4974	0.6338	0.8852	0.8542	0.9395
Q_3, P_3, P_2, P_1	0.4329	0.4980	0.6137	0.7393	0.8471	0.8906
Q_2, Q_1, P_3, P_2	0.5669	0.4933	0.5468	0.8822	0.6524	0.8844
Q_2, Q_1, P_3, P_1	0.5193	0.4978	0.5783	0.8878	0.7207	0.9065
Q_2, Q_1, P_2, P_1	0.4638	0.5037	0.6413	0.8944	0.8683	0.9383
Q_2, P_3, P_2, P_1	0.4868	0.4999	0.6142	0.7694	0.8503	0.9084
Q_1, P_3, P_2, P_1	0.4735	0.5025	0.6320	0.8807	0.8547	0.9353
Q_3, Q_2, Q_1, P_3, P_2	0.5005	0.5015	0.5387	0.8848	0.6322	0.8927
Q_3, Q_2, Q_1, P_3, P_1	0.4652	0.4962	0.5760	0.8875	0.7079	0.9093
Q_3, Q_2, Q_1, P_2, P_1	0.5108	0.4994	0.6088	0.8934	0.8313	0.9381
Q_3, Q_2, P_3, P_2, P_1	0.4994	0.5105	0.5808	0.7540	0.8463	0.9052
Q_3, Q_1, P_3, P_2, P_1	0.5330	0.5024	0.6108	0.8849	0.8380	0.9364
Q_2, Q_1, P_3, P_2, P_1	0.4899	0.5054	0.6026	0.8900	0.8371	0.9391
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4634	0.5018	0.5862	0.8878	0.7785	0.9343

Table K.1: The F_1 scores achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.1531	0.5527	0.4283	0.5572	0.6084	0.5736
Q_2	0.6641	0.5097	0.6418	0.6575	0.6228	0.6744
Q_1	0.5426	0.4525	0.5694	0.8704	0.4243	0.8547
P_3	0.5024	0.5098	0.4999	0.5139	0.5355	0.5158
P_2	0.6490	0.4979	0.5052	0.5366	0.7038	0.5491
P_1	0.2410	0.4992	0.6588	0.6510	0.7052	0.7215
Q_3, Q_2	0.5055	0.5035	0.5651	0.6655	0.5439	0.7044
Q_3, Q_1	0.7217	0.5094	0.5461	0.8681	0.6288	0.8661
Q_3, P_3	0.4462	0.5091	0.4944	0.5615	0.5226	0.5583
Q_3, P_2	0.3652	0.5090	0.5111	0.5731	0.5572	0.5777
Q_3, P_1	0.3032	0.5049	0.6510	0.6411	0.7334	0.7149
Q_2, Q_1	0.4543	0.5091	0.5770	0.8765	0.6372	0.8845
Q_2, P_3	0.5395	0.5117	0.5263	0.6555	0.6691	0.6926
Q_2, P_2	0.6043	0.4987	0.5436	0.6563	0.3867	0.7117
Q_2, P_1	0.5766	0.4905	0.6442	0.6962	0.6879	0.7668
Q_1, P_3	0.4718	0.5003	0.5461	0.8679	0.7167	0.8601
Q_1, P_2	0.3903	0.5068	0.5438	0.8672	0.7467	0.8664
Q_1, P_1	0.6488	0.4978	0.6672	0.8798	0.8290	0.8848
P_3, P_2	0.4744	0.5034	0.5251	0.5383	0.4863	0.5335
P_3, P_1	0.3334	0.4815	0.6303	0.6599	0.7469	0.7842
P_2, P_1	0.4943	0.5001	0.6699	0.7762	0.8995	0.9026
Q_3, Q_2, Q_1	0.4935	0.4812	0.5598	0.8789	0.6425	0.8896
Q_3, Q_2, P_3	0.5586	0.4913	0.5454	0.6647	0.5014	0.7219
Q_3, Q_2, P_2	0.4348	0.5091	0.5336	0.6788	0.5298	0.7313
Q_3, Q_2, P_1	0.2964	0.5165	0.6068	0.6975	0.6471	0.7783
Q_3, Q_1, P_3	0.4427	0.5021	0.5321	0.8620	0.6388	0.8677
Q_3, Q_1, P_2	0.4268	0.4974	0.5605	0.8718	0.6645	0.8721
Q_3, Q_1, P_1	0.5134	0.4687	0.6341	0.8735	0.7474	0.8936
Q_3, P_3, P_2	0.6823	0.4939	0.5076	0.5698	0.4853	0.5905
Q_3, P_3, P_1	0.3887	0.4941	0.5882	0.6691	0.6907	0.7866
Q_3, P_2, P_1	0.6288	0.4919	0.6251	0.7473	0.8593	0.8928
Q_2, Q_1, P_3	0.4599	0.5056	0.5474	0.8735	0.6999	0.8819
Q_2, Q_1, P_2	0.4296	0.5085	0.5728	0.8766	0.6703	0.8853
Q_2, Q_1, P_1	0.5235	0.5037	0.6393	0.8825	0.7635	0.8991
Q_2, P_3, P_2	0.3581	0.4902	0.5484	0.6566	0.5249	0.7098
Q_2, P_3, P_1	0.6042	0.4816	0.6079	0.7001	0.7140	0.8277
Q_2, P_2, P_1	0.5758	0.4826	0.6481	0.7516	0.8780	0.9114
Q_1, P_3, P_2	0.4530	0.4984	0.5521	0.8616	0.6733	0.8651
Q_1, P_3, P_1	0.4760	0.4859	0.6233	0.8721	0.7403	0.8993
Q_1, P_2, P_1	0.4610	0.4917	0.6744	0.8807	0.8797	0.9433
P_3, P_2, P_1	0.4001	0.5159	0.6442	0.7562	0.8731	0.8930
Q_3, Q_2, Q_1, P_3	0.5792	0.4903	0.5548	0.8700	0.6711	0.8916
Q_3, Q_2, Q_1, P_2	0.6211	0.4961	0.5726	0.8788	0.6622	0.8933
Q_3, Q_2, Q_1, P_1	0.5018	0.4831	0.5948	0.8837	0.7442	0.9015
Q_3, Q_2, P_3, P_2	0.5499	0.4981	0.4938	0.6722	0.4844	0.7277
Q_3, Q_2, P_3, P_1	0.4381	0.5052	0.5948	0.7055	0.6997	0.8337
Q_3, Q_2, P_2, P_1	0.7010	0.5049	0.6401	0.7463	0.8600	0.9129
Q_3, Q_1, P_3, P_2	0.4629	0.5000	0.5424	0.8597	0.6331	0.8747
Q_3, Q_1, P_3, P_1	0.4436	0.4937	0.6099	0.8747	0.7555	0.9017

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.5362	0.5062	0.6300	0.8761	0.8538	0.9417
Q_3, P_3, P_2, P_1	0.6073	0.5011	0.6165	0.7243	0.8391	0.8993
Q_2, Q_1, P_3, P_2	0.4973	0.5056	0.5779	0.8729	0.6427	0.8874
Q_2, Q_1, P_3, P_1	0.4225	0.5065	0.6115	0.8813	0.7596	0.9100
Q_2, Q_1, P_2, P_1	0.5115	0.4954	0.6345	0.8858	0.8618	0.9416
Q_2, P_3, P_2, P_1	0.5582	0.4877	0.6266	0.7498	0.8573	0.9133
Q_1, P_3, P_2, P_1	0.5769	0.4891	0.6309	0.8674	0.8667	0.9375
Q_3, Q_2, Q_1, P_3, P_2	0.5446	0.4929	0.5686	0.8759	0.6487	0.8949
Q_3, Q_2, Q_1, P_3, P_1	0.4676	0.4933	0.6021	0.8775	0.7169	0.9117
Q_3, Q_2, Q_1, P_2, P_1	0.5403	0.5015	0.6142	0.8858	0.8288	0.9415
Q_3, Q_2, P_3, P_2, P_1	0.6396	0.5042	0.6070	0.7375	0.8308	0.9120
Q_3, Q_1, P_3, P_2, P_1	0.5330	0.4920	0.6171	0.8795	0.8345	0.9399
Q_2, Q_1, P_3, P_2, P_1	0.4640	0.4919	0.6149	0.8774	0.8273	0.9416
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.6224	0.5116	0.6012	0.8747	0.7916	0.9364

Table K.2: The sensitivities achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.7090	0.4968	0.5462	0.5904	0.3886	0.5556
Q_2	0.4579	0.5075	0.4474	0.7006	0.3896	0.6479
Q_1	0.5776	0.5204	0.6063	0.8893	0.7517	0.8502
P_3	0.4896	0.4998	0.4989	0.5555	0.4632	0.5122
P_2	0.3826	0.5096	0.5335	0.5732	0.2983	0.5384
P_1	0.6628	0.4893	0.5363	0.6767	0.6993	0.7010
Q_3, Q_2	0.5935	0.4929	0.4917	0.7116	0.4745	0.6808
Q_3, Q_1	0.4072	0.5036	0.5576	0.8876	0.7293	0.8605
Q_3, P_3	0.5478	0.4884	0.5187	0.5912	0.4832	0.5690
Q_3, P_2	0.6764	0.4995	0.5153	0.5998	0.4687	0.5879
Q_3, P_1	0.5730	0.5030	0.5215	0.6696	0.6619	0.7081
Q_2, Q_1	0.4618	0.4937	0.5453	0.9032	0.6786	0.8808
Q_2, P_3	0.4618	0.4883	0.5057	0.6978	0.3494	0.6743
Q_2, P_2	0.5020	0.5055	0.5018	0.6978	0.6303	0.6798
Q_2, P_1	0.4055	0.5003	0.5269	0.7341	0.6672	0.7498
Q_1, P_3	0.4399	0.4987	0.5648	0.8774	0.6701	0.8560
Q_1, P_2	0.5865	0.5016	0.5704	0.8855	0.6883	0.8701
Q_1, P_1	0.4417	0.5032	0.6035	0.8869	0.7522	0.8853
P_3, P_2	0.5798	0.5146	0.5268	0.5659	0.5477	0.5467
P_3, P_1	0.5958	0.5013	0.5494	0.6961	0.7335	0.7356
P_2, P_1	0.5140	0.4926	0.6983	0.8054	0.8785	0.8889
Q_3, Q_2, Q_1	0.4125	0.5057	0.5361	0.9002	0.7054	0.8846
Q_3, Q_2, P_3	0.4580	0.5003	0.4925	0.7049	0.5404	0.6992
Q_3, Q_2, P_2	0.5711	0.4937	0.5158	0.7055	0.5181	0.6986
Q_3, Q_2, P_1	0.6091	0.4977	0.5190	0.7374	0.5915	0.7638
Q_3, Q_1, P_3	0.4700	0.5015	0.5316	0.8735	0.6857	0.8648
Q_3, Q_1, P_2	0.5508	0.4956	0.5571	0.8777	0.6386	0.8734
Q_3, Q_1, P_1	0.5302	0.5132	0.5699	0.8929	0.7568	0.8877
Q_3, P_3, P_2	0.3818	0.5064	0.5135	0.6018	0.5629	0.5882
Q_3, P_3, P_1	0.5399	0.4877	0.5392	0.6782	0.7300	0.7441
Q_3, P_2, P_1	0.3769	0.5084	0.6328	0.7620	0.8646	0.8920
Q_2, Q_1, P_3	0.5443	0.5037	0.5238	0.8924	0.6055	0.8743
Q_2, Q_1, P_2	0.5454	0.4916	0.5531	0.8913	0.7162	0.8838
Q_2, Q_1, P_1	0.5258	0.4909	0.5886	0.9035	0.7597	0.8950
Q_2, P_3, P_2	0.5319	0.5088	0.5156	0.6959	0.5367	0.6840
Q_2, P_3, P_1	0.3563	0.5015	0.5177	0.7390	0.7211	0.7921
Q_2, P_2, P_1	0.5374	0.4869	0.6385	0.7882	0.8701	0.8979
Q_1, P_3, P_2	0.4840	0.4965	0.5500	0.8766	0.7220	0.8651
Q_1, P_3, P_1	0.5131	0.4924	0.5866	0.8822	0.7795	0.8927
Q_1, P_2, P_1	0.5126	0.4991	0.6787	0.8925	0.8592	0.9351
P_3, P_2, P_1	0.5462	0.4944	0.6456	0.7777	0.8911	0.8942
Q_3, Q_2, Q_1, P_3	0.4208	0.4928	0.5357	0.8943	0.6053	0.8807
Q_3, Q_2, Q_1, P_2	0.3725	0.5043	0.5363	0.8938	0.6672	0.8852
Q_3, Q_2, Q_1, P_1	0.4907	0.5149	0.5716	0.8996	0.7008	0.8995
Q_3, Q_2, P_3, P_2	0.4675	0.5042	0.5340	0.6921	0.5725	0.7072
Q_3, Q_2, P_3, P_1	0.4510	0.4911	0.5196	0.7331	0.6572	0.7981
Q_3, Q_2, P_2, P_1	0.3589	0.5029	0.5888	0.7716	0.8632	0.8941
Q_3, Q_1, P_3, P_2	0.5034	0.4963	0.5392	0.8811	0.6285	0.8773
Q_3, Q_1, P_3, P_1	0.5706	0.5037	0.5681	0.8823	0.6997	0.8947

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.4969	0.4946	0.6357	0.8926	0.8545	0.9376
Q_3, P_3, P_2, P_1	0.3848	0.4970	0.6126	0.7481	0.8530	0.8837
Q_2, Q_1, P_3, P_2	0.5945	0.4893	0.5352	0.8896	0.6571	0.8822
Q_2, Q_1, P_3, P_1	0.5533	0.4950	0.5648	0.8930	0.6989	0.9037
Q_2, Q_1, P_2, P_1	0.4495	0.5065	0.6446	0.9015	0.8734	0.9355
Q_2, P_3, P_2, P_1	0.4639	0.5040	0.6088	0.7818	0.8452	0.9045
Q_1, P_3, P_2, P_1	0.4415	0.5070	0.6326	0.8912	0.8458	0.9335
Q_3, Q_2, Q_1, P_3, P_2	0.4859	0.5044	0.5277	0.8919	0.6246	0.8911
Q_3, Q_2, Q_1, P_3, P_1	0.4646	0.4972	0.5655	0.8956	0.7031	0.9073
Q_3, Q_2, Q_1, P_2, P_1	0.5008	0.4988	0.6065	0.8996	0.8332	0.9351
Q_3, Q_2, P_3, P_2, P_1	0.4528	0.5127	0.5701	0.7641	0.8577	0.8996
Q_3, Q_1, P_3, P_2, P_1	0.5331	0.5060	0.6081	0.8892	0.8406	0.9334
Q_2, Q_1, P_3, P_2, P_1	0.4984	0.5101	0.5974	0.9002	0.8442	0.9370
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4155	0.4986	0.5800	0.8984	0.7703	0.9325

Table K.3: The specificities achieved across the combination search by each of the six classification methods.

