Integrated Optimisation of Upstream and Downstream Processing in Biopharmaceutical Manufacturing under Uncertainty: A Chance Constrained Programming Approach

Songsong Liu[a], Suzanne S. Farid[b], Lazaros G. Papageorgiou[a,*]

[a] *Centre for Process Systems Engineering, Department of Chemical Engineering, UCL (University College London), London WC1E 7JE, United Kingdom*

[b] *The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, UCL (University College London), London WC1E 7JE, United Kingdom*

Abstract

This work addresses the integrated optimisation of upstream and downstream processing strategies of a monoclonal antibody (mAb) under uncertainty. In the upstream processing (USP), the bioreactor sizing strategies are optimised, while in the downstream processing (DSP), the chromatography sequencing and column sizing strategies, including the resin at each chromatography step, the number of columns, the column diameter and bed height, and the number of cycles per batch, are determined. Meanwhile, the product's purity requirement is considered. Under the uncertainties of both upstream titre and chromatography resin yields, a stochastic mixed integer linear programming (MILP) model is developed, using chance constrained programming (CCP) techniques, to minimise the total cost of goods (COG). The model is applied to an industrially-relevant example and the impact of different USP:DSP ratios is studied. The computational results of the stochastic optimisation model illustrate its advantage over the deterministic model. Also, the benefit of the integrated optimisation of both USP and DSP is demonstrated. The sensitivity analysis of both the confidence level used in the CCP model and the initial impurity level is investigated as well.

Keywords: biopharmaceutical manufacturing, integrated optimisation, uncertainty, chance constrained programming, mixed integer linear programming

* Corresponding author. E-mail: l.papageorgiou@ucl.ac.uk. Tel: +44 20 7679 2563. Fax: +44 20 7679 7092.

1. Introduction

Monoclonal antibodies (mAbs) represent one of the fastest growing groups of therapeutic biopharmaceutical drugs. The total global sales revenue of all mAb drugs in 2013 was almost US$75 billion, approximately half of the total sales of all biopharmaceutical drugs, and the growth of the sales of mAb drugs will continue in the near future[1]. The upstream processing (USP) of mAbs manufacturing process is becoming more challenging, due to the significant improvements in USP productivities with higher mAb upstream titres being achieved in cell culture. The downstream processing (DSP), in which chromatography operations are critical steps representing a large proportion of the total manufacturing cost, needs to match the increased titres in the USP[2]. Thus, it is critical to identify cost-effective manufacturing processes to overcome manufacturing bottlenecks through the optimisation of the design and operation of both upstream fermentation and downstream chromatography steps.

Optimisation-based decision-support tools have been developed for the biopharmaceutical manufacturing using mathematical programming techniques in the literature, e.g., on the biopharmaceutical production planning and scheduling[3-7], capacity planning[8], purification process synthesis[9-12]. Several optimisation-based approaches have been developed for the optimal design of downstream purification processes. A meta-heuristic optimisation approach was developed using genetic algorithms, and applied it to a case study on the production of mAbs, focusing on the optimal purification sequences and chromatography column sizing strategies, in order to cope with different configurations of upstream and downstream trains and product impurity loads[13]. A real-world problem concerned with the discovery of cost-effective equipment sizing strategies for purification processes was addressed, which was formulated as a combinatorial closed-loop optimisation problem and solved using evolutionary algorithms (EA)[14-16]. In addition, a mixed integer linear programming (MILP) model was developed for the optimal chromatography column sizing decisions in the mAb manufacturing, considering an objective of the minimisation of cost of goods per gram (COG/g) as the above work, for different facility configurations[17]. Later, the above work was extended to a mixed integer nonlinear programming (MINLP) model for the facility design with the current titre and facility fit with higher titres[18]. The authors further optimised both chromatography sequencing and column sizing decisions using mixed integer linear fractional

programming (MILFP) and Dinkelbach's algorithm with decision variables as discrete or continuous[19,20]. Another work developed a mixed-integer dynamic optimisation model for optimal development of bioprocesses, which was solved by a hybrid simulation-optimisation decomposition algorithm[21]. These works have focused on optimising DSP but not optimising USP and DSP simultaneously, which is one of the objectives of this paper.

In industrial practice, the upstream fermentation titre and downstream chromatography resin yield are usually uncertain, which can affect the performance of the whole manufacturing process significantly, and should be taken into account in the optimisation model. The chance constrained programming (CCP) approach, firstly introduced by Charnes and Cooper[22], and has been widely used in engineering applications to deal with uncertainties[3,23-28]. Under the CCP approach, the optimal solution ensures a specified probability of complying with constraints, i.e., the confidence level of being feasible. Compared to stochastic programming, the CCP approach usually requires much fewer constraints and equations that are similar in number to the deterministic model, and has a better performance in terms of ease of solution. This work aims to develop an optimisation framework by extending our previous work[17-20] to address the integrated optimisation of strategies for both USP and DSP for the mAb manufacturing, using CCP techniques to deal with the uncertainties of the upstream titre and chromatography resin yield. In the USP, the optimal design of the upstream facilities is considered by determining the optimal bioreactor volumes with given configurations of USP and DSP trains, while in the DSP we consider the optimisation of both the chromatography columns sizing and chromatography sequencing strategies. Due to the interactions among the manufacturing operations, these strategies should be optimised simultaneously to obtain the optimal decisions.

The rest of this paper is organised as follows: Section 2 describes the integrated optimisation problem of both USP and DSP design. The mathematical formulation of the proposed stochastic integrated optimisation model is presented in Section 3. In Section 4, an industrially-relevant example is presented, and its computational results and discussion are given in Section 5. Finally, the concluding remarks are made in Section 6.

2. Problem Statement

This work addresses the integrated optimisation of the mAb manufacturing strategies, including the bioreactor sizing strategies in the USP, and the column sequencing and sizing strategies in the DSP. Fig. 1 shows a typical mAb platform process studied in this work. After mammalian cells expressing the mAb are cultured in bioreactors in the USP (see "Cell culture" in Fig. 1), the mAb is recovered, purified and cleared from potential viruses by a number of DSP operations, including three packed-bed chromatography steps, i.e. capture, intermediate purification and polishing steps. These DSP steps operate in a non-overlapping mode, which means when a DSP step is operating, other DSP steps are idle.

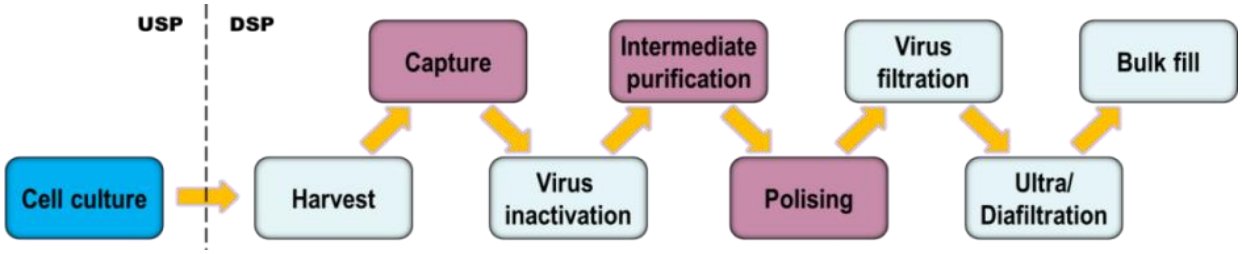

Fig. 1. A typical mAb manufacturing process

Here, the optimal volumes of bioreactors in the USP are considered as decision variables for optimisation, rather than pre-determined as in the literature. With given configurations of USP and DSP trains, it is assumed that all bioreactors have the same volume, and they operate in the staggered pattern.

There are a number of candidate resins from different types to be selected at each chromatography step in the DSP for the optimal chromatography sequencing decisions. In order to use the orthogonal separation mechanisms, it is assumed that no more than one resin from the same type can be used in the sequence. In addition, at each chromatography step, the column sizing strategies are to be optimised, including the bed height, diameter, number of cycles and number of columns to run in parallel, whose optimal values are to be chosen from a set of discrete candidate values.

In addition, the purity requirement of final product is considered. The host cell proteins (HCPs) are considered as the critical impurities, which may cause antigenic effects in patients and must be removed during DSP[29]. A target level of the HCPs in the final product specification is to be met by

the selected chromatography sequence, where each resin will have a different HCP reduction capability.

In this problem, the upstream titre and chromatography step yield with each resin are considered as uncertain parameters which follow triangular distributions[14-16,30]. It is assumed that the upstream titre and step yields remain the same for different batches. The given annual demand should be met by the annual output. In our previous work[17-20], the optimisation objective was to minimise the COG/g, which is equal to the annual total cost divided by the annual total output. However, in this problem, as the annual demand is required to be met, i.e. the total output is equal to the demand, the minimisation of COG/g is equivalent to the minimisation of the total cost of goods (COG), which comprises of both direct and indirect operating costs.

The problem addressed in this work can be described as follows.

Given:

- process sequence of a mAb product;
- a number of USP trains, and upstream titre;
- candidate chromatography resins at each step, and their key characteristics (e.g. yield, linear velocity, buffer usage, dynamic binding capacity, HCP reduction value);
- key characteristics of non-chromatography steps (e.g. yield, time, buffer usage);
- cost data (e.g., reference equipment costs, labour wage, resin, buffer and media prices);
- candidate column diameters and heights, numbers of cycles and columns;
- initial and target HCP levels;
- probability distributions of both titre and chromatography step yield;

determine:

- bioreactor sizing strategies;
- chromatography step sequencing strategies (i.e., resin for each step);
- chromatography column sizing strategies (i.e. column diameter and bed height, number of columns per step, and number of cycles per batch) at each chromatography step;
- product mass and volume, and buffer usage volume;
- number of total completed batches;
- annual total processing time;

so as to:

minimise the annual total COG, including both direct and indirect costs.

3. Mathematical Formulation

In this section, a deterministic MILP model for the optimal USP bioreactor sizing, DSP chromatography sequencing and sizing decisions is initially presented. Then, it is extended to a stochastic CCP model to deal with uncertainties in upstream titre and chromatography resin yields. Both models are developed based on the MILP model for DSP optimisation[19], which is presented in the Supporting Information. The proposed model for the whole biopharmaceutical manufacturing process involves an unavoidable complexity with a large number of constraints and variables to track the product mass, product volume, buffer volume and processing time of each downstream operation, calculate many direct and indirect cost terms, and linearise the nonlinear constraints. Only the newly developed equations in this work are presented in this section.

3.1. Deterministic Formulation

Firstly, a deterministic MILP model was developed using exact linearisation and piecewise linearisation techniques to maximise the total COG. The key difference between the proposed model and the literature model[19] lies in modelling the bioreactor size as a variable. Thus, only the constraints related to the bioreactor size are presented here. Please refer to the appendix for other constraints involved.

3.1.1. Protein Mass

In each batch, the initial protein mass going out of the upstream processes depends on the titre of the product, $titre$, and the working volume of bioreactor:

$$M_0 = titre \cdot \alpha \cdot BRV \tag{1}$$

where α is the working volume ratio, and BRV is a variable for single bioreactor volume.

An auxiliary variable, $\overline{UM}_{sn} \equiv U_{sr} \cdot M_{s-1}$, is introduced to calculate the protein mass at chromatography step s, M_s, supported by the following constraints:

$$\overline{UM}_{s-1,r} \leq titre \cdot \alpha \cdot maxbrv \cdot U_{sr}, \quad \forall s \in CS, r \in R_s \tag{2}$$

$$\sum_{r\in R_s} \overline{UM}_{s-1,r} = M_{s-1}, \quad \forall s \in CS \tag{3}$$

where M_s is the product mass per batch after operation *s*, and binary variable U_{sr} is equal to 1 if resin *r* is selected for chromatography step *s*.

Another auxiliary variable $\overline{ZM}_{sn} \equiv Z_n \cdot M_s$, is introduced for the calculation of annual product output, AP, with the following auxiliary constraints:

$$\overline{ZM}_{sn} \leq titre \cdot \alpha \cdot maxbrv \cdot Z_n, \ \ \forall s = bf, n = 1, \dots, q \tag{4}$$

$$\overline{ZM}_{sn} \leq M_s, \ \ \forall s = bf, n = 1, \dots, q \tag{5}$$

$$\overline{ZM}_{sn} \geq M_s - titre \cdot \alpha \cdot maxbrv \cdot (1 - Z_n), \ \ \forall s = bf, n = 1, \dots, q \tag{6}$$

where Z_n is a binary variable to indicate whether if the *n*th digit of the binary representation of variable BN is equal to 1.

Different from the literature models[17-20], the annual product output is required to meet the annual demand in this work, to avoid unreasonably large bioreactor size. In this case, more bioreactors may result in more batches with smaller batch size, i.e. smaller bioreactor volume.

$$AP = dem \tag{7}$$

3.1.2. Product Volume

For each batch, the initial product volume from the USP, PV_0, is equal to the working volume of the bioreactor.

$$PV_0 = \alpha \cdot BRV \tag{8}$$

3.1.3. Purity

The purity of the mAb product after the purification processes should meet certain requirements. HCPs, which are produced or encoded by the organisms and unrelated to the intended mAb product, are considered as the critical impurities in the considered processes. In the USP, the initial level of the HCPs, $inihcp$ (in ng of HCP per mg of mAb product), is known. In the DSP, the ability to remove HCPs by each resin is measured in terms of log reduction value (LRV), lrv_{sr}. The selected purification process should be able to reduce the HCP level below a target value, $tarhcp$, in order to meet the purity requirement of the final product.

$$\frac{inihcp}{10^{\sum_{s\in CS}\sum_{r\in R_s} lrv_{sr}\cdot U_{sr}}} \leq tarhcp \quad (9)$$

which is equivalent to the following linear constraint by applying the logarithm operation:

$$\lg inihcp - \sum_{s\in CS}\sum_{r\in R_s} lrv_{sr}\cdot U_{sr} \leq \lg tarhcp \quad (10)$$

3.1.4. Costs

The media cost is determined by the media overfill allowance, θ, the media price, $mepc$, the bioreactor working volume, $\alpha \cdot BRV$, and the number of batches, BN:

$$MEC = \theta \cdot mepc \cdot \alpha \cdot BRV \cdot BN \quad (11)$$

which is linearised by introducing $\overline{ZBRV_n} \equiv Z_n \cdot BRV$ and the following constraints:

$$\overline{ZBRV_n} \leq maxbrv \cdot Z_n, \ \ \forall n = 1, \dots, q \quad (12)$$

$$\overline{ZBRV_n} \leq BRV, \ \ \forall n = 1, \dots, q \quad (13)$$

$$\overline{ZBRV_n} \geq BRV - maxbrv \cdot (1 - Z_n), \ \ \forall n = 1, \dots, q \quad (14)$$

Thus, we rewrite Eq. (11) as follows:

$$MEC = \theta \cdot mepc \cdot \alpha \cdot \sum_{n=1}^{q} 2^{n-1} \cdot \overline{ZBRV_n} \quad (15)$$

The utilities cost, UC, is assumed to have three parts, which are proportional to the total bioreactor volume, $brn \cdot BRV$, number of completed batches by single bioreactor volume, $BRV \cdot BN$, and annual buffer volume, ABV, respectively.

$$UC = a \cdot brn \cdot BRV + b \cdot BRV \cdot BN + c \cdot ABV \quad (16)$$

where a, b, c are coefficient to calculate the utilities cost. Given Eqs. (12)-(14), the bilinear term $BRV \cdot BN$ can be linearised, and Eq. (16) can be replaced by:

$$UC = a \cdot brn \cdot BRV + b \cdot \sum_{n=1}^{q} 2^{n-1} \cdot \overline{ZBRV_n} + c \cdot ABV \quad (17)$$

The fixed capital investment consists of the capital investment for bioreactors, chromatography columns, as well as other equipment in DSP, whose cost is assumed to be proportional to the bioreactor cost. Here, Lang factor, $lang$, is used to approximate the fixed capital investment[31]. In addition, the general equipment factor, *gef*, is used to take into account the costs of support equipment, spares, utilities equipment and vessels.

$$FCI = lang \cdot (1 + gef) \cdot (brn \cdot BRC + \sum_{s\in CS}\sum_{i} cc_{si} \cdot CN_{si} + oe\lambda \cdot brn \cdot BRC) \quad (18)$$

where $cc_{si} = refcc \cdot \left(\frac{dm_{si}}{refdm}\right)^{cf}$, in which $refcc$ is the cost of a chromatography column with a reference diameter of $refdm$, and cf is the scale-up factor for chromatography column cost. Also, *gef* refers to the general equipment factor taking into account the costs of support equipment, spares, utilities equipment and vessels.

In addition, variable BRC in Eq. (18) is the cost of one bioreactor, which is dependent on the bioreactor volume. Here, we use piecewise linearisation approximation to calculate BRC. Given a number of discrete bioreactor volumes, $dbrv_m$, their corresponding bioreactor costs are given by $dbrc_m = refbrc \cdot \left(\frac{dbrv_m}{refbrv}\right)^{brf}$, where $refbrv$ is the reference volume of one bioreactor; $refbrc$ is the cost of one bioreactor with a volume of $refbrv$; and brf is the scale-up factor for bioreactor. Thus, by introducing a SOS2 variable, Λ_m, we have:

$$BRC = \sum_m dbrc_m \cdot \Lambda_m \tag{19}$$

$$BRV = \sum_m dbrv_m \cdot \Lambda_m \tag{20}$$

$$\sum_m \Lambda_m = 1 \tag{21}$$

The general utilities cost, GUC, included in the other indirect costs, depends on the total bioreactor volume, i.e., the product of the number of bioreactors and the bioreactor volume.

$$GUC = gu \cdot brn \cdot BRV \tag{22}$$

3.1.5. Objective Function

The objective of this optimisation problem is to minimise the total COG, including both direct and indirect costs (as defined in Eq. (S.84) in the Supporting Information).

$$OBJ = COG \tag{23}$$

Overall, the optimisation problem is formulated as a deterministic MILP model (denoted as DM) with Eqs. (1)-(8), (10), (12)-(15), (17)-(22), as well as Eqs. (S.1)-(S.7), (S.9), (S.10), (S.13), (S.17)-(S.28), (S.30)-(S.69), (S.71)-(S.75), (S.77), (S.79)-(S.81), (S.83) and (S.84) in the Supporting Information, as constraints, and Eq. (23) as the objective function.