Appendix L

SAS combination search results

The F_1 scores, sensitivities, and specificities achieved for SAS classification when using each of the six ML methods are shown in Table L.1, L.2, and L.3 respectively.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.5041	0.5288	0.4897	0.5723	0.5403	0.5592
Q_2	0.4681	0.5004	0.4839	0.7577	0.5691	0.7415
Q_1	0.3799	0.5028	0.4923	0.7779	0.6176	0.7529
P_3	0.4931	0.4972	0.5097	0.5530	0.5474	0.5331
P_2	0.4698	0.4990	0.5528	0.5627	0.4895	0.5453
P_1	0.5344	0.5023	0.5035	0.5171	0.5571	0.5060
Q_3, Q_2	0.4529	0.5136	0.5075	0.7623	0.4939	0.7608
Q_3, Q_1	0.4588	0.4893	0.5053	0.7814	0.5414	0.7758
Q_3, P_3	0.4992	0.4963	0.5207	0.5824	0.5463	0.5746
Q_3, P_2	0.5497	0.5068	0.5306	0.5869	0.5215	0.5850
Q_3, P_1	0.4195	0.5099	0.4992	0.5685	0.4776	0.5627
Q_2, Q_1	0.5064	0.5010	0.5025	0.8450	0.5853	0.8461
Q_2, P_3	0.4818	0.5020	0.5294	0.7555	0.6054	0.7694
Q_2, P_2	0.5116	0.5020	0.5405	0.7586	0.5454	0.7711
Q_2, P_1	0.5468	0.4913	0.5353	0.7568	0.5124	0.7609
Q_1, P_3	0.4564	0.4963	0.5252	0.7697	0.5067	0.7522
Q_1, P_2	0.5209	0.4986	0.5388	0.7708	0.5833	0.7606
Q_1, P_1	0.5186	0.5005	0.5327	0.7744	0.5426	0.7751
P_3, P_2	0.5450	0.5031	0.5256	0.5695	0.4960	0.5626
P_3, P_1	0.5464	0.4996	0.5282	0.5450	0.5510	0.5338
P_2, P_1	0.5399	0.5041	0.5447	0.5669	0.5133	0.5766
Q_3, Q_2, Q_1	0.4574	0.5081	0.5284	0.8447	0.5866	0.8552
Q_3, Q_2, P_3	0.5499	0.4925	0.5254	0.7624	0.5847	0.7830
Q_3, Q_2, P_2	0.4591	0.4936	0.5272	0.7629	0.5742	0.7829
Q_3, Q_2, P_1	0.4240	0.4980	0.5099	0.7627	0.4969	0.7800
Q_3, Q_1, P_3	0.4810	0.4994	0.5173	0.7808	0.5511	0.7691
Q_3, Q_1, P_2	0.4098	0.5069	0.5354	0.7749	0.5611	0.7750
Q_3, Q_1, P_1	0.5414	0.4999	0.5095	0.7761	0.5230	0.7880
Q_3, P_3, P_2	0.4492	0.5021	0.5330	0.5892	0.5636	0.5900
Q_3, P_3, P_1	0.4912	0.4971	0.5248	0.5767	0.5253	0.5759
Q_3, P_2, P_1	0.4476	0.4914	0.5259	0.5883	0.5758	0.5961
Q_2, Q_1, P_3	0.5243	0.5008	0.5154	0.8381	0.5874	0.8427
Q_2, Q_1, P_2	0.4994	0.5029	0.5349	0.8402	0.6139	0.8469
Q_2, Q_1, P_1	0.4988	0.5042	0.5279	0.8413	0.5861	0.8492
Q_2, P_3, P_2	0.5272	0.4992	0.5284	0.7549	0.5760	0.7802
Q_2, P_3, P_1	0.4351	0.5048	0.5351	0.7479	0.5724	0.7726
Q_2, P_2, P_1	0.5318	0.5081	0.5316	0.7563	0.5258	0.7752
Q_1, P_3, P_2	0.5152	0.5030	0.5454	0.7624	0.5782	0.7579
Q_1, P_3, P_1	0.4607	0.5022	0.5235	0.7690	0.5069	0.7680
Q_1, P_2, P_1	0.5437	0.5019	0.5319	0.7670	0.5930	0.7733
P_3, P_2, P_1	0.5314	0.4984	0.5352	0.5661	0.5518	0.5826

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_2, Q_1, P_3	0.4910	0.4925	0.5169	0.8407	0.5706	0.8541
Q_3, Q_2, Q_1, P_2	0.5113	0.5036	0.5301	0.8432	0.5952	0.8585
Q_3, Q_2, Q_1, P_1	0.5097	0.5078	0.5191	0.8404	0.5828	0.8558
Q_3, Q_2, P_3, P_2	0.4738	0.4968	0.5206	0.7549	0.5628	0.7879
Q_3, Q_2, P_3, P_1	0.4721	0.4944	0.5224	0.7545	0.5605	0.7857
Q_3, Q_2, P_2, P_1	0.5592	0.5081	0.5331	0.7616	0.5854	0.7911
Q_3, Q_1, P_3, P_2	0.4762	0.4987	0.5259	0.7738	0.5791	0.7711
Q_3, Q_1, P_3, P_1	0.4558	0.5108	0.5339	0.7749	0.5766	0.7850
Q_3, Q_1, P_2, P_1	0.4066	0.4957	0.5279	0.7719	0.5785	0.7813
Q_3, P_3, P_2, P_1	0.5257	0.4878	0.5395	0.5866	0.5695	0.5988
Q_2, Q_1, P_3, P_2	0.5318	0.4975	0.5487	0.8357	0.6064	0.8488
Q_2, Q_1, P_3, P_1	0.5348	0.4987	0.5326	0.8350	0.5879	0.8516
Q_2, Q_1, P_2, P_1	0.5537	0.5113	0.5337	0.8362	0.6258	0.8545
Q_2, P_3, P_2, P_1	0.4863	0.4966	0.5394	0.7458	0.6102	0.7797
Q_1, P_3, P_2, P_1	0.4711	0.5010	0.5358	0.7635	0.6088	0.7738
Q_3, Q_2, Q_1, P_3, P_2	0.4763	0.5038	0.5312	0.8330	0.5966	0.8534
Q_3, Q_2, Q_1, P_3, P_1	0.4953	0.4998	0.5212	0.8399	0.5809	0.8571
Q_3, Q_2, Q_1, P_2, P_1	0.4917	0.5099	0.5304	0.8390	0.6070	0.8600
Q_3, Q_2, P_3, P_2, P_1	0.5344	0.5069	0.5292	0.7540	0.5963	0.7913
Q_3, Q_1, P_3, P_2, P_1	0.5205	0.4991	0.5309	0.7734	0.5740	0.7828
Q_2, Q_1, P_3, P_2, P_1	0.4912	0.5012	0.5353	0.8325	0.6302	0.8502
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4642	0.5016	0.5301	0.8292	0.6040	0.8574

Table L.1: The F_1 scores achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.2997	0.4576	0.5129	0.5678	0.4059	0.5585
Q_2	0.5460	0.5348	0.6918	0.7517	0.3839	0.7366
Q_1	0.7074	0.4613	0.6338	0.7582	0.1873	0.7224
P_3	0.4402	0.5127	0.5616	0.5453	0.3978	0.5431
P_2	0.5140	0.4981	0.4783	0.5629	0.5704	0.5717
P_1	0.4446	0.4836	0.4741	0.5177	0.3803	0.5244
Q_3, Q_2	0.5683	0.4928	0.5411	0.7612	0.5901	0.7585
Q_3, Q_1	0.4887	0.4947	0.5036	0.7630	0.4709	0.7504
Q_3, P_3	0.6479	0.5019	0.5147	0.5720	0.4578	0.5808
Q_3, P_2	0.5719	0.4985	0.5163	0.5849	0.5223	0.5999
Q_3, P_1	0.6081	0.4947	0.4958	0.5633	0.5570	0.5788
Q_2, Q_1	0.6572	0.5008	0.6082	0.8374	0.4909	0.8293
Q_2, P_3	0.5785	0.4860	0.5626	0.7505	0.4320	0.7710
Q_2, P_2	0.4241	0.4801	0.5294	0.7560	0.6660	0.7763
Q_2, P_1	0.2405	0.5006	0.5127	0.7500	0.5838	0.7601
Q_1, P_3	0.5330	0.4970	0.5596	0.7534	0.5809	0.7305
Q_1, P_2	0.4943	0.5180	0.5282	0.7545	0.4884	0.7434
Q_1, P_1	0.5761	0.4991	0.5430	0.7516	0.6004	0.7549
P_3, P_2	0.4714	0.4939	0.5388	0.5668	0.6677	0.5744
P_3, P_1	0.5408	0.4954	0.5252	0.5406	0.4456	0.5421
P_2, P_1	0.4115	0.4958	0.4761	0.5761	0.6175	0.6056
Q_3, Q_2, Q_1	0.5695	0.5019	0.5106	0.8271	0.5303	0.8453
Q_3, Q_2, P_3	0.5651	0.5115	0.5075	0.7621	0.5044	0.7826
Q_3, Q_2, P_2	0.5768	0.5219	0.5101	0.7590	0.5941	0.7882
Q_3, Q_2, P_1	0.6416	0.5013	0.5350	0.7494	0.5963	0.7766
Q_3, Q_1, P_3	0.4649	0.5074	0.5237	0.7550	0.5783	0.7491
Q_3, Q_1, P_2	0.6031	0.50	0.5056	0.7584	0.5796	0.7514
Q_3, Q_1, P_1	0.3262	0.4942	0.5535	0.7527	0.6028	0.7677
Q_3, P_3, P_2	0.5316	0.4904	0.5184	0.5924	0.4985	0.6109
Q_3, P_3, P_1	0.3543	0.4949	0.5116	0.5765	0.5444	0.5855
Q_3, P_2, P_1	0.5225	0.5038	0.5041	0.5864	0.5018	0.6186
Q_2, Q_1, P_3	0.4531	0.4826	0.5427	0.8186	0.6309	0.8303
Q_2, Q_1, P_2	0.4642	0.5029	0.5481	0.8277	0.6178	0.8312
Q_2, Q_1, P_1	0.5179	0.5049	0.5544	0.8268	0.5788	0.8388
Q_2, P_3, P_2	0.5155	0.4806	0.5642	0.7500	0.6050	0.7757
Q_2, P_3, P_1	0.6119	0.4972	0.5365	0.7486	0.5358	0.7752
Q_2, P_2, P_1	0.5590	0.5214	0.5403	0.7578	0.7119	0.7791
Q_1, P_3, P_2	0.4890	0.5159	0.5345	0.7414	0.5886	0.7437
Q_1, P_3, P_1	0.5256	0.5041	0.5548	0.7498	0.6421	0.7479
Q_1, P_2, P_1	0.4038	0.5014	0.5175	0.7490	0.5995	0.7621
P_3, P_2, P_1	0.4461	0.4995	0.5216	0.5697	0.6360	0.6026
Q_3, Q_2, Q_1, P_3	0.6262	0.5155	0.5310	0.8274	0.6144	0.8411
Q_3, Q_2, Q_1, P_2	0.4646	0.5158	0.5531	0.8303	0.6113	0.8487
Q_3, Q_2, Q_1, P_1	0.4913	0.5011	0.5522	0.8242	0.5723	0.8466
Q_3, Q_2, P_3, P_2	0.5435	0.4831	0.54	0.7566	0.64	0.7924
Q_3, Q_2, P_3, P_1	0.5466	0.4884	0.5173	0.7534	0.5521	0.7874
Q_3, Q_2, P_2, P_1	0.4776	0.5022	0.5413	0.7555	0.5892	0.7900
Q_3, Q_1, P_3, P_2	0.5274	0.5010	0.5377	0.7587	0.5758	0.7545
Q_3, Q_1, P_3, P_1	0.4177	0.4823	0.5051	0.7560	0.5163	0.7675

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.5806	0.5103	0.5087	0.7550	0.5940	0.7735
Q_3, P_3, P_2, P_1	0.46	0.5052	0.5204	0.5857	0.6047	0.6121
Q_2, Q_1, P_3, P_2	0.4529	0.5117	0.5461	0.8241	0.6431	0.8413
Q_2, Q_1, P_3, P_1	0.2714	0.4964	0.5150	0.8186	0.6153	0.8437
Q_2, Q_1, P_2, P_1	0.5132	0.5057	0.5357	0.8214	0.6157	0.8386
Q_2, P_3, P_2, P_1	0.4464	0.5042	0.5606	0.7407	0.6294	0.7833
Q_1, P_3, P_2, P_1	0.4715	0.5032	0.5476	0.7439	0.6014	0.7599
Q_3, Q_2, Q_1, P_3, P_2	0.44	0.4889	0.5266	0.8175	0.5881	0.8510
Q_3, Q_2, Q_1, P_3, P_1	0.3896	0.4988	0.5447	0.8256	0.6080	0.8443
Q_3, Q_2, Q_1, P_2, P_1	0.5676	0.50	0.5270	0.8274	0.6084	0.8525
Q_3, Q_2, P_3, P_2, P_1	0.4376	0.5137	0.5454	0.7499	0.6264	0.7859
Q_3, Q_1, P_3, P_2, P_1	0.4463	0.4941	0.5332	0.7509	0.6137	0.7634
Q_2, Q_1, P_3, P_2, P_1	0.6175	0.4940	0.5561	0.8159	0.5996	0.8451
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.5047	0.4996	0.5342	0.8102	0.6133	0.8504

Table L.2: The sensitivities achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.5731	0.5544	0.4823	0.5742	0.5901	0.5596
Q_2	0.4444	0.4890	0.4176	0.7615	0.6429	0.7445
Q_1	0.3032	0.5168	0.4462	0.7905	0.8099	0.7714
P_3	0.5105	0.4921	0.4920	0.5560	0.6038	0.5295
P_2	0.4563	0.4993	0.5814	0.5627	0.4633	0.5355
P_1	0.5672	0.5087	0.5135	0.5170	0.6254	0.4999
Q_3, Q_2	0.4192	0.5208	0.4962	0.7631	0.4624	0.7623
Q_3, Q_1	0.4499	0.4876	0.5059	0.7932	0.5676	0.7920
Q_3, P_3	0.4498	0.4945	0.5229	0.5867	0.5796	0.5722
Q_3, P_2	0.5414	0.5097	0.5359	0.5878	0.5213	0.5789
Q_3, P_1	0.3695	0.5152	0.5004	0.5706	0.4528	0.5565
Q_2, Q_1	0.4553	0.5012	0.4671	0.8507	0.6244	0.8585
Q_2, P_3	0.4512	0.5074	0.5175	0.7586	0.6807	0.7685
Q_2, P_2	0.5418	0.5094	0.5447	0.7603	0.5002	0.7679
Q_2, P_1	0.6622	0.4884	0.5436	0.7610	0.4879	0.7614
Q_1, P_3	0.4338	0.4961	0.5130	0.7800	0.4816	0.7653
Q_1, P_2	0.5303	0.4922	0.5428	0.7811	0.6224	0.7712
Q_1, P_1	0.4985	0.5010	0.5290	0.7889	0.5211	0.7880
P_3, P_2	0.5726	0.5062	0.5209	0.5707	0.4394	0.5581
P_3, P_1	0.5486	0.5011	0.5293	0.5467	0.5912	0.5309
P_2, P_1	0.5874	0.5069	0.5704	0.5633	0.4774	0.5649
Q_3, Q_2, Q_1	0.4242	0.5103	0.5348	0.8576	0.6101	0.8626
Q_3, Q_2, P_3	0.5442	0.4864	0.5319	0.7626	0.6180	0.7834
Q_3, Q_2, P_2	0.4241	0.4844	0.5334	0.7654	0.5662	0.7795
Q_3, Q_2, P_1	0.3655	0.4970	0.5014	0.7710	0.4641	0.7822
Q_3, Q_1, P_3	0.4862	0.4968	0.5152	0.7974	0.5408	0.7816
Q_3, Q_1, P_2	0.3600	0.5093	0.5464	0.7854	0.5539	0.7900
Q_3, Q_1, P_1	0.6213	0.5018	0.4945	0.7911	0.4948	0.8013
Q_3, P_3, P_2	0.4254	0.5061	0.5384	0.5879	0.5892	0.5813
Q_3, P_3, P_1	0.5358	0.4979	0.5295	0.5769	0.5186	0.5721
Q_3, P_2, P_1	0.4261	0.4874	0.5338	0.5892	0.6058	0.5866
Q_2, Q_1, P_3	0.5497	0.5069	0.5060	0.8522	0.5694	0.8519
Q_2, Q_1, P_2	0.5112	0.5029	0.5301	0.8494	0.6123	0.8585
Q_2, Q_1, P_1	0.4925	0.5040	0.5184	0.8519	0.5892	0.8569
Q_2, P_3, P_2	0.5315	0.5055	0.5156	0.7579	0.5643	0.7831
Q_2, P_3, P_1	0.3860	0.5075	0.5347	0.7476	0.5871	0.7710
Q_2, P_2, P_1	0.5220	0.5036	0.5285	0.7555	0.4595	0.7728
Q_1, P_3, P_2	0.5244	0.4987	0.5495	0.7755	0.5740	0.7667
Q_1, P_3, P_1	0.4414	0.5016	0.5125	0.7810	0.4611	0.7806
Q_1, P_2, P_1	0.5960	0.5021	0.5372	0.7782	0.5904	0.7804
P_3, P_2, P_1	0.5624	0.4981	0.5403	0.5647	0.5198	0.5745
Q_3, Q_2, Q_1, P_3	0.4471	0.4850	0.5120	0.8504	0.5532	0.8638
Q_3, Q_2, Q_1, P_2	0.5274	0.4996	0.5219	0.8527	0.5884	0.8660
Q_3, Q_2, Q_1, P_1	0.5160	0.5102	0.5076	0.8522	0.5872	0.8627
Q_3, Q_2, P_3, P_2	0.4522	0.5014	0.5138	0.7540	0.5326	0.7851
Q_3, Q_2, P_3, P_1	0.4492	0.4964	0.5243	0.7553	0.5639	0.7847
Q_3, Q_2, P_2, P_1	0.5909	0.5102	0.5302	0.7654	0.5839	0.7919
Q_3, Q_1, P_3, P_2	0.4603	0.4980	0.5218	0.7834	0.5805	0.7816
Q_3, Q_1, P_3, P_1	0.4671	0.5206	0.5444	0.7869	0.6011	0.7964

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.3623	0.4910	0.5348	0.7826	0.5723	0.7864
Q_3, P_3, P_2, P_1	0.5492	0.4823	0.5466	0.5870	0.5555	0.5932
Q_2, Q_1, P_3, P_2	0.5604	0.4929	0.5497	0.8441	0.5905	0.8545
Q_2, Q_1, P_3, P_1	0.6311	0.4995	0.5390	0.8468	0.5765	0.8575
Q_2, Q_1, P_2, P_1	0.5693	0.5133	0.5330	0.8469	0.6304	0.8664
Q_2, P_3, P_2, P_1	0.4992	0.4941	0.5316	0.7489	0.6018	0.7774
Q_1, P_3, P_2, P_1	0.4710	0.5003	0.5315	0.7757	0.6121	0.7826
Q_3, Q_2, Q_1, P_3, P_2	0.4877	0.5089	0.5329	0.8442	0.6003	0.8552
Q_3, Q_2, Q_1, P_3, P_1	0.5301	0.5002	0.5130	0.8504	0.5699	0.8668
Q_3, Q_2, Q_1, P_2, P_1	0.4670	0.5133	0.5317	0.8475	0.6064	0.8657
Q_3, Q_2, P_3, P_2, P_1	0.5697	0.5046	0.5234	0.7566	0.5836	0.7950
Q_3, Q_1, P_3, P_2, P_1	0.5467	0.5008	0.5302	0.7876	0.5581	0.7954
Q_2, Q_1, P_3, P_2, P_1	0.4502	0.5037	0.5278	0.8444	0.6443	0.8540
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4520	0.5023	0.5287	0.8427	0.60	0.8627

Table L.3: The specificities achieved across the combination search by each of the six classification methods.

Appendix M

PAD combination search results

The F_1 scores, sensitivities, and specificities achieved for PAD classification when using each of the six ML methods are shown in Table M.1, M.2, and M.3 respectively.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.5017	0.5115	0.6645	0.8224	0.6897	0.8169
Q_2	0.5621	0.5222	0.5266	0.7127	0.4734	0.7076
Q_1	0.3927	0.4822	0.5310	0.8240	0.4713	0.8183
P_3	0.5162	0.5053	0.5182	0.5613	0.4131	0.5406
P_2	0.5030	0.4954	0.5242	0.5753	0.4741	0.5529
P_1	0.4290	0.5031	0.5038	0.5517	0.5487	0.5335
Q_3, Q_2	0.4740	0.5099	0.5926	0.8480	0.7040	0.8557
Q_3, Q_1	0.5355	0.4965	0.5786	0.8959	0.7254	0.9041
Q_3, P_3	0.4800	0.4932	0.5808	0.8050	0.6676	0.8151
Q_3, P_2	0.5118	0.4998	0.5824	0.8152	0.7057	0.8201
Q_3, P_1	0.5672	0.4979	0.5768	0.8103	0.7206	0.8221
Q_2, Q_1	0.5236	0.4962	0.5239	0.8556	0.5610	0.8637
Q_2, P_3	0.4929	0.4980	0.5069	0.7134	0.6117	0.7200
Q_2, P_2	0.5323	0.4956	0.5133	0.7126	0.5233	0.7255
Q_2, P_1	0.4602	0.5075	0.5222	0.7117	0.5585	0.7221
Q_1, P_3	0.5293	0.5116	0.5420	0.8136	0.5602	0.8204
Q_1, P_2	0.5335	0.4926	0.5406	0.8187	0.5818	0.8314
Q_1, P_1	0.5549	0.5011	0.5417	0.8181	0.6514	0.8307
P_3, P_2	0.4829	0.4996	0.5319	0.5810	0.5386	0.5733
P_3, P_1	0.4823	0.4976	0.5142	0.5624	0.5141	0.5559
P_2, P_1	0.5434	0.5035	0.5145	0.5904	0.4662	0.6002
Q_3, Q_2, Q_1	0.5209	0.4891	0.5619	0.9061	0.7004	0.9168
Q_3, Q_2, P_3	0.4717	0.5146	0.5605	0.8370	0.6864	0.8556
Q_3, Q_2, P_2	0.4651	0.5049	0.5640	0.8424	0.7074	0.8606
Q_3, Q_2, P_1	0.4643	0.5064	0.5610	0.8408	0.7040	0.8592
Q_3, Q_1, P_3	0.4947	0.4976	0.5679	0.8833	0.7148	0.9009
Q_3, Q_1, P_2	0.5615	0.4984	0.5741	0.8858	0.7100	0.9022
Q_3, Q_1, P_1	0.4149	0.4941	0.5760	0.8850	0.7361	0.9046
Q_3, P_3, P_2	0.4800	0.5065	0.5598	0.8005	0.6804	0.8215
Q_3, P_3, P_1	0.5214	0.5050	0.5642	0.8005	0.6886	0.8179
Q_3, P_2, P_1	0.4792	0.5065	0.5630	0.8004	0.7104	0.8178
Q_2, Q_1, P_3	0.5208	0.5006	0.5334	0.8469	0.6300	0.8617
Q_2, Q_1, P_2	0.4874	0.4974	0.5318	0.8472	0.5992	0.8703
Q_2, Q_1, P_1	0.5340	0.4938	0.5311	0.8472	0.6472	0.8682
Q_2, P_3, P_2	0.5306	0.4996	0.5162	0.7147	0.5581	0.7379
Q_2, P_3, P_1	0.5012	0.4989	0.5152	0.7062	0.5165	0.7311
Q_2, P_2, P_1	0.5165	0.4983	0.5232	0.7118	0.5659	0.7322
Q_1, P_3, P_2	0.5324	0.4941	0.5382	0.8086	0.6117	0.8302
Q_1, P_3, P_1	0.4632	0.5047	0.5322	0.8116	0.6127	0.8324
Q_1, P_2, P_1	0.4524	0.4930	0.5429	0.8146	0.6441	0.8380
P_3, P_2, P_1	0.5016	0.5023	0.5262	0.5838	0.5654	0.6078

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_2, Q_1, P_3	0.5480	0.5086	0.5600	0.8992	0.6988	0.9138
Q_3, Q_2, Q_1, P_2	0.4505	0.4997	0.5564	0.8997	0.7017	0.9164
Q_3, Q_2, Q_1, P_1	0.4973	0.5053	0.5601	0.8990	0.7030	0.9196
Q_3, Q_2, P_3, P_2	0.3998	0.4993	0.5601	0.8376	0.6688	0.8612
Q_3, Q_2, P_3, P_1	0.5253	0.4973	0.5558	0.8330	0.6738	0.8556
Q_3, Q_2, P_2, P_1	0.4726	0.4972	0.5650	0.8385	0.6811	0.8597
Q_3, Q_1, P_3, P_2	0.5030	0.4976	0.5684	0.8803	0.6845	0.8999
Q_3, Q_1, P_3, P_1	0.5189	0.5019	0.5595	0.8839	0.6849	0.9013
Q_3, Q_1, P_2, P_1	0.5692	0.4994	0.5715	0.8805	0.6962	0.9025
Q_3, P_3, P_2, P_1	0.4801	0.4991	0.5576	0.7940	0.6746	0.8170
Q_2, Q_1, P_3, P_2	0.4681	0.4966	0.5404	0.8417	0.6239	0.8624
Q_2, Q_1, P_3, P_1	0.5009	0.5015	0.5278	0.8378	0.6146	0.8677
Q_2, Q_1, P_2, P_1	0.5278	0.4979	0.5304	0.8433	0.6327	0.8690
Q_2, P_3, P_2, P_1	0.5242	0.5024	0.5180	0.7022	0.5806	0.7376
Q_1, P_3, P_2, P_1	0.4996	0.5033	0.5355	0.8087	0.6158	0.8328
Q_3, Q_2, Q_1, P_3, P_2	0.5012	0.5006	0.5495	0.8971	0.6889	0.9169
Q_3, Q_2, Q_1, P_3, P_1	0.5025	0.4969	0.5562	0.8952	0.6887	0.9151
Q_3, Q_2, Q_1, P_2, P_1	0.5023	0.5019	0.5502	0.8969	0.6895	0.9170
Q_3, Q_2, P_3, P_2, P_1	0.4946	0.4923	0.5488	0.8279	0.6545	0.8597
Q_3, Q_1, P_3, P_2, P_1	0.4489	0.4972	0.5666	0.8758	0.6688	0.9042
Q_2, Q_1, P_3, P_2, P_1	0.5377	0.4995	0.5391	0.8389	0.6154	0.8655
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4479	0.4974	0.5573	0.8935	0.6681	0.9187

Table M.1: The F_1 scores achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.3598	0.5048	0.6806	0.8219	0.5998	0.8188
Q_2	0.5441	0.4878	0.5879	0.6858	0.5536	0.6922
Q_1	0.5735	0.5026	0.6065	0.8126	0.5959	0.8140
P_3	0.4246	0.4935	0.5472	0.5358	0.6388	0.5425
P_2	0.4565	0.4985	0.5368	0.5532	0.5572	0.5576
P_1	0.6253	0.5001	0.5571	0.5245	0.3899	0.5261
Q_3, Q_2	0.5595	0.4912	0.6297	0.8414	0.7176	0.8532
Q_3, Q_1	0.4753	0.5087	0.6324	0.8825	0.7460	0.8950
Q_3, P_3	0.6086	0.5025	0.5980	0.8021	0.6523	0.8173
Q_3, P_2	0.3310	0.4895	0.5919	0.8089	0.7679	0.8269
Q_3, P_1	0.3079	0.5280	0.6021	0.8051	0.7461	0.8266
Q_2, Q_1	0.4323	0.4902	0.5878	0.8346	0.6016	0.8521
Q_2, P_3	0.5419	0.4877	0.5744	0.6826	0.2813	0.7126
Q_2, P_2	0.5505	0.5051	0.5776	0.6862	0.5169	0.7275
Q_2, P_1	0.6100	0.4976	0.5697	0.6875	0.4716	0.7127
Q_1, P_3	0.3309	0.4971	0.5476	0.8001	0.5911	0.8168
Q_1, P_2	0.5495	0.5063	0.5827	0.8019	0.5508	0.8288
Q_1, P_1	0.3834	0.4930	0.5778	0.8059	0.6787	0.8272
P_3, P_2	0.4789	0.4946	0.5458	0.5569	0.5443	0.5709
P_3, P_1	0.5309	0.5066	0.5642	0.5425	0.5406	0.5484
P_2, P_1	0.5325	0.4961	0.5863	0.5651	0.6096	0.5998
Q_3, Q_2, Q_1	0.4948	0.5163	0.5976	0.8885	0.7801	0.9055
Q_3, Q_2, P_3	0.3895	0.4985	0.5568	0.8323	0.7286	0.8572
Q_3, Q_2, P_2	0.5612	0.5051	0.5851	0.8388	0.6953	0.8545
Q_3, Q_2, P_1	0.4521	0.4890	0.5787	0.8278	0.7259	0.8559
Q_3, Q_1, P_3	0.5637	0.5045	0.5826	0.8707	0.7050	0.8913
Q_3, Q_1, P_2	0.4240	0.5030	0.5974	0.8710	0.7409	0.8923
Q_3, Q_1, P_1	0.6578	0.5094	0.6104	0.8663	0.6902	0.8928
Q_3, P_3, P_2	0.3869	0.4995	0.5834	0.7984	0.6967	0.8211
Q_3, P_3, P_1	0.2820	0.5009	0.5706	0.7914	0.6994	0.8208
Q_3, P_2, P_1	0.5814	0.4880	0.5824	0.7970	0.6789	0.8163
Q_2, Q_1, P_3	0.3260	0.4775	0.5663	0.8303	0.5969	0.8540
Q_2, Q_1, P_2	0.4239	0.4959	0.5625	0.8309	0.6028	0.8636
Q_2, Q_1, P_1	0.3205	0.5176	0.5610	0.8289	0.6418	0.8595
Q_2, P_3, P_2	0.4276	0.4900	0.5714	0.6920	0.5968	0.7328
Q_2, P_3, P_1	0.5554	0.4896	0.5560	0.6859	0.6252	0.7136
Q_2, P_2, P_1	0.4250	0.5134	0.5664	0.6845	0.5546	0.7245
Q_1, P_3, P_2	0.5668	0.4987	0.5330	0.7935	0.5752	0.8208
Q_1, P_3, P_1	0.4876	0.5104	0.5537	0.7998	0.6082	0.8287
Q_1, P_2, P_1	0.6109	0.4885	0.5572	0.8022	0.5978	0.8313
P_3, P_2, P_1	0.3959	0.4901	0.5652	0.5688	0.5532	0.6022
Q_3, Q_2, Q_1, P_3	0.3678	0.4879	0.5510	0.8819	0.7136	0.9035
Q_3, Q_2, Q_1, P_2	0.4522	0.5111	0.5909	0.8868	0.7224	0.9085
Q_3, Q_2, Q_1, P_1	0.5593	0.4867	0.5680	0.8846	0.7250	0.9068
Q_3, Q_2, P_3, P_2	0.5688	0.4972	0.5879	0.8231	0.7166	0.8574
Q_3, Q_2, P_3, P_1	0.4517	0.5112	0.5707	0.8201	0.7036	0.8504
Q_3, Q_2, P_2, P_1	0.5414	0.4904	0.5642	0.8247	0.7091	0.8526
Q_3, Q_1, P_3, P_2	0.6603	0.4851	0.5512	0.8655	0.7055	0.8936
Q_3, Q_1, P_3, P_1	0.3708	0.4993	0.5781	0.8655	0.7178	0.8951