3.2. Stochastic Formulation

Based on the above deterministic formulation, the uncertain upstream titre and chromatography resin yields, which follow a triangular probability distribution, were taken into account for the stochastic formulation. The classic CCP approach was used to tackle these uncertainties. In the CCP approach, the decision maker firstly expresses a risk tolerance, i.e., a permissible probability of violation in the constraints involving uncertain parameters. Expressed by the inverse cumulative distribution function, the developed chance constraints are transformed into their deterministic equivalent formulations. The details of the new chance constraints under uncertainty in the CCP approach are presented in this section.

3.2.1. Chance constraints under uncertain titre

To develop a chance constraint for the uncertain parameter, *titre*, the constraint Eq. (1) in the previous section, should be firstly converted into an inequality as shown in Eq. (24), as M_0 is maximised to satisfy the demand.

$$M_0 \leq titre \cdot \alpha \cdot BRV \tag{24}$$

The corresponding chance constraint of Eq. (24) is formulated as follows:

$$\Pr(M_0 \leq titre \cdot \alpha \cdot BRV) \geq A^t \tag{25}$$

where A^t is a minimum prespecified probability that Eq. (24) will hold true, greater than 50%.

The above Eq. (25) can be written as follows:

$$\Pr\left(titre \geq \frac{M_0}{\alpha \cdot BRV}\right) \geq A^t \tag{26}$$

which is equivalent to

$$1 - \Pr\left(titre \leq \frac{M_0}{\alpha \cdot BRV}\right) \geq A^t \tag{27}$$

Here, it is assumed that the upstream titre follows a triangular probability distribution, $\mathrm{Tr}(titre^{lo}, titre^{mo}, titre^{up})$, where $titre^{lo}$ and $titre^{up}$ are lower and upper bounds of the fermentation titre, and $titre^{mo}$ is the mode. Its cumulative distribution function is denoted as $\Phi(x)$. Thus, we rewrite Eq. (27) as:

$$\Phi\left(\frac{M_0}{\alpha \cdot BRV}\right) \leq 1 - A^t \tag{28}$$

Thus, the deterministic equivalent formulation of Eq. (24) is as follows:

$$M_0 \leq \Phi^{-1}(1 - A^t) \cdot \alpha \cdot BRV \tag{29}$$

For a special case of the isosceles triangular distribution that $titre^{up} - titre^{mo} = titre^{mo} - titre^{lo} = \Delta titre$, when $A^t > 50\%$, we have:

$$\Phi^{-1}(1 - A^t) = titre^{lo} + \sqrt{2(1 - A^t)} \cdot \Delta titre \tag{30}$$

3.2.2. Chance constraints under uncertain yields

It is assumed that the realisation of uncertain chromatography yield for each resin depends on the chromatography step it is selected for. To model the uncertainty of chromatography yield, we introduce an uncertain parameter, cyd_s, to denote the deviation of the selected resin from its standard value at chromatography step *s*. Thus, we can rewrite the constraint involving the resin yield, (Eq. S.10 in Supporting Information), as follows:

$$M_s = \sum_{r \in R_s} cy_{sr} \cdot \overline{UM}_{s-1,r} \cdot cyd_s, \quad \forall s \in CS \tag{31}$$

Similarly, the above constraint is converted into inequality, and its corresponding chance constraint can be formulated as:

$$\Pr\left(M_s \leq \sum_{r \in R_s} cy_{sr} \cdot \overline{UM}_{s-1,r} \cdot cyd_s\right) \geq A_s^y, \quad \forall s \in CS \tag{32}$$

It is assumed that cyv_s follows a triangular distribution $\mathrm{Tr}\left(cyd_s^{lo}, cyd_s^{mo}, cyd_s^{up}\right)$, in which the mode $cyd_s^{mo} = 100\%$, cyd_s^{lo} and cyd_s^{up} are lower and upper bounds of the yield deviation at chromatography step *s*, and its cumulative distribution function is denoted as $\overline{\Phi}_s(x)$. Thus, Eq. (32) can be reformulated as:

$$\overline{\Phi}_s\left(\frac{M_s}{\sum_{r \in R_s} cy_{sr} \cdot \overline{UM}_{s-1,r}}\right) \leq 1 - A_s^y, \quad \forall s \in CS \tag{33}$$

Using the inverse cumulative distribution function expression, we have:

$$M_s \leq \overline{\Phi}_s^{-1}\left(1 - A_s^y\right) \cdot \sum_{r \in R_s} cy_{sr} \cdot \overline{UM}_{s-1,r}, \quad \forall s \in CS \tag{34}$$

In particular, if cyv_s follows an isosceles triangular distribution that $cyd_s^{up} - cyd_s^{mo} = cyd_s^{mo} - cyd_s^{lo} = \Delta cyd_s$, when $A_s^y > 50\%$, we have:

$$\bar{\Phi}_s^{-1}\left(1-A_s^y\right) = cyd_s^{lo} + \sqrt{2\left(1-A_s^y\right)} \cdot \Delta cyd_s, \quad \forall s \in CS \tag{35}$$

In summary, the deterministic equivalent formulation is an MILP model (denoted as SM) with Eqs. (2)-(8), (10), (12)-(15), (17)-(22), (29), (34), as well as Eqs. (S.1)-(S.7), (S.9), (S.13), (S.17)-(S.28), (S.30)-(S.69), (S.71)-(S.75), (S.77), (S.79)-(S.81), (S.83) and (S.84) in the Supporting Information, as constraints, and Eq. (23) as the objective function.

4. An Illustrative Example

In this section, an industrially-relevant example is presented, based on a biopharmaceutical company using a platform process for mAb purification. The considered mAb product has an annual demand of 500 kg. There are 11 candidate commercial resins from five chromatography types, including affinity chromatography (AFF), cation-exchange chromatography (CEX), anion-exchange chromatography (AEX), mixed-mode chromatography (MM) and hydrophobic interaction chromatography (HIC). The characteristics of the resin candidates are shown in Table 1. Here, each resin/resin type candidate can be suitable for more than one chromatography step, and its yield, as well as the HCP reduction capability, varies with the chromatography step where it is used. The resin utilisation factor is 0.95, and the resin overpacking factor is 1.1.

Table 1. Characteristics of the candidate resins

Resin	Type	Mode	Binding capacity (g/L)	Eluate volume (CV)	Buffer volume (CV)	Linear velocity (cm/h)	Matrix lifetime (cycle)	Matrix price (£/L)	Standard yield			HCP LRV		
									Cap.	Int.	Pol.	Cap.	Int.	Pol.
R1	AFF	BE	50	2.3	37	150	100	9200	91%	95%	-	3	1.5	-
R2	AFF	BE	30	2.3	37	300	100	6400	91%	95%	-	3	1.5	-
R3	AFF	BE	50	2.3	37	800	100	9900	91%	95%	-	3	1.5	-
R4	AFF	BE	30	2.3	37	1000	100	9000	91%	95%	-	3	1.5	-
R5	CEX	BE	120	1.4	26	500	100	2500	86%	-	-	2	-	-
R6	CEX	BE	40	1.4	26	300	100	400	86%	92%	92%	2	1	0.5
R7	AEX	FT	100	0	10	300	100	700	-	95%	95%	-	0.5	0.3
R8	MM	FT	150	0	10	375	100	3500	-	90%	90%	-	1.2	0.6
R9	MM	BE	50	1.4	26	100	100	1900	-	90%	90%	-	1.5	0.8
R10	MM	BE	35	1.4	26	250	12	2700	-	90%	90%	-	2	1
R11	HIC	BE	27.5	1.4	26	175	100	2500	-	89%	89%	-	2	0.5

As to the chromatography column sizing decisions, there are 11 discrete potential bed heights and 10 discrete potential diameters (as given in Table 2), and therefore one column has 110 potential volumes. The number of cycles per batch can be up to 10, while at most 4 parallel columns are permitted at each chromatography step.

Table 2. Candidate values of the chromatography column sizing strategies

Decision	Candidate values
Bed height (cm)	15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25
Diameter (cm)	50, 60, 70, 80, 90, 100, 120, 160, 180, 200
Number of cycles	1, 2, 3, 4, 5 , 6, 7, 8, 9, 10
Number of columns	1, 2, 3, 4

To calculate the fixed capital investment for the bioreactors and chromatography columns, the reference equipment sizes, costs, and scale-up factors in Table 3 are used. To approximate the bioreactor cost, the bioreactor volumes and costs at given discrete points in the piecewise linear function are presented in Table 4.