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.4094	0.4967	0.5752	0.8612	0.7042	0.8926
Q_3, P_3, P_2, P_1	0.5180	0.5097	0.5724	0.7834	0.6593	0.8182
Q_2, Q_1, P_3, P_2	0.3984	0.4901	0.5564	0.8199	0.6451	0.8568
Q_2, Q_1, P_3, P_1	0.3787	0.5159	0.5556	0.8243	0.6639	0.8587
Q_2, Q_1, P_2, P_1	0.4432	0.5153	0.5587	0.8324	0.6442	0.8633
Q_2, P_3, P_2, P_1	0.4612	0.4878	0.5385	0.6811	0.5837	0.7262
Q_1, P_3, P_2, P_1	0.4762	0.4917	0.5679	0.7953	0.6449	0.8315
Q_3, Q_2, Q_1, P_3, P_2	0.3675	0.5049	0.5659	0.8802	0.6844	0.9133
Q_3, Q_2, Q_1, P_3, P_1	0.3552	0.4925	0.5784	0.8766	0.6848	0.9073
Q_3, Q_2, Q_1, P_2, P_1	0.4635	0.4996	0.5754	0.8829	0.6910	0.9041
Q_3, Q_2, P_3, P_2, P_1	0.4797	0.5169	0.5518	0.8142	0.6891	0.8544
Q_3, Q_1, P_3, P_2, P_1	0.5274	0.5069	0.5507	0.8625	0.6738	0.8986
Q_2, Q_1, P_3, P_2, P_1	0.3947	0.4911	0.5493	0.8258	0.6190	0.8556
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.6385	0.4859	0.5511	0.8813	0.6588	0.9102

Table M.2: The sensitivities achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.5493	0.5139	0.6566	0.8228	0.7371	0.8157
Q_2	0.5692	0.5344	0.5047	0.7276	0.4486	0.7161
Q_1	0.3486	0.4758	0.5038	0.8320	0.4329	0.8214
P_3	0.5481	0.5093	0.5081	0.5713	0.3544	0.5399
P_2	0.5187	0.4945	0.5198	0.5843	0.4484	0.5512
P_1	0.3754	0.5042	0.4859	0.5622	0.6088	0.5362
Q_3, Q_2	0.4475	0.5164	0.5770	0.8529	0.6967	0.8576
Q_3, Q_1	0.5576	0.4926	0.5568	0.9068	0.7137	0.9117
Q_3, P_3	0.4395	0.4902	0.5738	0.8070	0.6754	0.8137
Q_3, P_2	0.5740	0.5033	0.5785	0.8196	0.6718	0.8155
Q_3, P_1	0.6699	0.4880	0.5666	0.8140	0.7063	0.8190
Q_2, Q_1	0.5561	0.4983	0.5013	0.8714	0.5452	0.8726
Q_2, P_3	0.4769	0.5015	0.4840	0.7305	0.7573	0.7243
Q_2, P_2	0.5257	0.4926	0.4912	0.7273	0.5257	0.7245
Q_2, P_1	0.4155	0.5109	0.5055	0.7252	0.5922	0.7275
Q_1, P_3	0.6008	0.5166	0.5400	0.8229	0.5482	0.8230
Q_1, P_2	0.5277	0.4882	0.5251	0.8305	0.5946	0.8334
Q_1, P_1	0.6209	0.5039	0.5284	0.8266	0.6383	0.8333
P_3, P_2	0.4842	0.5013	0.5269	0.5910	0.5365	0.5743
P_3, P_1	0.4669	0.4947	0.4970	0.5703	0.5050	0.5589
P_2, P_1	0.5476	0.5061	0.4897	0.6010	0.4227	0.6004
Q_3, Q_2, Q_1	0.5302	0.4803	0.5480	0.9208	0.6575	0.9265
Q_3, Q_2, P_3	0.4972	0.5203	0.5620	0.8405	0.6644	0.8545
Q_3, Q_2, P_2	0.4360	0.5049	0.5558	0.8451	0.7141	0.8653
Q_3, Q_2, P_1	0.4681	0.5123	0.5541	0.8504	0.6922	0.8618
Q_3, Q_1, P_3	0.4721	0.4954	0.5622	0.8933	0.7204	0.9088
Q_3, Q_1, P_2	0.6153	0.4970	0.5648	0.8976	0.6931	0.9104
Q_3, Q_1, P_1	0.3514	0.4892	0.5621	0.9000	0.7629	0.9144
Q_3, P_3, P_2	0.5095	0.5090	0.5507	0.8020	0.6720	0.8218
Q_3, P_3, P_1	0.6059	0.5064	0.5617	0.8066	0.6830	0.8159
Q_3, P_2, P_1	0.4470	0.5129	0.5555	0.8028	0.7279	0.8189
Q_2, Q_1, P_3	0.5894	0.5084	0.5215	0.8592	0.6453	0.8677
Q_2, Q_1, P_2	0.5079	0.4979	0.5207	0.8593	0.5977	0.8755
Q_2, Q_1, P_1	0.6118	0.4860	0.5203	0.8607	0.6498	0.8749
Q_2, P_3, P_2	0.5679	0.5029	0.4971	0.7274	0.5432	0.7410
Q_2, P_3, P_1	0.4831	0.5021	0.5011	0.7173	0.4787	0.7413
Q_2, P_2, P_1	0.5484	0.4934	0.5080	0.7270	0.5704	0.7368
Q_1, P_3, P_2	0.5200	0.4927	0.5402	0.8190	0.6278	0.8369
Q_1, P_3, P_1	0.4559	0.5028	0.5245	0.8198	0.6148	0.8351
Q_1, P_2, P_1	0.4061	0.4945	0.5376	0.8232	0.6662	0.8430
P_3, P_2, P_1	0.5371	0.5064	0.5123	0.5900	0.5703	0.6103
Q_3, Q_2, Q_1, P_3	0.6161	0.5157	0.5636	0.9135	0.6910	0.9226
Q_3, Q_2, Q_1, P_2	0.4501	0.4960	0.5432	0.9104	0.6906	0.9231
Q_3, Q_2, Q_1, P_1	0.4769	0.5116	0.5571	0.9108	0.6911	0.9306
Q_3, Q_2, P_3, P_2	0.3576	0.5001	0.5493	0.8481	0.6448	0.8642
Q_3, Q_2, P_3, P_1	0.5516	0.4927	0.5502	0.8423	0.6587	0.8596
Q_3, Q_2, P_2, P_1	0.4514	0.4995	0.5654	0.8485	0.6667	0.8652
Q_3, Q_1, P_3, P_2	0.4502	0.5018	0.5753	0.8920	0.6736	0.9052
Q_3, Q_1, P_3, P_1	0.5708	0.5029	0.5523	0.8986	0.6678	0.9065

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.6328	0.5003	0.5701	0.8957	0.6920	0.9107
Q_3, P_3, P_2, P_1	0.4682	0.4956	0.5519	0.8011	0.6825	0.8163
Q_2, Q_1, P_3, P_2	0.4894	0.4988	0.5346	0.8577	0.6144	0.8667
Q_2, Q_1, P_3, P_1	0.5418	0.4968	0.5179	0.8477	0.5928	0.8747
Q_2, Q_1, P_2, P_1	0.5582	0.4922	0.5202	0.8513	0.6274	0.8734
Q_2, P_3, P_2, P_1	0.5467	0.5073	0.5109	0.7137	0.5794	0.7443
Q_1, P_3, P_2, P_1	0.5075	0.5073	0.5237	0.8179	0.6029	0.8338
Q_3, Q_2, Q_1, P_3, P_2	0.5460	0.4992	0.5434	0.9110	0.6913	0.9201
Q_3, Q_2, Q_1, P_3, P_1	0.5520	0.4984	0.5477	0.9103	0.6909	0.9218
Q_3, Q_2, Q_1, P_2, P_1	0.5154	0.5028	0.5407	0.9084	0.6888	0.9281
Q_3, Q_2, P_3, P_2, P_1	0.4996	0.4843	0.5478	0.8377	0.6378	0.8638
Q_3, Q_1, P_3, P_2, P_1	0.4262	0.4940	0.5729	0.8862	0.6664	0.9089
Q_2, Q_1, P_3, P_2, P_1	0.5904	0.5024	0.5354	0.8485	0.6138	0.8732
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.3930	0.5013	0.5597	0.9035	0.6729	0.9261

Table M.3: The specificities achieved across the combination search by each of the six classification methods.

Appendix N

AAA combination search results

The F_1 scores, sensitivities, and specificities achieved for AAA classification when using each of the six ML methods are shown in Table N.1, N.2, and N.3 respectively.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.4670	0.4881	0.8454	0.9095	0.8606	0.9294
Q_2	0.5754	0.4952	0.8246	0.9516	0.9092	0.9640
Q_1	0.4440	0.4843	0.9481	0.9741	0.9697	0.9805
P_3	0.4999	0.5102	0.8664	0.9027	0.8692	0.9226
P_2	0.5782	0.4944	0.8717	0.9087	0.8793	0.9311
P_1	0.4790	0.4826	0.8212	0.8771	0.8416	0.8884
Q_3, Q_2	0.3850	0.4983	0.8895	0.9753	0.9249	0.9843
Q_3, Q_1	0.4982	0.5029	0.9521	0.9840	0.9749	0.9919
Q_3, P_3	0.5126	0.4960	0.9215	0.9483	0.9249	0.9767
Q_3, P_2	0.6111	0.4958	0.9355	0.9543	0.9385	0.9770
Q_3, P_1	0.4737	0.4971	0.9286	0.9498	0.9448	0.9702
Q_2, Q_1	0.5523	0.4970	0.9523	0.9868	0.9718	0.9928
Q_2, P_3	0.5080	0.4994	0.9305	0.9604	0.9430	0.9805
Q_2, P_2	0.4756	0.4996	0.9371	0.9712	0.9552	0.9849
Q_2, P_1	0.4032	0.4975	0.9168	0.9689	0.9413	0.9828
Q_1, P_3	0.5350	0.5046	0.9630	0.9808	0.9741	0.9870
Q_1, P_2	0.4613	0.4981	0.9681	0.9820	0.9756	0.9900
Q_1, P_1	0.4909	0.5003	0.9747	0.9798	0.9801	0.9852
P_3, P_2	0.5343	0.5018	0.9247	0.9335	0.9305	0.9677
P_3, P_1	0.4857	0.5078	0.9321	0.9345	0.9311	0.9675
P_2, P_1	0.5431	0.5039	0.9213	0.9365	0.9405	0.9625
Q_3, Q_2, Q_1	0.4890	0.5164	0.9603	0.9912	0.9729	0.9962
Q_3, Q_2, P_3	0.5485	0.4993	0.9452	0.9771	0.9436	0.9905
Q_3, Q_2, P_2	0.5359	0.4998	0.9542	0.9791	0.9568	0.9910
Q_3, Q_2, P_1	0.4374	0.5070	0.9518	0.9803	0.9503	0.9906
Q_3, Q_1, P_3	0.5193	0.5085	0.9663	0.9861	0.9740	0.9936
Q_3, Q_1, P_2	0.5325	0.5034	0.9747	0.9884	0.9784	0.9939
Q_3, Q_1, P_1	0.4819	0.4943	0.9781	0.9850	0.9796	0.9936
Q_3, P_3, P_2	0.4106	0.4991	0.9479	0.9586	0.9434	0.9807
Q_3, P_3, P_1	0.4291	0.4901	0.9560	0.9598	0.9491	0.9846
Q_3, P_2, P_1	0.4537	0.4948	0.9492	0.9647	0.9515	0.9804
Q_2, Q_1, P_3	0.5071	0.5051	0.9685	0.9877	0.9795	0.9944
Q_2, Q_1, P_2	0.4853	0.4951	0.9724	0.9893	0.9797	0.9957
Q_2, Q_1, P_1	0.4459	0.4994	0.9752	0.9885	0.9816	0.9952
Q_2, P_3, P_2	0.4060	0.4932	0.9566	0.9714	0.9576	0.9873
Q_2, P_3, P_1	0.5857	0.4972	0.9577	0.9722	0.9582	0.9882
Q_2, P_2, P_1	0.4776	0.5030	0.9497	0.9755	0.9671	0.9892
Q_1, P_3, P_2	0.4224	0.4974	0.9729	0.9823	0.9788	0.9904
Q_1, P_3, P_1	0.4944	0.4987	0.9747	0.9813	0.9797	0.9897
Q_1, P_2, P_1	0.5362	0.5051	0.9756	0.9828	0.9827	0.9917
P_3, P_2, P_1	0.4406	0.5001	0.9479	0.9455	0.9517	0.9750

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_2, Q_1, P_3	0.5284	0.5135	0.9711	0.9914	0.9756	0.9965
Q_3, Q_2, Q_1, P_2	0.5279	0.5066	0.9784	0.9923	0.9794	0.9972
Q_3, Q_2, Q_1, P_1	0.4331	0.4983	0.9790	0.9903	0.9792	0.9961
Q_3, Q_2, P_3, P_2	0.5090	0.5041	0.9636	0.9797	0.9582	0.9930
Q_3, Q_2, P_3, P_1	0.5250	0.4963	0.9665	0.9784	0.9633	0.9922
Q_3, Q_2, P_2, P_1	0.4600	0.4887	0.9646	0.9829	0.9724	0.9937
Q_3, Q_1, P_3, P_2	0.4994	0.5003	0.9759	0.9880	0.9771	0.9939
Q_3, Q_1, P_3, P_1	0.5058	0.5060	0.9779	0.9867	0.9782	0.9942
Q_3, Q_1, P_2, P_1	0.4981	0.4974	0.9781	0.9869	0.9778	0.9950
Q_3, P_3, P_2, P_1	0.4679	0.5050	0.9634	0.9651	0.9595	0.9856
Q_2, Q_1, P_3, P_2	0.4910	0.4989	0.9776	0.9901	0.9759	0.9954
Q_2, Q_1, P_3, P_1	0.4893	0.5041	0.9794	0.9892	0.9772	0.9948
Q_2, Q_1, P_2, P_1	0.4849	0.4994	0.9771	0.9911	0.9800	0.9957
Q_2, P_3, P_2, P_1	0.4963	0.5081	0.9644	0.9748	0.9684	0.9903
Q_1, P_3, P_2, P_1	0.5090	0.5054	0.9763	0.9857	0.9788	0.9910
Q_3, Q_2, Q_1, P_3, P_2	0.4588	0.4997	0.9781	0.9915	0.9739	0.9970
Q_3, Q_2, Q_1, P_3, P_1	0.5224	0.4957	0.9800	0.9920	0.9767	0.9970
Q_3, Q_2, Q_1, P_2, P_1	0.5003	0.4947	0.9823	0.9912	0.9808	0.9966
Q_3, Q_2, P_3, P_2, P_1	0.4667	0.4900	0.9708	0.9828	0.9668	0.9948
Q_3, Q_1, P_3, P_2, P_1	0.5322	0.4962	0.9801	0.9874	0.9775	0.9938
Q_2, Q_1, P_3, P_2, P_1	0.4450	0.5064	0.9801	0.9892	0.9808	0.9961
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.5083	0.4991	0.9820	0.9912	0.9785	0.9970

Table N.1: The F_1 scores achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.5683	0.5120	0.8568	0.8878	0.8661	0.9300
Q_2	0.5738	0.5089	0.8136	0.9355	0.9100	0.9638
Q_1	0.4451	0.4962	0.9517	0.9654	0.9673	0.9799
P_3	0.4846	0.5035	0.8785	0.8765	0.8660	0.9202
P_2	0.4451	0.5110	0.8712	0.9005	0.8818	0.9352
P_1	0.6616	0.4902	0.8491	0.8514	0.8308	0.8770
Q_3, Q_2	0.5927	0.4676	0.8868	0.9652	0.9308	0.9835
Q_3, Q_1	0.5541	0.5333	0.9508	0.9757	0.9747	0.9907
Q_3, P_3	0.4269	0.4894	0.9222	0.9282	0.9266	0.9746
Q_3, P_2	0.4746	0.5016	0.9325	0.9382	0.9379	0.9819
Q_3, P_1	0.5850	0.4760	0.9213	0.9317	0.9462	0.9694
Q_2, Q_1	0.2504	0.5034	0.9534	0.9810	0.9738	0.9919
Q_2, P_3	0.4111	0.4591	0.9285	0.9464	0.9439	0.9793
Q_2, P_2	0.5865	0.5093	0.9345	0.9604	0.9544	0.9836
Q_2, P_1	0.5669	0.4940	0.9227	0.9552	0.9471	0.9817
Q_1, P_3	0.4266	0.4741	0.9626	0.9729	0.9743	0.9850
Q_1, P_2	0.5075	0.4991	0.9664	0.9743	0.9780	0.9895
Q_1, P_1	0.5143	0.5055	0.9742	0.9715	0.9806	0.9841
P_3, P_2	0.4414	0.4981	0.9287	0.9209	0.9379	0.9673
P_3, P_1	0.5355	0.4956	0.9461	0.9109	0.9337	0.9631
P_2, P_1	0.4090	0.4957	0.9311	0.9260	0.9359	0.9596
Q_3, Q_2, Q_1	0.6548	0.5014	0.9592	0.9864	0.9760	0.9954
Q_3, Q_2, P_3	0.4363	0.4885	0.9445	0.9689	0.9482	0.9897
Q_3, Q_2, P_2	0.5720	0.5284	0.9506	0.9704	0.9620	0.9904
Q_3, Q_2, P_1	0.4962	0.5110	0.9455	0.9723	0.9511	0.9914
Q_3, Q_1, P_3	0.5329	0.4857	0.9666	0.9793	0.9774	0.9913
Q_3, Q_1, P_2	0.3570	0.4931	0.9701	0.9820	0.9794	0.9929
Q_3, Q_1, P_1	0.3667	0.5022	0.9771	0.9755	0.9805	0.9924
Q_3, P_3, P_2	0.6250	0.5064	0.9434	0.9445	0.9426	0.9822
Q_3, P_3, P_1	0.4716	0.4865	0.9564	0.9413	0.9473	0.9843
Q_3, P_2, P_1	0.5103	0.4982	0.9447	0.9522	0.9575	0.9819
Q_2, Q_1, P_3	0.4499	0.4986	0.9676	0.9815	0.9797	0.9933
Q_2, Q_1, P_2	0.6389	0.4936	0.9689	0.9838	0.9795	0.9947
Q_2, Q_1, P_1	0.6675	0.5043	0.9741	0.9817	0.9811	0.9945
Q_2, P_3, P_2	0.5890	0.4948	0.9564	0.9609	0.9598	0.9864
Q_2, P_3, P_1	0.4238	0.5033	0.9606	0.9619	0.9578	0.9868
Q_2, P_2, P_1	0.5582	0.5024	0.9540	0.9660	0.9686	0.9881
Q_1, P_3, P_2	0.5561	0.4904	0.9703	0.9736	0.9786	0.9898
Q_1, P_3, P_1	0.6229	0.5165	0.9753	0.9725	0.9799	0.9881
Q_1, P_2, P_1	0.4489	0.5084	0.9753	0.9750	0.9837	0.9896
P_3, P_2, P_1	0.6036	0.5139	0.9563	0.9278	0.9522	0.9726
Q_3, Q_2, Q_1, P_3	0.4318	0.5058	0.9684	0.9870	0.9803	0.9953
Q_3, Q_2, Q_1, P_2	0.5271	0.4841	0.9751	0.9879	0.9791	0.9959
Q_3, Q_2, Q_1, P_1	0.6257	0.4871	0.9768	0.9848	0.9794	0.9944
Q_3, Q_2, P_3, P_2	0.4330	0.5113	0.9615	0.9692	0.9620	0.9922
Q_3, Q_2, P_3, P_1	0.4955	0.4973	0.9639	0.9675	0.9661	0.9925
Q_3, Q_2, P_2, P_1	0.4783	0.4925	0.9610	0.9737	0.9660	0.9930
Q_3, Q_1, P_3, P_2	0.4914	0.4957	0.9741	0.9818	0.9795	0.9932
Q_3, Q_1, P_3, P_1	0.5768	0.5028	0.9778	0.9794	0.9788	0.9928

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.4613	0.4924	0.9749	0.9805	0.9771	0.9940
Q_3, P_3, P_2, P_1	0.6938	0.5114	0.9619	0.9516	0.9633	0.9856
Q_2, Q_1, P_3, P_2	0.5969	0.4915	0.9770	0.9861	0.9772	0.9944
Q_2, Q_1, P_3, P_1	0.5361	0.5044	0.9800	0.9846	0.9770	0.9938
Q_2, Q_1, P_2, P_1	0.5999	0.5042	0.9753	0.9867	0.9815	0.9944
Q_2, P_3, P_2, P_1	0.4892	0.4885	0.9676	0.9650	0.9693	0.9887
Q_1, P_3, P_2, P_1	0.3810	0.5027	0.9761	0.9790	0.9791	0.9887
Q_3, Q_2, Q_1, P_3, P_2	0.5180	0.5006	0.9749	0.9866	0.9752	0.9959
Q_3, Q_2, Q_1, P_3, P_1	0.4600	0.4811	0.9805	0.9873	0.9794	0.9963
Q_3, Q_2, Q_1, P_2, P_1	0.4965	0.5034	0.9824	0.9870	0.9808	0.9952
Q_3, Q_2, P_3, P_2, P_1	0.4020	0.5030	0.9704	0.9745	0.9692	0.9944
Q_3, Q_1, P_3, P_2, P_1	0.4284	0.5086	0.9809	0.9804	0.9763	0.9925
Q_2, Q_1, P_3, P_2, P_1	0.5795	0.4863	0.9812	0.9836	0.9811	0.9949
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.4242	0.5024	0.9802	0.9861	0.9778	0.9959

Table N.2: The sensitivities achieved across the combination search by each of the six classification methods.

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3	0.4362	0.4805	0.8371	0.9276	0.8565	0.9290
Q_2	0.5761	0.4908	0.8324	0.9663	0.9087	0.9643
Q_1	0.4437	0.4806	0.9450	0.9825	0.9720	0.9811
P_3	0.5050	0.5126	0.8572	0.9244	0.8718	0.9248
P_2	0.6324	0.4890	0.8722	0.9156	0.8775	0.9277
P_1	0.4215	0.4803	0.8018	0.8972	0.8496	0.8976
Q_3, Q_2	0.3355	0.5086	0.8917	0.9851	0.9200	0.9851
Q_3, Q_1	0.4797	0.4927	0.9533	0.9922	0.9751	0.9931
Q_3, P_3	0.5422	0.4982	0.9210	0.9666	0.9235	0.9788
Q_3, P_2	0.6712	0.4939	0.9383	0.9691	0.9392	0.9724
Q_3, P_1	0.4392	0.5041	0.9351	0.9662	0.9436	0.97103
Q_2, Q_1	0.6675	0.4949	0.9514	0.9926	0.9701	0.9938
Q_2, P_3	0.5410	0.5129	0.9324	0.9735	0.9423	0.9817
Q_2, P_2	0.4411	0.4964	0.9394	0.9815	0.9561	0.9862
Q_2, P_1	0.3619	0.4987	0.9119	0.9819	0.9363	0.9840
Q_1, P_3	0.5747	0.5149	0.9635	0.9885	0.9741	0.9890
Q_1, P_2	0.4475	0.4979	0.9697	0.9896	0.9734	0.9906
Q_1, P_1	0.4834	0.4986	0.9753	0.9879	0.9797	0.9863
P_3, P_2	0.5682	0.5031	0.9213	0.9447	0.9241	0.9681
P_3, P_1	0.4698	0.5120	0.9199	0.9552	0.9289	0.9718
P_2, P_1	0.5932	0.5067	0.9130	0.9459	0.9446	0.9652
Q_3, Q_2, Q_1	0.4354	0.5217	0.9615	0.9961	0.9700	0.9970
Q_3, Q_2, P_3	0.5910	0.5030	0.9460	0.9850	0.9396	0.9914
Q_3, Q_2, P_2	0.5227	0.4904	0.9575	0.9876	0.9522	0.9917
Q_3, Q_2, P_1	0.4210	0.5057	0.9576	0.9880	0.9496	0.9899
Q_3, Q_1, P_3	0.5146	0.5163	0.9662	0.9928	0.9708	0.9960
Q_3, Q_1, P_2	0.5963	0.5069	0.9792	0.9947	0.9775	0.9950
Q_3, Q_1, P_1	0.5186	0.4918	0.9792	0.9944	0.9789	0.9949
Q_3, P_3, P_2	0.3553	0.4967	0.9520	0.9716	0.9442	0.9794
Q_3, P_3, P_1	0.4176	0.4913	0.9557	0.9769	0.9509	0.9850
Q_3, P_2, P_1	0.4371	0.4938	0.9533	0.9764	0.9461	0.9791
Q_2, Q_1, P_3	0.5266	0.5074	0.9695	0.9939	0.9794	0.9956
Q_2, Q_1, P_2	0.4362	0.4957	0.9758	0.9948	0.9799	0.9967
Q_2, Q_1, P_1	0.3824	0.4979	0.9764	0.9952	0.9822	0.9959
Q_2, P_3, P_2	0.3595	0.4928	0.9568	0.9814	0.9557	0.9882
Q_2, P_3, P_1	0.6529	0.4952	0.9552	0.9821	0.9586	0.9896
Q_2, P_2, P_1	0.4524	0.5033	0.9460	0.9847	0.9658	0.9903
Q_1, P_3, P_2	0.3867	0.4998	0.9754	0.9908	0.9791	0.9910
Q_1, P_3, P_1	0.4523	0.4929	0.9743	0.9898	0.9796	0.9913
Q_1, P_2, P_1	0.5683	0.5040	0.9759	0.9904	0.9819	0.9939
P_3, P_2, P_1	0.3946	0.4955	0.9405	0.9614	0.9513	0.9774
Q_3, Q_2, Q_1, P_3	0.5631	0.5162	0.9738	0.9958	0.9713	0.9977
Q_3, Q_2, Q_1, P_2	0.5282	0.5143	0.9816	0.9967	0.9797	0.9986
Q_3, Q_2, Q_1, P_1	0.3799	0.5021	0.9813	0.9958	0.9792	0.9978
Q_3, Q_2, P_3, P_2	0.5350	0.5018	0.9657	0.9899	0.9548	0.9939
Q_3, Q_2, P_3, P_1	0.5356	0.4960	0.9690	0.9890	0.9608	0.9921
Q_3, Q_2, P_2, P_1	0.4546	0.4875	0.9680	0.9918	0.9785	0.9944
Q_3, Q_1, P_3, P_2	0.5021	0.5019	0.9777	0.9941	0.9749	0.9947
Q_3, Q_1, P_3, P_1	0.4818	0.5072	0.9781	0.9939	0.9778	0.9956

Input combination	Classification method					
	NB	LR	SVM	RF	MLP	GB
Q_3, Q_1, P_2, P_1	0.5104	0.4991	0.9813	0.9932	0.9785	0.9961
Q_3, P_3, P_2, P_1	0.3990	0.5029	0.9648	0.9777	0.9561	0.9856
Q_2, Q_1, P_3, P_2	0.4566	0.5014	0.9782	0.9942	0.9748	0.9964
Q_2, Q_1, P_3, P_1	0.4742	0.5041	0.9789	0.9938	0.9774	0.9958
Q_2, Q_1, P_2, P_1	0.4481	0.4979	0.9790	0.9955	0.9786	0.9970
Q_2, P_3, P_2, P_1	0.4987	0.5149	0.9615	0.9843	0.9677	0.9919
Q_1, P_3, P_2, P_1	0.5527	0.5064	0.9765	0.9924	0.9787	0.9933
Q_3, Q_2, Q_1, P_3, P_2	0.4412	0.4995	0.9813	0.9965	0.9727	0.9981
Q_3, Q_2, Q_1, P_3, P_1	0.5446	0.5006	0.9796	0.9967	0.9743	0.9978
Q_3, Q_2, Q_1, P_2, P_1	0.5016	0.4919	0.9823	0.9955	0.9808	0.9981
Q_3, Q_2, P_3, P_2, P_1	0.4864	0.4858	0.9712	0.9910	0.9647	0.9952
Q_3, Q_1, P_3, P_2, P_1	0.5699	0.4922	0.9794	0.9944	0.9788	0.9951
Q_2, Q_1, P_3, P_2, P_1	0.4066	0.5133	0.9792	0.9947	0.9806	0.9973
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.5370	0.4981	0.9839	0.9964	0.9792	0.9981

Table N.3: The specificities achieved across the combination search by each of the six classification methods.

Appendix O

AAA-L combination search results

The F_1 scores, sensitivities, and specificities achieved for AAA-L classification when employing the GB method are shown in Table O.1.

Input combination	F_1	Sen.	Spec.
Q_3	0.8633	0.8561	0.8689
Q_2	0.9010	0.9103	0.8934
Q_1	0.9528	0.9630	0.9436
P_3	0.8305	0.8383	0.8250
P_2	0.8380	0.8529	0.8274
P_1	0.8005	0.7700	0.8209
Q_3, Q_2	0.9387	0.9390	0.9385
Q_3, Q_1	0.9683	0.9681	0.9685
Q_3, P_3	0.9045	0.8968	0.9109
Q_3, P_2	0.9151	0.9127	0.9172
Q_3, P_1	0.8989	0.8942	0.9028
Q_2, Q_1	0.9711	0.9741	0.9683
Q_2, P_3	0.9176	0.9256	0.9109
Q_2, P_2	0.9229	0.9328	0.9145
Q_2, P_1	0.9234	0.9258	0.9215
Q_1, P_3	0.9569	0.9558	0.9580
Q_1, P_2	0.9606	0.9645	0.9570
Q_1, P_1	0.9618	0.9609	0.9628
P_3, P_2	0.8852	0.8889	0.8824
P_3, P_1	0.8877	0.8889	0.8869
P_2, P_1	0.884	0.8858	0.8838
Q_3, Q_2, Q_1	0.9777	0.9788	0.9767
Q_3, Q_2, P_3	0.9454	0.9513	0.9402
Q_3, Q_2, P_2	0.9455	0.9498	0.9417
Q_3, Q_2, P_1	0.9481	0.9537	0.9431
Q_3, Q_1, P_3	0.9693	0.9743	0.9647
Q_3, Q_1, P_2	0.9695	0.9748	0.9647
Q_3, Q_1, P_1	0.9668	0.9642	0.9693
Q_3, P_3, P_2	0.9148	0.9105	0.9186
Q_3, P_3, P_1	0.9178	0.9232	0.9133
Q_3, P_2, P_1	0.9217	0.9163	0.9265
Q_2, Q_1, P_3	0.9770	0.9788	0.9753
Q_2, Q_1, P_2	0.9715	0.9729	0.9702
Q_2, Q_1, P_1	0.9737	0.9762	0.9714
Q_2, P_3, P_2	0.9327	0.9434	0.9234
Q_2, P_3, P_1	0.9285	0.9299	0.9273
Q_2, P_2, P_1	0.9345	0.9304	0.9381
Q_1, P_3, P_2	0.9606	0.9640	0.9575
Q_1, P_3, P_1	0.9637	0.9676	0.9601
Q_1, P_2, P_1	0.9607	0.9625	0.9592
P_3, P_2, P_1	0.8996	0.9038	0.8963
Q_3, Q_2, Q_1, P_3	0.9767	0.9781	0.9755

Input combination	F_1	Sen.	Spec.
Q_3, Q_2, Q_1, P_2	0.9788	0.9786	0.9791
Q_3, Q_2, Q_1, P_1	0.9759	0.9791	0.9729
Q_3, Q_2, P_3, P_2	0.9484	0.9510	0.9462
Q_3, Q_2, P_3, P_1	0.9487	0.9525	0.9453
Q_3, Q_2, P_2, P_1	0.9472	0.9529	0.9421
Q_3, Q_1, P_3, P_2	0.9670	0.9654	0.9685
Q_3, Q_1, P_3, P_1	0.9673	0.9678	0.9669
Q_3, Q_1, P_2, P_1	0.9704	0.9683	0.9724
Q_3, P_3, P_2, P_1	0.9217	0.9227	0.9210
Q_2, Q_1, P_3, P_2	0.9754	0.9781	0.9729
Q_2, Q_1, P_3, P_1	0.9774	0.9784	0.9765
Q_2, Q_1, P_2, P_1	0.9772	0.9776	0.9770
Q_2, P_3, P_2, P_1	0.9352	0.9436	0.9280
Q_1, P_3, P_2, P_1	0.9587	0.9659	0.9522
Q_3, Q_2, Q_1, P_3, P_2	0.9744	0.9731	0.9758
Q_3, Q_2, Q_1, P_3, P_1	0.9820	0.9834	0.9808
Q_3, Q_2, Q_1, P_2, P_1	0.9802	0.9796	0.9808
Q_3, Q_2, P_3, P_2, P_1	0.9513	0.9541	0.9489
Q_3, Q_1, P_3, P_2, P_1	0.9725	0.9712	0.9738
Q_2, Q_1, P_3, P_2, P_1	0.9757	0.9815	0.9702
$Q_3, Q_2, Q_1, P_3, P_2, P_1$	0.9809	0.9808	0.9810

Table O.1: The F_1 scores, sensitivities and specificities achieved across the combination search by the GB method.