Table 3. Reference volumes, costs and scale-up factors

	Reference volume	Reference cost (£)	Scale-up factor
Bioreactor	Volume = 2000 L	612,000	0.6
Chromatography column	Diameter = 100 cm	170,000	0.8

Table 4. Discrete bioreactor volumes and costs for piecewise linear approximation

m	Bioreactor volume, brv_m (L)	Bioreactor cost, brc_m (£)
1	2,000	612,000
2	10,000	1,607,500
3	20,000	2,436,500
4	50,000	4,222,000

The equipment lifetime, el, is 10 years, and the interest rate, r, is 10%. Here, several cases of multiple USP trains feeding one DSP train are investigated, considering four scenarios: 1USP:1DSP, 2USP:1DSP, 4USP:1DSP and 6USP:1DSP, i.e. brn =1, 2, 4 and 6, respectively. As to the HCP levels, the initial value at USP is 100,000 ng/mg, and the target value after DSP is set to 100 ng/mg. More data of case study are given in the Supporting Information (Tables S1 and S2).

The lower and upper bounds of fermentation titre are 2 and 4 g/L, respectively, i.e., the uncertain parameter, *titre*, follows the triangular distribution $\mathrm{Tr}(2, 3, 4)$. Also, at each chromatography step, the chromatography resin yield can be at most 5% higher or lower than its standard value, which means that the yield deviation, cyd_s, follows the distribution, $\mathrm{Tr}(95\%, 100\%, 105\%)$. The unsatisfied demand penalty, γ, is set to £1000/g.

5. Results and Discussion

The proposed MILP models are implemented in GAMS 24.4[32] on a 64-bit Windows 7 based machine with 3.20 GHz six-core Intel Xeon processor W3670 and 12.0 GB RAM, using CPLEX as MILP solver. The optimality gap is set to 1%.

5.1. Optimisation and Monte Carlo Simulation

In this section, we solve the optimisation models SM and DM, and conduct stochastic analysis to examine the impact of variability on their solutions by implementing Monte Carlo (MC) simulation[33].

5.1.1. Optimal Solutions

To solve model SM, the confidence level of chance constraint feasibility in model SM was set to 90%, i.e., $A^t = A_s^y = 90\%$. Under the given 90% confidence level, $\Phi^{-1}(1 - A^t)$ in Eq. (29) is approximately equal to 2.45, and $\bar{\Phi}_s^{-1}\left(1 - A_s^y\right)$ in Eq. (34) is approximately equal to 97.24%, according to Eqs. (30) and (35), respectively.

The model statistics and computational performance of model SM in all four scenarios are presented in Table 5. Similar to the findings in our previous work[17-20], the optimal COG increases with the number of bioreactors. The computational time spent for solving the problems varies from 339 to 1613 seconds (s).

Table 5. Computational performance of model SM in all scenarios

	1USP:1DSP	2USP:1DSP	4USP:1DSP	6USP:1DSP
No of equations	26,940	27,612	28,284	28,284
No of continuous variables	11,172	11,396	11,620	11,620
No of discrete variables	28.661	28,662	28,663	28,663
COG (million £, m£)	43.0	49.1	57.7	66.2
CPU (s)	339	423	1613	1287

The optimal solutions obtained by model SM are compared with those of model DM, in which the mean values of uncertain parameters are used. The key decisions obtained by models SM and DM are given in Fig. 2. As the number of cycles per batch are related to the process operation and can be determined after the realisation of uncertain parameters, here we focus on the decisions related the process design which should be implemented in advance, including bioreactor volume, chromatography sequence, and diameter, bed height and number of chromatography columns. In each scenario, the gridded cylinder represents the bioreactor. The area of each gridded cylinder is proportional to the bioreactor volume, and the number in the gridded cylinder is the single bioreactor volume (in L) in the optimal solution. The filled cylinders represent the chromatography columns. Here, the number of the filled cylinders represents the number of columns at each chromatography step, while its width and height are proportional to the column diameter and bed height, respectively. The number in the top area is the column diameter (in cm), and that at the side is the column bed height (in cm). Comparing the optimal solutions of both models, to deal with the variability of the upstream titre and chromatography resin yields, model SM selects 30% larger bioreactor volumes in all scenarios. In the DSP, both models select only one column at each step in all scenarios. As to the chromatography sequence, CEX (R5), AEX (R7) and MM (R8) are selected in all scenarios, in which CEX (R5) is selected at the capture step in all scenarios by both models. The difference among the selected sequences comes from the switch between AEX (R7) and MM (R8) at the last two chromatography steps. In scenarios 2USP:1DSP and 6USP:1DSP, models SM and DM choose different sequences.

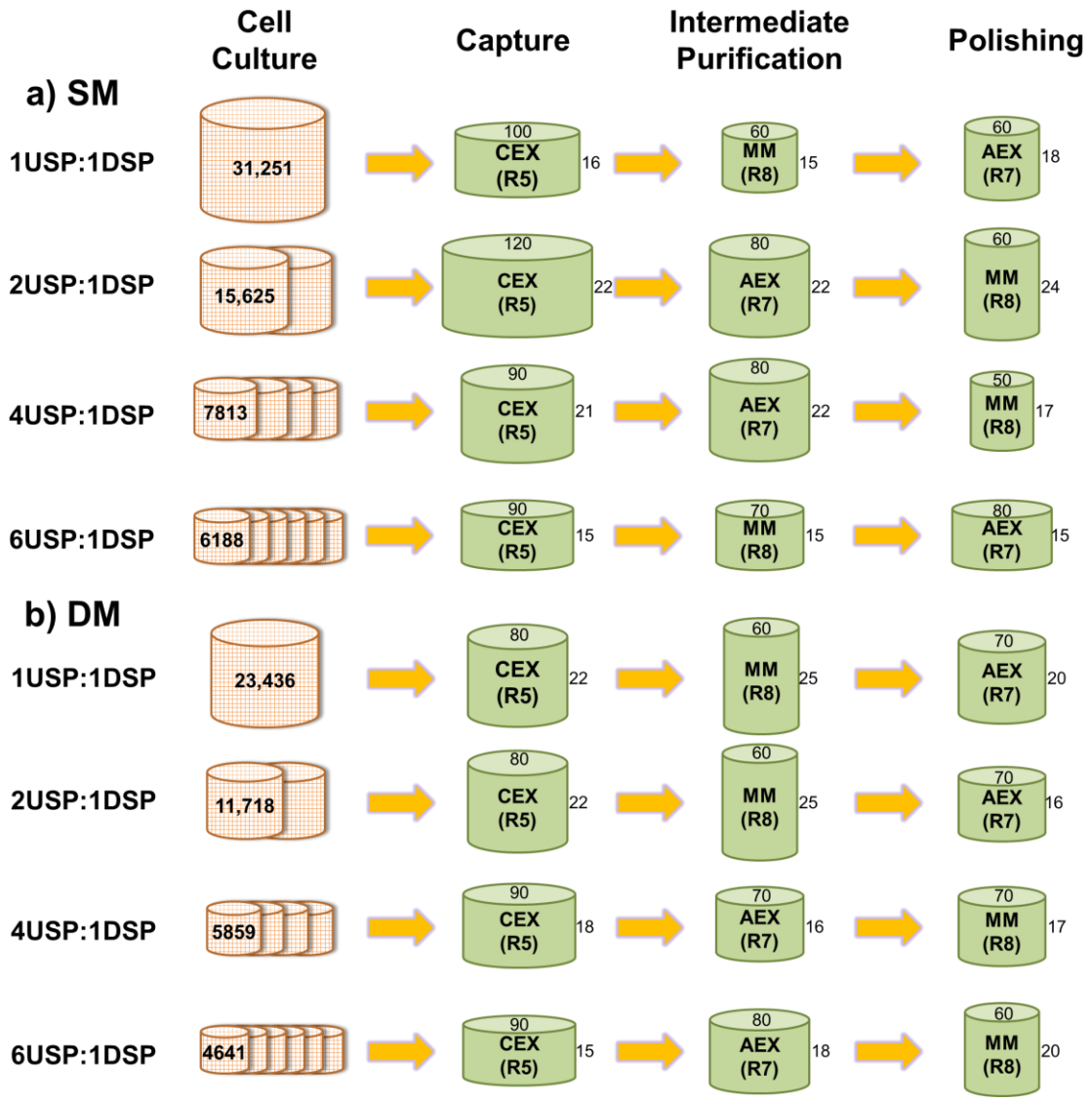

Fig. 2. Optimal solutions by models a) SM and b) DM in all scenarios (initial HCP level = 100,000 ng/mg)

5.1.2. MC Simulation

MC simulation analysis was implemented on the solutions obtained by both models SM and DM. After obtaining the optimal solutions of optimisation models DM and SM, an MC simulation analysis was conducted by solving a deterministic optimisation model with fixed design variables, including variables for bioreactor volume, BRV, chromatography sequence, U_{sr}, column volume, X_{si}, and number of columns, CN_{si}, while all other operational variables are free to be re-optimised, with different realisations of uncertain parameters, $titre$ and cyd_s. The performance measures in the MC analysis are the average values of COGs and unsatisfied demands obtained in all

simulation runs, which mimics the expected value of COG and unsatisfied demand. The procedure of MC simulation (as illustrated in Fig. 3) is described as follows:

1. Solve the proposed MILP model (DM or SM) to minimise COG, and then fix the obtained bioreactor volume, chromatography sequence, column volume and the number of columns;
2. Generate the realisations of uncertain titre and yields, both following triangular probability distributions;
3. Solve the modified MILP model (denoted as DM-USD), which considers unsatisfied demand by replacing Eqs. (7) and (23) in model DM by the following Eqs. (36) and (37), respectively:

$$USD \geq dem - AP \tag{36}$$

$$OBJ = AC + \gamma \cdot USD \tag{37}$$

where the unsatisfied demand, USD, is formulated and penalised in the objective function to be included in the annual cost, and γ is the unit penalty cost for unsatisfied demand.

4. Repeat the above two steps for *maxiter* times. In this problem, after 200 simulation runs, the average objective of all runs has converged to a stable status. Thus, *maxiter* is set to 200 here.

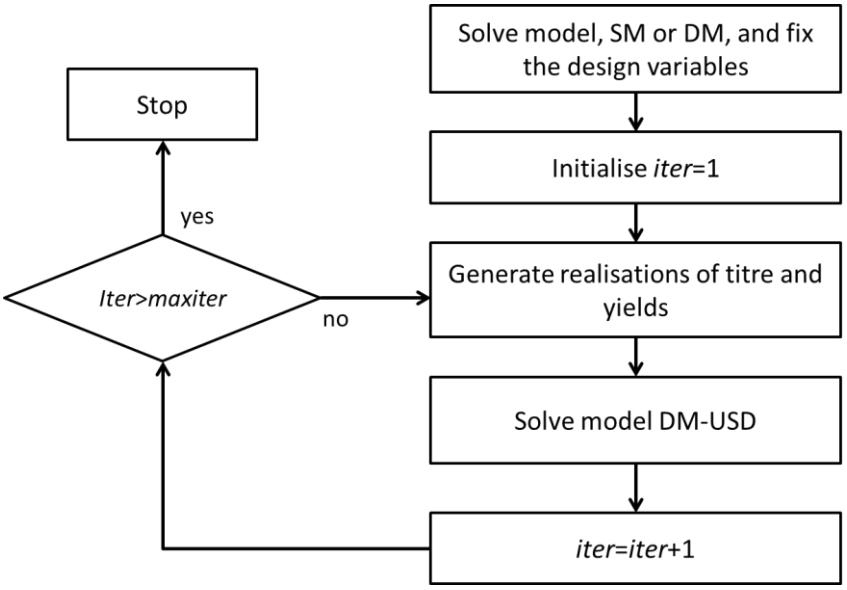

Fig. 3. Procedure of MC simulation

Fig. 4 shows the optimal COG by model SM is slightly higher than the average COG obtained in the MC simulation, because the chance constraints in model SM slightly overestimate the realisation of uncertain parameters. In contrast, the average COG values of the MC simulation become much higher than the optimal solutions by model DM, mainly due to the high penalty cost of unsatisfied demand. Comparing to model DM, the solutions of model SM obtain smaller average

COG values in the MC simulation, with 30-40% difference, and smaller variance, even though model SM generates much higher costs in the optimisation. Thus, model SM is much more robust than model DM, facing the uncertainties in titre and yields.

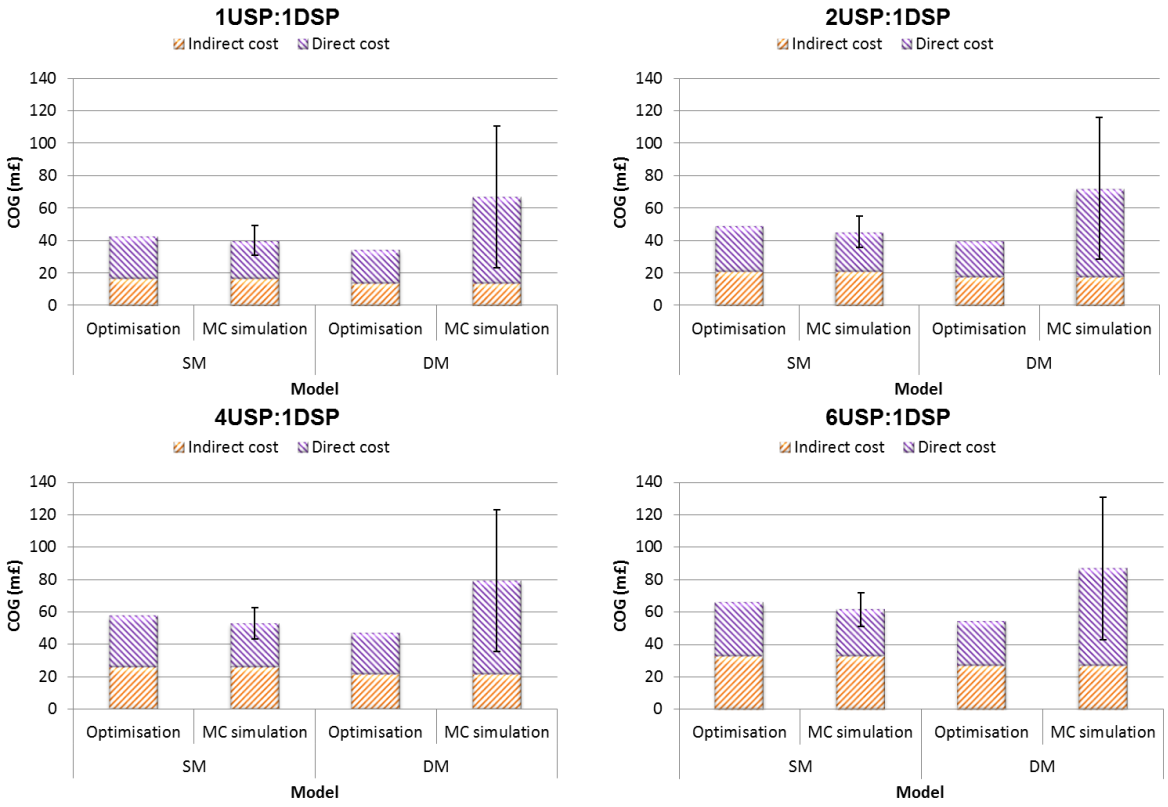

Fig. 4. COG of the optimisation and MC simulation of models SM and DM in all scenarios.

Focusing on the two extreme scenarios 1USP:1DSP and 6USP:1DSP, the comparison of unsatisfied demand in the MC simulation results between the solutions of models DM and SM are presented in Fig. 5. The MC simulation of model SM in scenario 1USP:1DSP achieves an annual unsatisfied demand of 1.5 kg, while model DM loses over 33.0 kg demand, nearly 7% of the annual demand, which is one order of magnitude higher than model SM. In scenario 6USP:1DSP, the average unsatisfied demand of MC simulation for model SM is 1.8 kg, which is also one order of magnitude lower than that of model DM, 33.0 kg. Thus, it can be observed that the optimal USP and DSP designs of model DM cannot cope with the uncertainties very well, and lose more sales than the optimal design of model SM, which leads to much higher penalty cost and COG.

Meanwhile, the designs of model SM have much better performance and only lose less than 0.5% of the annual demand.

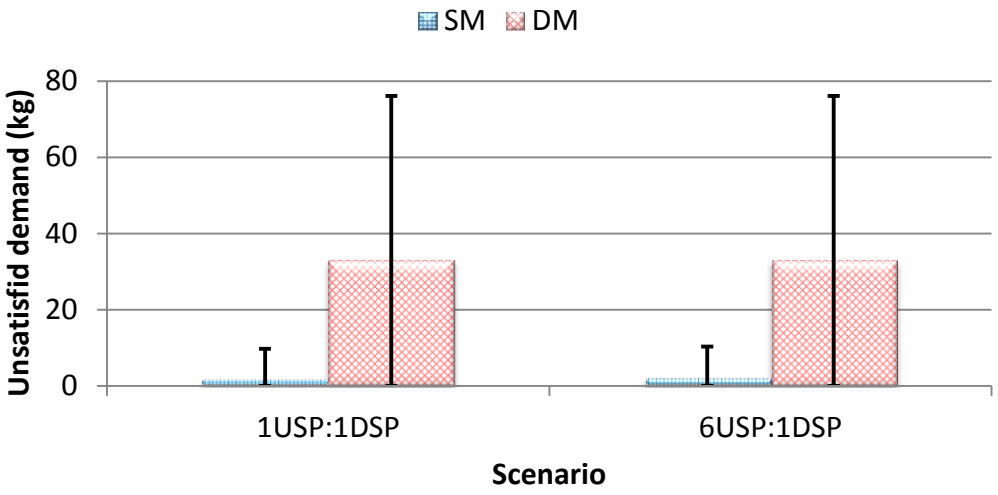

Fig. 5. Unsatisfied demand of the MC simulation on the optimal solutions of models SM and DM in scenarios 1USP:1DSP and 6USP:1DSP

In the MC simulation, the number of cycles per batch is optimised in each run, with fixed design variables and given random generations of titre and yields. Thus, we examine its value in the MC simulation, as shown in Fig. 6. In scenario 1USP:1DSP, for the solutions of model SM, the latter chromatography step requires more cycles, while for model DM, an average of 4-5 cycles are needed for each step. In scenario 6USP:1DSP, 1 or 2 cycles are enough at each step for all the MC simulation runs, while the solutions of SM require more cycles than those of model DM.