Appendix P

GB results for all disease forms

The F_1 scores achieved for all forms of disease classification (including AAA-L) when providing each combination of input measurements are shown when employing the GB method in Figure P.1.

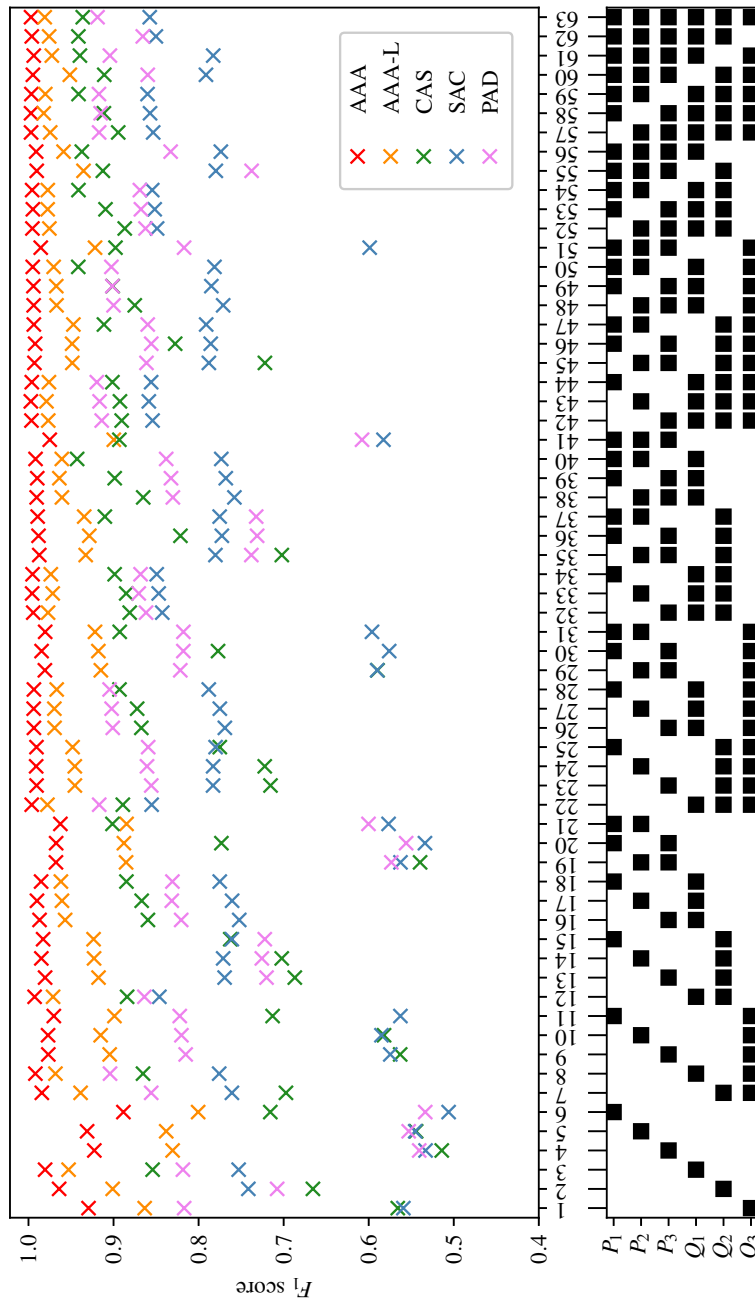


Figure P.1: The F_1 scores achieved for all disease forms when employing the GB method. Measurements included within each combination are highlighted with a black square.

References

- [1] Victor Aboyans et al. “The general prognosis of patients with peripheral arterial disease differs according to the disease localization”. In: *Journal of the American College of Cardiology* 55.9 (2010), pp. 898–903.
- [2] Victor Aboyans et al. “The vital prognosis of subclavian stenosis”. In: *Journal of the American College of Cardiology* 49.14 (2007), pp. 1540–1545.
- [3] Audrey Adji, Kozo Hirata, and Michael F O’rourke. “Clinical use of indices determined non-invasively from the radial and carotid pressure waveforms”. In: *Blood pressure monitoring* 11.4 (2006), pp. 215–221.
- [4] Anthony K Akobeng. “Understanding diagnostic tests 3: receiver operating characteristic curves”. In: *Acta paediatrica* 96.5 (2007), pp. 644–647.
- [5] Mo Al-Qaisi et al. “Ankle brachial pressure index (ABPI): An update for practitioners”. In: *Vascular health and risk management* 5 (2009), p. 833.
- [6] Jordi Alastruey, Kim H Parker, Spencer J Sherwin, et al. “Arterial pulse wave haemodynamics”. In: *11th International Conference on Pressure Surges*. Virtual PiE Led t/a BHR Group: Lisbon, Portugal. 2012, pp. 401–442.
- [7] Anthony J Alberg et al. “The use of “overall accuracy” to evaluate the validity of screening or diagnostic tests”. In: *Journal of general internal medicine* 19.5p1 (2004), pp. 460–465.
- [8] Natalia A Alexandrova et al. “Carotid artery stenosis in peripheral vascular disease”. In: *Journal of vascular surgery* 23.4 (1996), pp. 645–649.
- [9] J Alihanka, K Vaahtoranta, and I Saarikivi. “A new method for long-term monitoring of the ballistocardiogram, heart rate, and respiration”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 240.5 (1981), R384–R392.
- [10] Naomi S Altman. “An introduction to kernel and nearest-neighbor nonparametric regression”. In: *The American Statistician* 46.3 (1992), pp. 175–185.

- [11] Joachim O Arndt, Jürgen Klauske, and Frank Mersch. “The diameter of the intact carotid artery in man and its change with pulse pressure”. In: *Pflüger’s Archiv für die gesamte Physiologie des Menschen und der Tiere* 301.3 (1968), pp. 230–240.
- [12] Roland Asmar et al. “Assessment of arterial distensibility by automatic pulse wave velocity measurement: validation and clinical application studies”. In: *Hypertension* 26.3 (1995), pp. 485–490.
- [13] AP Avolio. “Multi-branched model of the human arterial system”. In: *Medical and Biological Engineering and Computing* 18.6 (1980), pp. 709–718.
- [14] Richard Bellman. “Dynamic programming”. In: *Science* 153.3731 (1966), pp. 34–37.
- [15] Jacob Benesty et al. “Pearson correlation coefficient”. In: *Noise reduction in speech processing*. Springer, 2009, pp. 1–4.
- [16] RLT Bevan et al. “Patient-specific blood flow simulation through an aneurysmal thoracic aorta with a folded proximal neck”. In: *International Journal for Numerical Methods in Biomedical Engineering* 27.8 (2011), pp. 1167–1184.
- [17] Pablo J Blanco et al. “An anatomically detailed arterial network model for one-dimensional computational hemodynamics”. In: *IEEE Transactions on biomedical engineering* 62.2 (2015), pp. 736–753.
- [18] Pablo J Blanco et al. “Blood flow distribution in an anatomically detailed arterial network model: criteria and algorithms”. In: *Biomechanics and modeling in mechanobiology* 13.6 (2014), pp. 1303–1330.
- [19] Mary L Boas. *Mathematical methods in the physical sciences*. Wiley, 2006.
- [20] Etienne Boileau et al. “A benchmark study of numerical schemes for one-dimensional arterial blood flow modelling”. In: *International journal for numerical methods in biomedical engineering* 31.10 (2015).
- [21] Etienne Boileau et al. “Estimating the accuracy of a reduced-order model for the calculation of fractional flow reserve (FFR)”. In: *International journal for numerical methods in biomedical engineering* 34.1 (2018), e2908.

- [22] Antoine Brault, Laurent Dumas, and Didier Lucor. “Uncertainty quantification of inflow boundary condition and proximal arterial stiffness–coupled effect on pulse wave propagation in a vascular network”. In: *International journal for numerical methods in biomedical engineering* 33.10 (2017), e2859.
- [23] Leo Breiman. “Random forests”. In: *Machine learning* 45.1 (2001), pp. 5–32.
- [24] Anthony E Brockwell. “Parallel Markov chain Monte Carlo simulation by pre-fetching”. In: *Journal of Computational and Graphical Statistics* 15.1 (2006), pp. 246–261.
- [25] Steve Brooks et al. *Handbook of markov chain monte carlo*. CRC press, 2011.
- [26] Steven J Burnham, Paul Jaques, and Cynthia B Burnham. “Noninvasive detection of iliac artery stenosis in the presence of superficial femoral artery obstruction”. In: *Journal of vascular surgery* 16.3 (1992), pp. 445–452.
- [27] S Byström et al. “Ultrasound-Doppler technique for monitoring blood flow in the brachial artery compared with occlusion plethysmography of the forearm”. In: *Scandinavian journal of clinical and laboratory investigation* 58.7 (1998), pp. 569–576.
- [28] Chain Monte Carlo. “Markov chain monte carlo and gibbs sampling”. In: *Lecture notes for EEB 581* (2004).
- [29] Jason Carson et al. “Personalising cardiovascular network models in pregnancy: A two-tiered parameter estimation approach”. In: *International Journal for Numerical Methods in Biomedical Engineering* (2019), e3267.
- [30] Jason M Carson. “Development of a cardiovascular and lymphatic network model during human pregnancy”. PhD thesis. Swansea University, 2019.
- [31] Numan Celik et al. “Deep-Channel uses deep neural networks to detect single-molecule events from patch-clamp data”. In: *Communications Biology* 3.1 (2020), pp. 1–10.
- [32] Elliot L Chaikof et al. “SVS practice guidelines for the care of patients with an abdominal aortic aneurysm: executive summary”. In: *Journal of vascular surgery* 50.4 (2009), pp. 880–896.

- [33] Neeraj Kavan Chakshu, Igor Sazonov, and Perumal Nithiarasu. “Towards enabling a cardiovascular digital twin for human systemic circulation using inverse analysis”. In: *Biomechanics and Modeling in Mechanobiology* (2020), pp. 1–17.
- [34] Neeraj Kavan Chakshu et al. “A semi-active human digital twin model for detecting severity of carotid stenoses from head vibration—A coupled computational mechanics and computer vision method”. In: *International journal for numerical methods in biomedical engineering* 35.5 (2019), e3180.
- [35] P Anthony Chandraratna et al. “Determination of cardiac output by transcutaneous continuous-wave ultrasonic Doppler computer”. In: *The American journal of cardiology* 53.1 (1984), pp. 234–237.
- [36] Peter H Charlton et al. “Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 317.5 (2019), H1062–H1085.
- [37] Eugene Charniak. “Bayesian networks without tears.” In: *AI magazine* 12.4 (1991), pp. 50–50.
- [38] Chen-Huan Chen et al. “Estimation of central aortic pressure waveform by mathematical transformation of radial tonometry pressure: validation of generalized transfer function”. In: *Circulation* 95.7 (1997), pp. 1827–1836.
- [39] Qian Chen et al. “Disease location is associated with survival in patients with peripheral arterial disease”. In: *Journal of the American Heart Association* 2.5 (2013), e000304.
- [40] Siddhartha Chib and Edward Greenberg. “Understanding the metropolis-hastings algorithm”. In: *The american statistician* 49.4 (1995), pp. 327–335.
- [41] J Andrés Christen, Colin Fox, et al. “A general purpose sampling algorithm for continuous distributions (the t-walk)”. In: *Bayesian Analysis* 5.2 (2010), pp. 263–281.
- [42] Emre Çomak, Ahmet Arslan, and İbrahim Türkoğlu. “A decision support system based on support vector machines for diagnosis of the heart valve diseases”. In: *Computers in biology and Medicine* 37.1 (2007), pp. 21–27.

- [43] Paul A Cosford, Gillian C Leng, and Justyn Thomas. “Screening for abdominal aortic aneurysm”. In: *Cochrane database of systematic reviews* 2 (2007).
- [44] Rosie Darwood et al. “Twenty-year review of abdominal aortic aneurysm screening in men in the county of Gloucestershire, United Kingdom”. In: *Journal of vascular surgery* 56.1 (2012), pp. 8–13.
- [45] Manoranjan Dash and Huan Liu. “Feature selection for classification”. In: *Intelligent data analysis* 1.3 (1997), pp. 131–156.
- [46] Wilbur B Davenport, William L Root, et al. *An introduction to the theory of random signals and noise*. Vol. 159. McGraw-Hill New York, 1958.
- [47] Justine Ina Davies and Allan D Struthers. “Pulse wave analysis and pulse wave velocity: a critical review of their strengths and weaknesses”. In: *Journal of hypertension* 21.3 (2003), pp. 463–472.
- [48] Meryl Davis, Mike Harris, and Jonathan J Earnshaw. “Implementation of the national health service abdominal aortic aneurysm screening program in England”. In: *Journal of vascular surgery* 57.5 (2013), pp. 1440–1445.
- [49] William Dock and Felix Taubman. “Some technics for recording the ballistocardiogram directly from the body”. In: *The American journal of medicine* 7.6 (1949), pp. 751–755.
- [50] Thomas J Donohue et al. “Assessing the hemodynamic significance of coronary artery stenosis: Analysis of translesional pressure-flow velocity relations in patients”. In: *Journal of the American College of Cardiology* 22.2 (1993), pp. 449–458.
- [51] Sahibsingh A Dudani. “The distance-weighted k-nearest-neighbor rule”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 4 (1976), pp. 325–327.
- [52] Ron Dueck, Oliver Goedje, and Paul Clopton. “Noninvasive continuous beat-to-beat radial artery pressure via TL-200 applanation tonometry”. In: *Journal of clinical monitoring and computing* 26.2 (2012), pp. 75–83.
- [53] Mark L Dyken et al. “Complete occlusion of common or internal carotid arteries: Clinical significance”. In: *Archives of Neurology* 30.5 (1974), pp. 343–346.