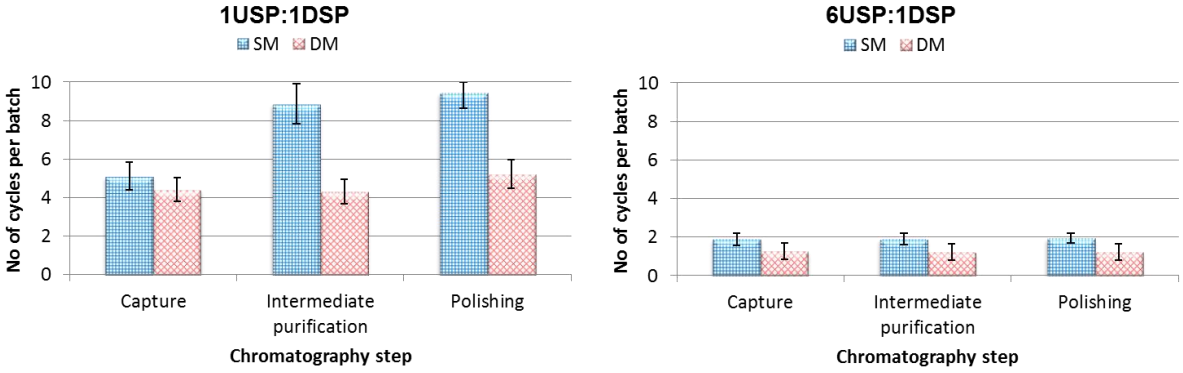

Fig. 6. Number of cycles per batch of the MC simulation on the optimal solutions of models SM and DM in scenarios 1USP:1DSP and 6USP:1DSP

The number of batches is also significantly affected by the uncertainties of titre and yield. When both titre and yield are higher, fewer batches are required to satisfy the annual demand, while if both titre and yield become lower, more batches need to complete with unsatisfied demand occurred. The mean and variability of number of batches in the MC simulation are presented in Fig. 6. In both scenarios, the solution of model SM requires fewer average batches, 16.4 vs 19.2 in scenario 1USP:1DSP and 78.4 vs 96.4 in scenario 6USP:1DSP. In scenario 1USP:1DSP, the average number of batches of model DM is very close to the maximum value, which is 20, while the solution of SM requires fewer batches to satisfy all demands due to the larger bioreactor volumes. In scenario 6USP:1DSP, a maximum of 120 batches are available. However, because of the higher DSP time than the tight DSP window in scenario 6USP:1DSP, nearly 100 batches are finished by model DM, while model SM meets most of the demand using only about 80 batches in average. Note that the variability of the solution of model DM is much less than that of model SM. The reason is that the solution of model DM is driven to meet the demand by completing all possible batches in many MC simulation runs, and has less flexibility to change the completed number of batches. Meanwhile, the number of batches required by the solution of model SM is affected more significantly by the realised values of uncertain parameters.

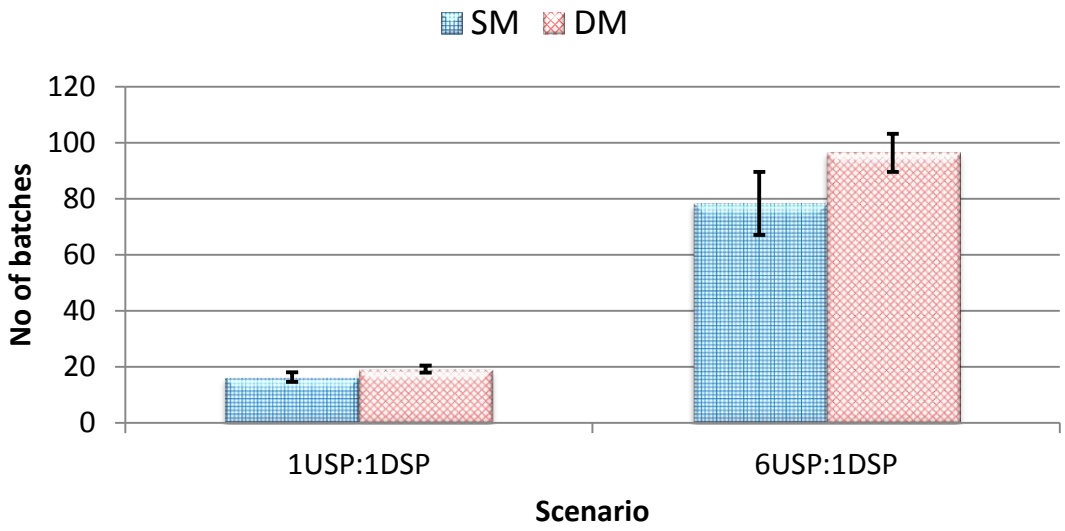

Fig. 7. Number of batches of the MC simulation on the optimal solutions of models SM and DM in scenarios 1USP:1DSP and 6USP:1DSP

Overall, the developed CCP based model SM is able to cope with the uncertainties of the parameters, upstream titre and downstream resin yields in this problem, achieving lower average COG and unsatisfied demand than model DM.

5.1.3. Sensitivity Analysis of Confidence Levels

The confidence levels used in the chance constraints indicate the probability of the obtained solutions being feasible, and have a significant impact on the optimal solutions of the CCP model. The example of 1USP:1DSP scenario was taken to investigate the effect of confidence level on the performance of the optimal solutions. Without loss of generality, assuming that all chance constraints use the same confidence level, we test five different confidence levels, i.e. 80%, 85%, 90%, 95% and 99%. Table 6 shows the optimal solutions of model SM with the above five confidence levels. A higher confidence level, preferred by a risk-averse decision maker, allows the chance constraints to remain hold with higher probability. Thus, as shown in Table 6, in order to cope with lower titres, the optimal solution of model SM with higher confidence levels select larger bioreactor volumes, and incur higher COG and indirect costs. Note that the chromatography sequence is affected by the choice of confidence level. Higher conference levels (≥90%) result in different chromatography sequences from lower ones (<90%).

Table 6. Optimal solutions of model SM in scenario 1USP:1DSP with different confidence levels

	Confidence level				
	80%	**85%**	**90%**	**95%**	**99%**
COG (m£)	39.8	41.2	43.0	45.6	49.7
Indirect costs (k£)	15.8	16.1	16.7	17.5	18.8
Bioreactor volume (L)	28,237	29,557	31,251	33,694	37,452
Resin	R5/R7/R8[a]	R5/R7/R8[a]	R5/R8/R7[a]	R5/R8/R7[a]	R5/R8/R7[a]
Column diameter (cm)	160/50/50[a]	120/50/50[a]	100/60/60[a]	100/60/60[a]	100/50/50[a]
Column bed height (cm)	24/24/15[a]	22/25/17[a]	16/15/18[a]	21/21/15[a]	16/19/20[a]
No. of columns	1/1/1[a]	1/1/1[a]	1/1/1[a]	1/1/1[a]	1/1/1[a]

[a] Capture/intermediate purification/polishing step

The MC simulation results with different confidence levels are shown in Fig. 8. As expected, the higher the confidence level is, the lower average unsatisfied demand achieved from MC simulation is. The average unsatisfied demand reduces from 5.5 kg to 0.05 kg, with increasing confidence levels from 80% to 99%. Meanwhile, comparing the average COG, a higher confidence level

requires larger bioreactors and indirect costs, while a lower confidence level incurs higher penalty cost due to higher unsatisfied demand. In addition, with increasing higher confidence level, the variance of COG in MC simulation decreases. It can be seen from Fig. 8 that the confidence level of 90% can achieve a balance between the indirect costs and unsatisfied demand penalty cost, obtaining the lowest average COG in MC simulation, and thus is recommended for this problem.

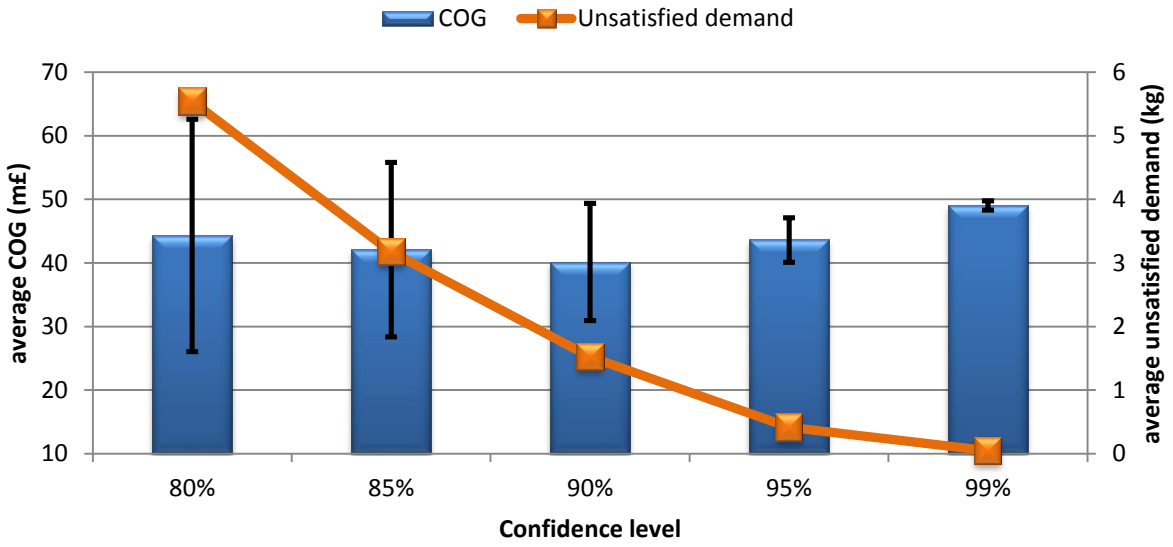

Fig. 8. COG and unsatisfied demand of MC simulation on SM solutions in scenario 1USP:1DSP

5.2. Scenario Analysis

In this section, two specific scenarios are investigated to further discuss the proposed model SM. Firstly, the model is compared to an approach with a rule-based pre-determined bioreactor volume and an optimisation model with only DSP optimisation, to demonstrate the benefit of integrated optimisation of both USP and DSP design. Secondly, the effect of initial HCP level is studied by examining a higher initial HCP level.