- [54] Joseph S Eckerle. “Tonometry, arterial”. In: *Encyclopedia of medical devices and instrumentation* (2006).
- [55] Bradley Efron. “Bayes’ theorem in the 21st century”. In: *Science* 340.6137 (2013), pp. 1177–1178.
- [56] Sturla H Eik-Nes, Karel Maršál, and Kjell Kristoffersen. “Methodology and basic problems related to blood flow studies in the human fetus”. In: *Ultrasound in medicine & biology* 10.3 (1984), pp. 329–337.
- [57] Jane Elith, John R Leathwick, and Trevor Hastie. “A working guide to boosted regression trees”. In: *Journal of Animal Ecology* 77.4 (2008), pp. 802–813.
- [58] Sally Epstein et al. “Reducing the number of parameters in 1D arterial blood flow modeling: less is more for patient-specific simulations”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 309.1 (2015), H222–H234.
- [59] Calvin B Ernst. “Abdominal aortic aneurysm”. In: *New England Journal of Medicine* 328.16 (1993), pp. 1167–1172.
- [60] Matthias Feurer et al. “Auto-sklearn: efficient and robust automated machine learning”. In: *Automated Machine Learning*. Springer, Cham, 2019, pp. 113–134.
- [61] Mark F Fillinger et al. “Prediction of rupture risk in abdominal aortic aneurysm during observation: wall stress versus diameter”. In: *Journal of vascular surgery* 37.4 (2003), pp. 724–732.
- [62] Joachim E Fischer, Lucas M Bachmann, and Roman Jaeschke. “A readers’ guide to the interpretation of diagnostic test properties: clinical example of sepsis”. In: *Intensive care medicine* 29.7 (2003), pp. 1043–1051.
- [63] Dominik Fleischmann, Richard L Hallett, and Geoffrey D Rubin. “CT angiography of peripheral arterial disease”. In: *Journal of vascular and interventional radiology* 17.1 (2006), pp. 3–26.
- [64] Luca Formaggia, Daniele Lamponi, and Alfio Quarteroni. “One-dimensional models for blood flow in arteries”. In: *Journal of engineering mathematics* 47.3-4 (2003), pp. 251–276.

- [65] Kenneth R Foster, Robert Koprowski, and Joseph D Skufca. “Machine learning, medical diagnosis, and biomedical engineering research-commentary”. In: *Biomedical engineering online* 13.1 (2014), p. 94.
- [66] F Gerald R Fowkes et al. “Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis”. In: *The Lancet* 382.9901 (2013), pp. 1329–1340.
- [67] Jerome H Friedman. “Greedy function approximation: a gradient boosting machine”. In: *Annals of statistics* (2001), pp. 1189–1232.
- [68] Pierre Geurts. “Bias vs variance decomposition for regression and classification”. In: *Data mining and knowledge discovery handbook*. Springer, 2009, pp. 733–746.
- [69] Charles J Geyer. “Practical markov chain monte carlo”. In: *Statistical science* (1992), pp. 473–483.
- [70] Sonia Gilaberte et al. “Heart and respiratory rate detection on a bathroom scale based on the ballistocardiogram and the continuous wavelet transform”. In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, 2010, pp. 2557–2560.
- [71] Walter R Gilks. “M arkov Chain M onte C arlo”. In: *Encyclopedia of biostatistics* 4 (2005).
- [72] Robert W Gill. “Measurement of blood flow by ultrasound: accuracy and sources of error”. In: *Ultrasound in Medicine and Biology* 11.4 (1985), pp. 625–641.
- [73] MJ Glover et al. “Cost-effectiveness of the National Health Service abdominal aortic aneurysm screening programme in England”. In: *The British journal of surgery* 101.8 (2014), p. 976.
- [74] Gene H Golub and John H Welsch. “Calculation of Gauss quadrature rules”. In: *Mathematics of computation* 23.106 (1969), pp. 221–230.
- [75] Jianping Gou et al. “A new distance-weighted k-nearest neighbor classifier”. In: *J. Inf. Comput. Sci* 9.6 (2012), pp. 1429–1436.

- [76] Ilja Guelen et al. “Finometer, finger pressure measurements with the possibility to reconstruct brachial pressure”. In: *Blood pressure monitoring* 8.1 (2003), pp. 27–30.
- [77] Ilja Guelen et al. “Validation of brachial artery pressure reconstruction from finger arterial pressure”. In: *Journal of hypertension* 26.7 (2008), pp. 1321–1327.
- [78] Godofredo R Gutierrez et al. “Prevalence of subclavian artery stenosis in patients with peripheral vascular disease”. In: *Angiology* 52.3 (2001), pp. 189–194.
- [79] Isabelle Guyon and André Elisseeff. “An introduction to variable and feature selection”. In: *Journal of machine learning research* 3.Mar (2003), pp. 1157–1182.
- [80] F Hansen et al. “Diameter and compliance in the human common carotid artery—variations with age and sex”. In: *Ultrasound in medicine & biology* 21.1 (1995), pp. 1–9.
- [81] Nancy L Harthun. “Current issues in the treatment of women with abdominal aortic aneurysm”. In: *Gender medicine* 5.1 (2008), pp. 36–43.
- [82] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
- [83] Robert Hecht-Nielsen. “Theory of the backpropagation neural network”. In: *Neural networks for perception*. Elsevier, 1992, pp. 65–93.
- [84] Mathilde C Hemon and Justin P Phillips. “Comparison of foot finding methods for deriving instantaneous pulse rates from photoplethysmographic signals”. In: *Journal of clinical monitoring and computing* 30.2 (2016), pp. 157–168.
- [85] Bertha Hidalgo and Melody Goodman. “Multivariate or multivariable regression?” In: *American journal of public health* 103.1 (2013), pp. 39–40.
- [86] Joseph M Hilbe. *Logistic regression models*. CRC press, 2009.
- [87] Alan T Hirsch et al. “Peripheral arterial disease detection, awareness, and treatment in primary care”. In: *Jama* 286.11 (2001), pp. 1317–1324.

-
- [88] Arthur E Hoerl, Robert W Kannard, and Kent F Baldwin. “Ridge regression: some simulations”. In: *Communications in Statistics-Theory and Methods* 4.2 (1975), pp. 105–123.
- [89] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*. Vol. 398. John Wiley & Sons, 2013.
- [90] Alun D Hughes and Kim H Parker. “Forward and backward waves in the arterial system: impedance or wave intensity analysis?” In: *Medical & biological engineering & computing* 47.2 (2009), pp. 207–210.
- [91] Ildar Abdulovich Ibragimov and Yurii Antol’evich Rozanov. *Gaussian random processes*. Vol. 9. Springer Science & Business Media, 2012.
- [92] HALFDAN Ihlen et al. “Determination of cardiac output by Doppler echocardiography.” In: *Heart* 51.1 (1984), pp. 54–60.
- [93] OT Inan et al. “Robust ballistocardiogram acquisition for home monitoring”. In: *Physiological measurement* 30.2 (2009), p. 169.
- [94] Vikramaditya Jakkula. “Tutorial on support vector machine (svm)”. In: *School of EECS, Washington State University* 37 (2006).
- [95] Finn V Jensen et al. *An introduction to Bayesian networks*. Vol. 210. UCL press London, 1996.
- [96] Weiwei Jin and Jordi Alastruey. “Arterial pulse wave propagation across stenoses and aneurysms: assessment of one-dimensional simulations against three-dimensional simulations and in vitro measurements”. In: *Journal of the Royal Society Interface* 18.177 (2021), p. 20200881.
- [97] Oddmund Joakimsen et al. “Prediction of mortality by ultrasound screening of a general population for carotid stenosis: the Tromsø Study”. In: *Stroke* 31.8 (2000), pp. 1871–1876.
- [98] Brian F Johnson et al. “Clinical outcome in patients with mild and moderate carotid artery stenosis”. In: *Journal of vascular surgery* 21.1 (1995), pp. 120–126.
- [99] Gareth Jones et al. *A physiologically realistic virtual patient database for the study of arterial haemodynamics*. 2021. arXiv: 2102 . 10655 [physics.med-ph].

- [100] Gareth Jones et al. “A proof of concept study for machine learning application to stenosis detection”. In: *Arxiv* 0.0 (2021), pp. 0–0.
- [101] Gareth Jones et al. *Machine learning for detection of stenoses and aneurysms: application in a physiologically realistic virtual patient database*. 2021. arXiv: 2103.00599 [cs.LG].
- [102] Vojislav Kecman. *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models*. MIT press, 2001.
- [103] Vojislav Kecman. “Support vector machines—an introduction”. In: *Support vector machines: theory and applications*. Springer, 2005, pp. 1–47.
- [104] K Craig Kent. “Abdominal aortic aneurysms”. In: *New England journal of medicine* 371.22 (2014), pp. 2101–2108.
- [105] Eamonn Keogh and Abdullah Mueen. “Curse of dimensionality”. In: *Encyclopedia of Machine Learning and Data Mining* (2017), pp. 314–315.
- [106] Ahsan H Khandoker, Marimuthu Palaniswami, and Chandan K Karmakar. “Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings”. In: *IEEE transactions on information technology in biomedicine* 13.1 (2009), pp. 37–48.
- [107] AW Khir et al. “Determination of wave speed and wave separation in the arteries”. In: *Journal of biomechanics* 34.9 (2001), pp. 1145–1155.
- [108] Chang-Sei Kim et al. “Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring”. In: *Scientific reports* 6 (2016), p. 31297.
- [109] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).
- [110] Kenji Kira and Larry A Rendell. “A practical approach to feature selection”. In: *Machine Learning Proceedings 1992*. Elsevier, 1992, pp. 249–256.
- [111] T Koivistoinen et al. “A new method for measuring the ballistocardiogram using EMFi sensors in a normal chair”. In: *The 26th annual international conference of the IEEE engineering in medicine and biology society*. Vol. 1. IEEE. 2004, pp. 2026–2029.

- [112] Igor Kononenko. “Machine learning for medical diagnosis: history, state of the art and perspective”. In: *Artificial Intelligence in medicine* 23.1 (2001), pp. 89–109.
- [113] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. “Supervised machine learning: A review of classification techniques”. In: *Emerging artificial intelligence applications in computer engineering* 160.1 (2007), pp. 3–24.
- [114] Mariusz Kubus. “Discriminant stepwise procedure”. In: *Acta Universitatis Lodzianensis. Folia Oeconomica* 3.302 (2014).
- [115] Iftikhar J Kullo and Thom W Rooke. “Peripheral artery disease”. In: *New England Journal of Medicine* 374.9 (2016), pp. 861–871.
- [116] Frances Y Kuo and Ian H Sloan. “Lifting the curse of dimensionality”. In: *Notices of the AMS* 52.11 (2005), pp. 1320–1328.
- [117] Ben Lambert. *A student’s guide to Bayesian statistics*. Sage, 2018.
- [118] Nicolas Langwieser et al. “Radial artery applanation tonometry for continuous noninvasive arterial blood pressure monitoring in the cardiac intensive care unit”. In: *Clinical Research in Cardiology* 104.6 (2015), pp. 518–524.
- [119] George R Leopold, Lawrence E Goldberger, and Eugene F Bernstein. “Ultrasonic detection and evaluation of abdominal aortic aneurysms”. In: *Surgery* 72.6 (1972), pp. 939–945.
- [120] JA Levenson et al. “Pulsed Doppler: determination of diameter, blood flow velocity, and volumic flow of brachial artery in man”. In: *Cardiovascular Research* 15.3 (1981), pp. 164–170.
- [121] P Lewis et al. “Measurement of volume flow in the human common femoral artery using a duplex ultrasound system”. In: *Ultrasound in medicine & biology* 12.10 (1986), pp. 777–784.
- [122] Xi Li et al. “Prevalence and trends of the abdominal aortic aneurysms epidemic in general population—a meta-analysis”. In: *PLoS One* 8.12 (2013), e81260.
- [123] Andy Liaw, Matthew Wiener, et al. “Classification and regression by randomForest”. In: *R news* 2.3 (2002), pp. 18–22.

- [124] Andrew W Litt et al. “Diagnosis of carotid artery stenosis: comparison of 2DFT time-of-flight MR angiography with contrast angiography in 50 patients.” In: *American journal of neuroradiology* 12.1 (1991), pp. 149–154.
- [125] Melvin Lobo et al. “A novel non-invasive cuff-less optoelectronic sensor to measure blood pressure: comparison against intra-arterial measurement”. In: *Journal of Hypertension* 37 (2019), e158.
- [126] Tayler H Long et al. “The correlation between the severity of peripheral arterial disease and carotid occlusive disease”. In: *Vascular Medicine* 4.3 (1999), pp. 135–142.
- [127] K Low et al. “An improved baseline model for a human arterial network to study the impact of aneurysms on pressure-flow waveforms”. In: *International journal for numerical methods in biomedical engineering* 28.12 (2012), pp. 1224–1246.
- [128] Albert Maroto et al. “Brachial and femoral artery blood flow in cirrhosis: relationship to kidney dysfunction”. In: *Hepatology* 17.5 (1993), pp. 788–793.
- [129] Ellisiv B Mathiesen, Oddmund Joakimsen, and Kaare H Bønaa. “Prevalence of and risk factors associated with carotid artery stenosis: the Tromsø Study”. In: *Cerebrovascular diseases* 12.1 (2001), pp. 44–51.
- [130] Ajay Mathur and Giles M Foody. “Multiclass and binary SVM classification: Implications for training and classification users”. In: *IEEE Geoscience and remote sensing letters* 5.2 (2008), pp. 241–245.
- [131] Koen S Matthys et al. “Pulse wave propagation in a model human arterial network: assessment of 1-D numerical simulations against in vitro measurements”. In: *Journal of biomechanics* 40.15 (2007), pp. 3476–3486.
- [132] L Daniel Maxim, Ron Niebo, and Mark J Utell. “Screening tests: a review with examples”. In: *Inhalation toxicology* 26.13 (2014), pp. 811–828.
- [133] Allyn G May, James A Dewese, and Charles G Rob. “Hemodynamic effects of arterial stenosis”. In: *Surgery* 53.4 (1963), pp. 513–524.
- [134] Henry C McGill and Jack P Strong. “The geographic pathology of atherosclerosis”. In: *Annals of the New York Academy of Sciences* 149.1 (1968), pp. 923–927.

- [135] Matthew McKenna, Sidney Wolfson, and Lewis Kuller. “The ratio of ankle and arm arterial pressure as an independent predictor of mortality”. In: *Atherosclerosis* 87.2 (1991), pp. 119–128.
- [136] Ismail Bin Mohamad and Dauda Usman. “Standardization and its effects on K-means clustering algorithm”. In: *Research Journal of Applied Sciences, Engineering and Technology* 6.17 (2013), pp. 3299–3303.
- [137] Wesley S Moore et al. “Abdominal aortic aneurysm: a 6-year comparison of endovascular versus transabdominal repair”. In: *Annals of surgery* 230.3 (1999), p. 298.
- [138] JOSEPH P Murgu et al. “Aortic input impedance in normal man: relationship to pressure wave forms.” In: *Circulation* 62.1 (1980), pp. 105–116.
- [139] Kevin P Murphy et al. “Naive bayes classifiers”. In: *University of British Columbia* 18.60 (2006).
- [140] Fionn Murtagh. “Multilayer perceptrons for classification and regression”. In: *Neurocomputing* 2.5-6 (1991), pp. 183–197.
- [141] Yurii E Nesterov. “A method for solving the convex programming problem with convergence rate $O(1/k^2)$ ”. In: *Dokl. Akad. Nauk SSSR*. Vol. 269. 1983, pp. 543–547.
- [142] Lars Norgren et al. “Inter-society consensus for the management of peripheral arterial disease (TASC II)”. In: *Journal of vascular surgery* 45.1 (2007), S5–S67.
- [143] Greg Novack. “A Defense of the Principle of Indifference”. In: *Journal of philosophical logic* 39.6 (2010), pp. 655–678.
- [144] Ammar A Oglat et al. “A review of medical doppler ultrasonography of blood flow in general and especially in common carotid artery”. In: *Journal of medical ultrasound* 26.1 (2018), p. 3.
- [145] Mette S Olufsen et al. “Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions”. In: *Annals of biomedical engineering* 28.11 (2000), pp. 1281–1299.
- [146] Michael F O’rourke. “Carotid Artery Tonometry: Pros and Cons”. In: *American journal of hypertension* 29.3 (2015), pp. 296–298.

- [147] Thomas Lee Paez and Jerome Scot Cap. *Generation of Random Process Realizations Via Markov Chain Monte Carlo*. Tech. rep. Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2006.
- [148] Sanjay Pant et al. “A methodological paradigm for patient-specific multi-scale CFD simulations: from clinical measurements to parameter estimates for individual analysis”. In: *International journal for numerical methods in biomedical engineering* 30.12 (2014), pp. 1614–1648.
- [149] Francesco Pappalardo et al. “In silico clinical trials: concepts and early adoptions”. In: *Briefings in Bioinformatics* 20.5 (2019), pp. 1699–1708.
- [150] Gianfranco Parati et al. “Comparison of finger and intra-arterial blood pressure monitoring at rest and during laboratory testing.” In: *Hypertension* 13.6_pt.1 (1989), pp. 647–655.
- [151] Kim H Parker. “An introduction to wave intensity analysis”. In: *Medical & biological engineering & computing* 47.2 (2009), p. 175.
- [152] Alfredo L Pauca et al. “Does radial artery pressure accurately reflect aortic pressure?” In: *Chest* 102.4 (1992), pp. 1193–1198.
- [153] L Mihaela Paun and Dirk Husmeier. “Markov chain Monte Carlo with Gaussian processes for fast parameter estimation and uncertainty quantification in a 1D fluid-dynamics model of the pulmonary circulation”. In: *International journal for numerical methods in biomedical engineering* (2020), e3421.
- [154] L Mihaela Paun et al. “Assessing model mismatch and model selection in a Bayesian uncertainty quantification analysis of a fluid-dynamics model of pulmonary blood circulation”. In: *Journal of the Royal Society Interface* 17.173 (2020), p. 20200886.
- [155] L Mihaela Paun et al. “MCMC methods for inference in a mathematical model of pulmonary circulation”. In: *Statistica Neerlandica* 72.3 (2018), pp. 306–338.
- [156] Fabian Pedregosa et al. “Scikit-learn: Machine learning in Python”. In: *the Journal of machine Learning research* 12 (2011), pp. 2825–2830.

-
- [157] Patricia A Pellikka et al. “Outcome of 622 adults with asymptomatic, hemodynamically significant aortic stenosis during prolonged follow-up”. In: *Circulation* 111.24 (2005), pp. 3290–3295.
- [158] Noud Peppelenbosch et al. “Diameter of abdominal aortic aneurysm and outcome of endovascular aneurysm repair: does size matter? A report from EUROSTAR”. In: *Journal of vascular surgery* 39.2 (2004), pp. 288–297.
- [159] Kemal Polat and Salih Güneş. “Breast cancer diagnosis using least square support vector machine”. In: *Digital signal processing* 17.4 (2007), pp. 694–701.
- [160] Lutz Prechelt. “Early stopping-but when?” In: *Neural Networks: Tricks of the trade*. Springer, 1998, pp. 55–69.
- [161] Ning Qian. “On the momentum term in gradient descent learning algorithms”. In: *Neural networks* 12.1 (1999), pp. 145–151.
- [162] Alfio Quarteroni and Luca Formaggia. “Mathematical modelling and numerical simulation of the cardiovascular system”. In: *Handbook of numerical analysis* 12 (2004), pp. 3–127.
- [163] Denis S Quill, Mary Paula Colgan, and David S Sumner. “Ultrasonic screening for the detection of abdominal aortic aneurysms”. In: *Surgical Clinics of North America* 69.4 (1989), pp. 713–720.
- [164] G Radegran. “Ultrasound Doppler estimates of femoral artery blood flow during dynamic knee extensor exercise in humans”. In: *Journal of Applied Physiology* 83.4 (1997), pp. 1383–1388.
- [165] Adrian E Raftery and Steven M Lewis. “Implementing mcmc”. In: *Markov chain Monte Carlo in practice* (1996), pp. 115–130.
- [166] Calyampudi Radhakrishna Rao and Helge Toutenburg. “Linear models”. In: *Linear models*. Springer, 1995, pp. 3–18.
- [167] Carl Edward Rasmussen. “Gaussian processes in machine learning”. In: *Advanced lectures on machine learning*. Springer, 2004, pp. 63–71.
- [168] Philippe Reymond et al. “Validation of a one-dimensional model of the systemic arterial tree”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 297.1 (2009), H208–H222.
-

-
- [169] Philippe Reymond et al. “Validation of a patient-specific one-dimensional model of the systemic arterial tree”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 301.3 (2011), H1173–H1182.
- [170] Ryan Rifkin and Aldebaro Klautau. “In defense of one-vs-all classification”. In: *Journal of machine learning research* 5.Jan (2004), pp. 101–141.
- [171] Irina Rish, Joseph Hellerstein, and Jayram Thathachar. “An analysis of data characteristics that affect naive Bayes performance”. In: *IBM TJ Watson Research Center* 30 (2001), pp. 1–8.
- [172] Irina Rish et al. “An empirical study of the naive Bayes classifier”. In: *IJCAI 2001 workshop on empirical methods in artificial intelligence*. Vol. 3. 22. 2001, pp. 41–46.
- [173] Gareth O Roberts, Andrew Gelman, Walter R Gilks, et al. “Weak convergence and optimal scaling of random walk Metropolis algorithms”. In: *The annals of applied probability* 7.1 (1997), pp. 110–120.
- [174] Anderson Rocha and Siome Klein Goldenstein. “Multiclass from binary: Expanding one-versus-all, one-versus-one and ecoc-based approaches”. In: *IEEE Transactions on Neural Networks and Learning Systems* 25.2 (2013), pp. 289–302.
- [175] John Ross and Eugene Braunwald. “Aortic stenosis”. In: *Circulation* 38.1S5 (1968), pp. V–61.
- [176] Vivekananda Roy. “Convergence diagnostics for Markov chain Monte Carlo”. In: *Annual Review of Statistics and Its Application* 7 (2020), pp. 387–412.
- [177] Sebastian Ruder. “An overview of gradient descent optimization algorithms”. In: *arXiv preprint arXiv:1609.04747* (2016).
- [178] Joseph W Rudmin. “Calculating the exact pooled variance”. In: *arXiv preprint arXiv:1007.1012* (2010).
- [179] Natzi Sakalihasan, Raymond Limet, and Olivier Damien Defawe. “Abdominal aortic aneurysm”. In: *The Lancet* 365.9470 (2005), pp. 1577–1589.
- [180] Paolo Salvi. “Pulse waves”. In: *How vascular hemodynamics affects Blood pressure* (2012).
-

- [181] Yutaka Sasaki. “The truth of the F-measure”. In: *Teach Tutor Mater* (Jan. 2007).
- [182] Yutaka Sasaki et al. “The truth of the F-measure”. In: *Teach Tutor mater* 1.5 (2007), pp. 1–5.
- [183] Igor Sazonov et al. “A novel method for non-invasively detecting the severity and location of aortic aneurysms”. In: *Biomechanics and modeling in mechanobiology* 16.4 (2017), pp. 1225–1242.
- [184] Bernhard Scholkopf et al. “Comparing support vector machines with Gaussian kernels to radial basis function classifiers”. In: *IEEE transactions on Signal Processing* 45.11 (1997), pp. 2758–2765.
- [185] Howard D Sesso et al. “Systolic and diastolic blood pressure, pulse pressure, and mean arterial pressure as predictors of cardiovascular disease risk in men”. In: *Hypertension* 36.5 (2000), pp. 801–807.
- [186] Ramin Shadman et al. “Subclavian artery stenosis: prevalence, risk factors, and association with cardiovascular diseases”. In: *Journal of the American College of Cardiology* 44.3 (2004), pp. 618–623.
- [187] Matthew Shardlow. “An analysis of feature selection techniques”. In: ().
- [188] Margaret A Shipp et al. “Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning”. In: *Nature medicine* 8.1 (2002), p. 68.
- [189] Maria Siebes et al. “Potential and limitations of wave intensity analysis in coronary arteries”. In: *Medical & biological engineering & computing* 47.2 (2009), pp. 233–239.
- [190] Birgitta Sigvant et al. “A population-based study of peripheral arterial disease prevalence with special focus on critical limb ischemia and sex differences”. In: *Journal of vascular surgery* 45.6 (2007), pp. 1185–1191.
- [191] Karri Silventoinen et al. “Heritability of adult body height: a comparative study of twin cohorts in eight countries”. In: *Twin Research and Human Genetics* 6.5 (2003), pp. 399–408.

- [192] Björn Sonesson et al. “Intra-aneurysm pressure measurements in successfully excluded abdominal aortic aneurysm after endovascular repair”. In: *Journal of vascular surgery* 37.4 (2003), pp. 733–738.
- [193] Mi Hye Song et al. “Support vector machine based arrhythmia classification using reduced features”. In: *International Journal of Control, Automation, and Systems* 3.4 (2005), pp. 571–579.
- [194] Sandro Sperandei. “Understanding logistic regression analysis”. In: *Biochemia medica: Biochemia medica* 24.1 (2014), pp. 12–18.
- [195] N Stergiopoulos, DF Young, and TR Rogge. “Computer simulation of arterial flow with applications to arterial and aortic stenoses”. In: *Journal of biomechanics* 25.12 (1992), pp. 1477–1488.
- [196] Ingvar Strid. “Efficient parallelisation of Metropolis–Hastings algorithms using a prefetching approach”. In: *Computational Statistics & Data Analysis* 54.11 (2010), pp. 2814–2835.
- [197] Rajesh Subramanian et al. “Renal fractional flow reserve: a hemodynamic evaluation of moderate renal artery stenoses”. In: *Catheterization and cardiovascular interventions* 64.4 (2005), pp. 480–486.
- [198] Jun Sugawara et al. “Brachial–ankle pulse wave velocity: an index of central arterial stiffness?” In: *Journal of human hypertension* 19.5 (2005), pp. 401–406.
- [199] Motoaki Sugawara et al. “Relationship between the pressure and diameter of the carotid artery in humans”. In: *Heart and vessels* 15.1 (2000), pp. 49–51.
- [200] Michael J Sweeting et al. “Analysis of clinical benefit, harms, and cost-effectiveness of screening women for abdominal aortic aneurysm”. In: *The Lancet* 392.10146 (2018), pp. 487–495.
- [201] Abigail Swillens et al. “Effect of an abdominal aortic aneurysm on wave reflection in the aorta”. In: *IEEE Transactions on Biomedical Engineering* 55.5 (2008), pp. 1602–1611.
- [202] Kenji Takazawa et al. “Relationship between radial and central arterial pulse wave and evaluation of central aortic pressure using the radial arterial pulse wave”. In: *Hypertension Research* 30.3 (2007), p. 219.

- [203] Robert F Tate. “Correlation between a discrete and a continuous variable. Point-biserial correlation”. In: *The Annals of mathematical statistics* 25.3 (1954), pp. 603–607.
- [204] SG Thompson et al. “Screening men for abdominal aortic aneurysm: 10 year mortality and cost effectiveness results from the randomised Multicentre Aneurysm Screening Study”. In: *Bmj* 338 (2009), b2307.
- [205] M Titi et al. “Comparison of carotid Doppler ultrasound and computerised tomographic angiography in the evaluation of carotid artery stenosis”. In: *The Surgeon* 5.3 (2007), pp. 132–136.
- [206] Georgi P Tolstov. *Fourier series*. Courier Corporation, 2012.
- [207] Eric J Udoff et al. “Hemodynamic significance of iliac artery stenosis: pressure measurements during angiography”. In: *Radiology* 132.2 (1979), pp. 289–293.
- [208] Hiroshi Ujiie et al. “Effects of size and shape (aspect ratio) on the hemodynamics of saccular aneurysms: a possible index for surgical treatment of intracranial aneurysms”. In: *Neurosurgery* 45.1 (1999), pp. 119–130.
- [209] Padmini Varadarajan et al. “Clinical profile and natural history of 453 nonsurgically managed patients with severe aortic stenosis”. In: *The Annals of thoracic surgery* 82.6 (2006), pp. 2111–2115.
- [210] Ramachandran S Vasan et al. “Impact of high-normal blood pressure on the risk of cardiovascular disease”. In: *New England journal of medicine* 345.18 (2001), pp. 1291–1297.
- [211] Eric Vittinghoff and Charles E McCulloch. “Relaxing the rule of ten events per variable in logistic and Cox regression”. In: *American journal of epidemiology* 165.6 (2007), pp. 710–718.
- [212] Charalambos Vlachopoulos, Michael O’Rourke, and Wilmer W Nichols. *McDonald’s blood flow in arteries: theoretical, experimental and clinical principles*. CRC press, 2011.
- [213] Eberhard O Voit. “Perspective: Dimensions of the scientific method”. In: *PLoS Computational Biology* 15.9 (2019), e1007279.

- [214] Tianqi Wang et al. “Machine learning-based pulse wave analysis for early detection of abdominal aortic aneurysms using in silico pulse waves”. In: *Symmetry* 13.5 (2021), p. 804.
- [215] Nico Westerhof, Jan-Willem Lankhaar, and Berend E Westerhof. “The arterial windkessel”. In: *Medical & biological engineering & computing* 47.2 (2009), pp. 131–141.
- [216] Richard M Wiard et al. “Automatic detection of motion artifacts in the ballistocardiogram measured on a modified bathroom scale”. In: *Medical & biological engineering & computing* 49.2 (2011), pp. 213–220.
- [217] Marie Willemet, Phil Chowienczyk, and Jordi Alastruey. “A database of virtual healthy subjects to assess the accuracy of foot-to-foot pulse wave velocities for estimation of aortic stiffness”. In: *American Journal of Physiology-Heart and Circulatory Physiology* 309.4 (2015), H663–H675.
- [218] Dachun Xu et al. “Sensitivity and specificity of the ankle—brachial index to diagnose peripheral artery disease: a structured review”. In: *Vascular Medicine* 15.5 (2010), pp. 361–369.
- [219] Akira Yamashina et al. “Validity, reproducibility, and clinical significance of noninvasive brachial-ankle pulse wave velocity measurement”. In: *Hypertension research* 25.3 (2002), pp. 359–364.
- [220] Yuan Yao, Lorenzo Rosasco, and Andrea Caponnetto. “On early stopping in gradient descent learning”. In: *Constructive Approximation* 26.2 (2007), pp. 289–315.
- [221] Dongping Yin et al. “Cost-effectiveness of MR angiography in cases of limb-threatening peripheral vascular disease.” In: *Radiology* 194.3 (1995), pp. 757–764.
- [222] Evangelia I Zacharaki et al. “Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme”. In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 62.6 (2009), pp. 1609–1618.
- [223] Matthew D Zeiler. “ADADELTA: an adaptive learning rate method”. In: *arXiv preprint arXiv:1212.5701* (2012).

- [224] Zhengze Zhou and Giles Hooker. “Unbiased measurement of feature importance in tree-based methods”. In: *arXiv preprint arXiv:1903.05179* (2019).