5.2.1. Pre-determined Bioreactor Volume

One of the novelties of this work is the integrated optimisation of both USP and DSP design, including the bioreactor size at the USP and chromatography decisions at the DSP, which is missing in our previous work[17-20]. In this section, we investigate the advantage of the simultaneous

integrated optimisation. Firstly, we introduce a rule-based method to estimate the single bioreactor volume[12,17,18], using the formulation as follows:

$$BRV = \frac{dem}{\alpha \cdot titre^{mo} \cdot maxbn \cdot \sigma \cdot \prod_{s \in CS} \min_r cy_{sr} \cdot \prod_{s \notin CS} ncy_s} \tag{38}$$

where dem is the annual demand; α is the bioreactor working volume ratio; $titre^{mo}$ is the mode of the titre triangular distribution, which is also its mean value for the case study; $maxbn$ is the maximum number of batches allowed; σ is the batch success rate; $cy_{sr}$is the standard value of the yield of resin r at chromatography step $s \in CS$, while ncy_s is the yield of non-chromatography step $s \notin CS$. Thus, the multiplication of the two product function is an estimation of the overall manufacturing yield. The above pre-determined bioreactor volume and its optimal value by model SM for our case study are given in Table 7. The estimated bioreactor volume is 20% less than the optimal volume in the first three scenarios, while in the last scenario 6USP:1DSP, it is over 30% less.

Table 7. Pre-determined and optimal value of single bioreactor volume (L) in all scenarios

	1USP:1DSP	**2USP:1DSP**	**4USP:1DSP**	**6USP:1DSP**
Pre-determined bioreactor volume by Eq. (38)	25,016	12,508	6254	4169
Optimal bioreactor volume by model SM	31,251	15,625	7813	6188

Then, the proposed integrated optimisation (model SM) is compared with the DSP optimisation, in which the DSP chromatography strategies are optimised by model SM with variable *BRV* fixed to the pre-determined value obtained by Eq. (38). In order to allow the unsatisfied demand with the fixed bioreactor volume, Eqs. (36) and (37) are used to replace Eqs. (7) and (23) in model SM. Here, we still take the two extreme scenarios, 1USP:1DSP and 6USP:1DSP, as examples to compare the optimal solutions of model SM for DSP optimisation and the integrated optimisation of both USP and DSP. The optimal chromatography decisions obtained by DSP optimisation are given in Fig. 9. It can be observed that compared with the optimal solution of integrated optimisation as decision values in Fig. 2a, different chromatography sequences are selected in both scenarios, resulting in different column diameters and bed height as well.

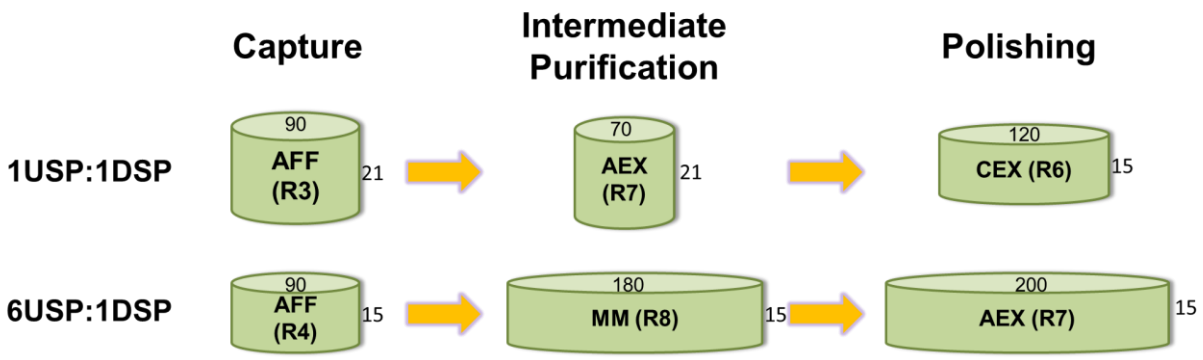

Fig. 9. Optimal solution by model SM with pre-determined bioreactor volume in scenarios 1USP:1DSP and 6USP:1DSP

The MC simulation is conducted on the solutions of both the integrated optimisation and DSP optimisation with bioreactor volume pre-determined, as descried in section 5.1.2. The comparison on the running average COG and unsatisfied demand obtained in the MC simulation is shown in Fig. 10. In scenario 1USP:1DSP, the average unsatisfied demand (9.1 kg) of MC simulation of the solution of DSP optimisation is much higher than that of integrated optimisation (1.5 kg), which increases the average COG by 17% from £40.2m to £47.1m. While in scenario 6USP:1DSP, the advantage of the integrated optimisation is much more significant, as the average unsatisfied demand of MC simulation of DSP optimisation is (30.7 kg) is one order of magnitude higher than that of the integrated optimisation (1.8 kg). As to the obtained average COG in MC simulation, DSP optimisation obtains an average COG of £93.4m, which is over 50% higher than that of the integrated optimisation (£61.5m).

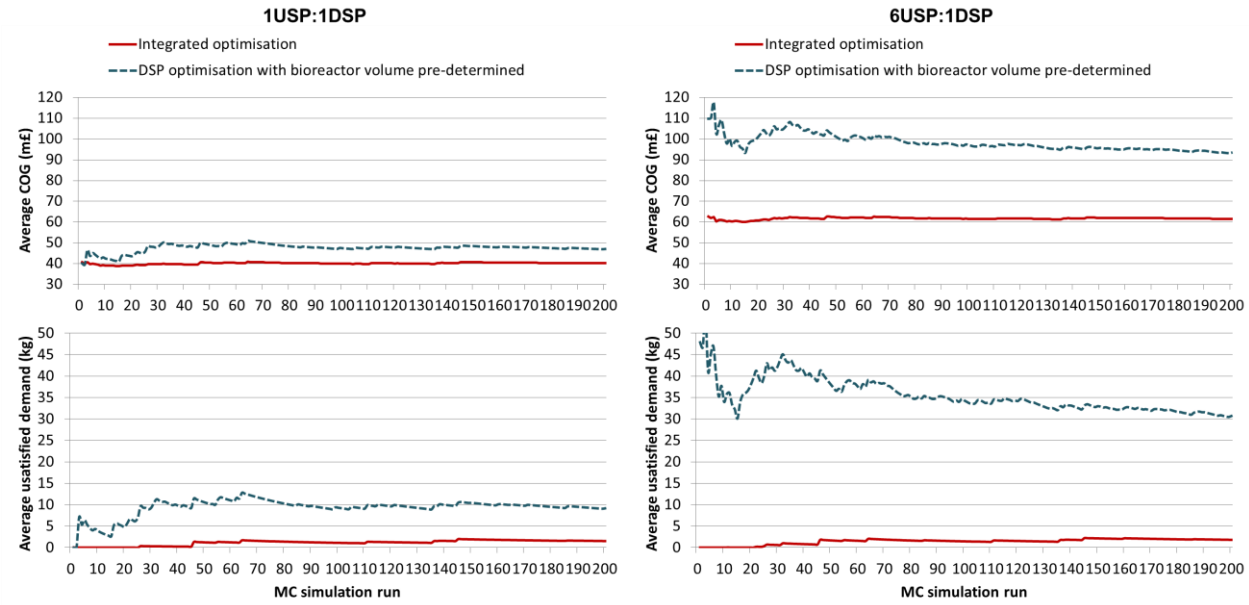

Fig. 10. Convergence of the average COG and unsatisfied demand of MC simulation on the optimal solutions of USP/DSP integrated optimisation and DSP optimisation with bioreactor volume pre-determined by model SM in scenarios 1USP:1DSP and 6USP:1DSP

The above comparison demonstrates a clear benefit of the integrated optimisation, which synchronises the demand, bioreactor volume and DSP time, and is able to find an optimal solution to meet the demand using larger bioreactor volume and fewer batches, while the rule-based method for bioreactor volume calculation (Eq. 38) underestimates the bioreactor volume and leads to much higher unsatisfied demand, especially when there are more bioreactors and tighter DSP window, such as the scenario 6USP:1DSP. Compared to the DSP optimisation with the pre-determined bioreactor volumes, the proposed integrated optimisation of both USP and DSP is able to achieve much lower unsatisfied demand and COG. Thus, it is important and necessary to consider the optimisation of USP bioreactor volumes simultaneously together with the DSP chromatography strategies.

Note that if the worst case scenario is considered by replacing the mean by the lower bound of uncertain parameters either in the pre-calculation of bioreactor sizes or in the mode DM, the unsatisfied demand of both DSP optimisation and model DM in the MC simulation will become lower than that of model SM, but the overall COG is still higher than that of model SM. The conclusions on the benefits of model SM and the integrated optimisation are still valid.

5.2.2. Higher Initial HCP Level

Lastly, we investigate the effect of the initial HCP level on the optimal solution of SM. We let the initial HCP level increase by 15 times to 1,500,000 ng/mg. The optimal solution obtained by model SM with this high initial HCP level is given in Fig. 11. Compared with Fig. 2a whose initial HCP level is 100,000 ng/mg, a different chromatography sequence is selected, which includes AFF (R3) for capture, CEX (R6) for intermediate purification, and AEX (R7) for polishing for all scenarios. R3 selected at the capture step is much more expensive than R5 selected with lower initial HCP level, and has a higher HCPs reduction ability with a HCP LRV of 3, higher than 2 for R5. Although, at the intermediate purification step, R6 is cheaper and has less HCP reduction ability than, R8, the total HCP LRV of the optimal sequence in Fig. 11 is 4.3, higher than the sequences in Fig. 2a (3.1 for the sequence of R5/R7/R8 or 3.5 for the sequence of R5/R8/R7). So it is clear that when the initial HCP level is increased, in order to enforce that the same target HCP level is met, the resins with higher HCP reduction abilities are selected, even though they may cost more. In addition, as the selected resins in Fig. 11 have higher yields than those in Fig. 2a, which leads to less protein loss during processes, the optimal single bioreactor volume becomes smaller for each scenario in Fig. 11, but the same annual demand is achieved.

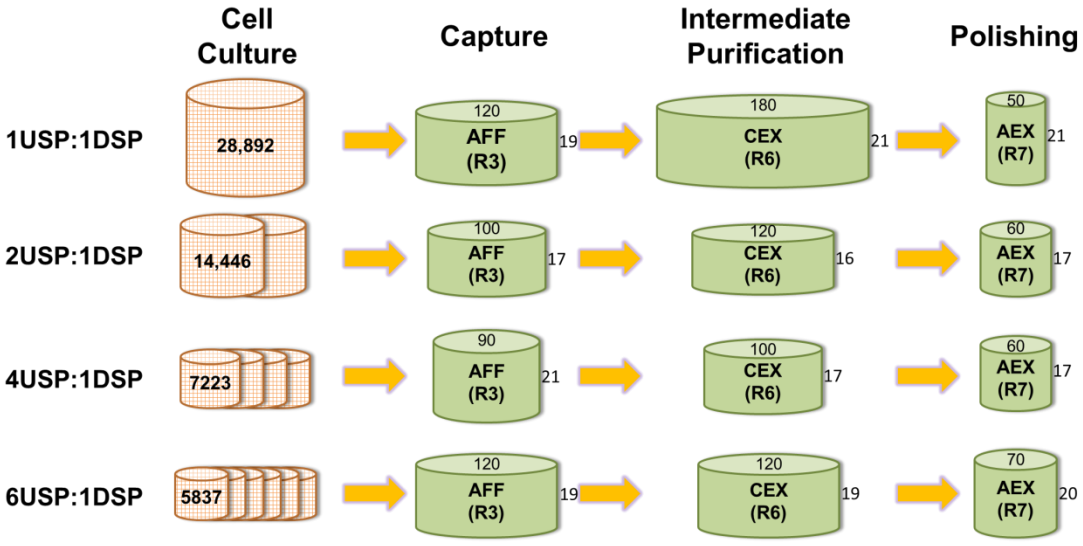

Fig. 11. Optimal solution by model SM in all scenarios (initial HCP level = 1,500,000 ng/mg)

Due to the more expensive price of the selected resins in Fig. 11, especially the resin R3, the direct cost for the optimal process with high initial HCP level is higher than that with the low initial HCP

level. Meanwhile, as smaller bioreactor volumes are selected with the high initial HCP level in Fig. 11, lower indirect cost is incurred as shown in Fig. 12. According to the above discussions, the initial HCP level has a significant impact on the optimal bioreactor sizing and chromatography sequencing decisions.

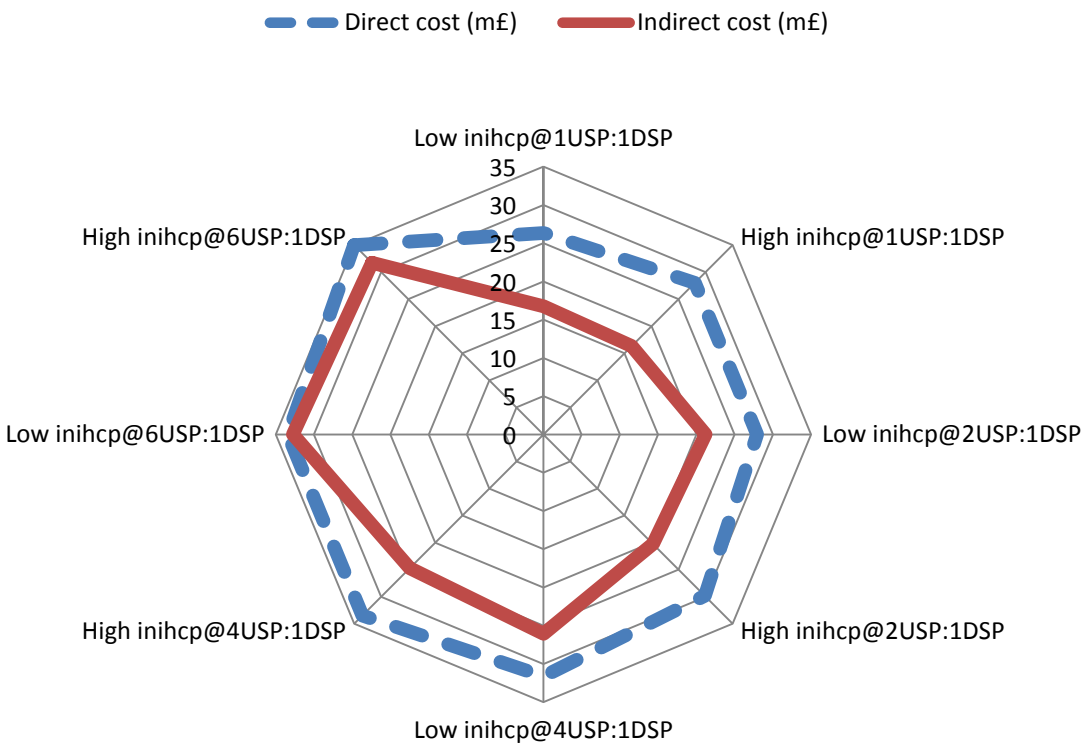

Fig. 12. COG of the optimisation of model SM with low (100,000 ng/mg) and high (1,500,000 ng/mg) initial HCP levels, *inihcp*, in all scenarios.

6. Concluding Remarks

This paper addressed the integrated optimisation of the both upstream and downstream processing of mAb products, considering bioreactor sizing, chromatography sequencing and column sizing strategies, simultaneously. The purity requirement of the final product has been taken into account as well, to make sure that the final HCP level is below the target. Extended from our previous work[17-20], a stochastic CCP model has been developed, based on a deterministic MILP model, to minimise the total COG considering the uncertainties of upstream titre and chromatography resin yield. An industrially-relevant example has been investigated for cases with different USP:DSP

configurations. The computational results have shown that the stochastic CCP model results in larger bioreactor volume and higher indirect cost, but is able to deal with the variability of uncertain parameters in a much better manner than the deterministic model, validated by MC simulation approach. Also, sensitivity analysis on the confidence level shows its effects on the expected cost and unsatisfied demand in the MC simulation. Through scenario analysis, the advantage of the integrated optimisation of both USP and DSP decisions is demonstrated through MC simulation, achieving lower average COG and unsatisfied demand, compared to the DSP optimisation with pre-determined bioreactor volume, which is underestimated by a rule-based estimation method. Finally, the initial HCP level is demonstrated to have significant impacts on the optimal bioreactor sizing and chromatography sequencing decisions. This work can be extended by modelling each batch separately to account for the random realisation of the uncertain parameters in each batch, while the proposed stochastic CCP model ignores the variations among batches in the real practice.

Supporting Information

More data of the case study and the literature model[19] are provided in the Supporting Information. This information is available free of charge via the Internet at http://pubs.acs.org/

Acknowledgements

Funding from the UK Engineering & Physical Sciences Research Council (EPSRC) for the EPSRC Centre for Innovative Manufacturing in Emergent Macromolecular Therapies (EP/I033270/1), and to LGP (EP/M027856/1) is gratefully acknowledged. Financial support from the consortium of industrial and governmental users is also acknowledged. The authors thank Dr Ana S. Simaria for useful discussions.

Nomenclature

Indices

i	column size
m	bioreactor volume
n	digit of the binary representation
r	resin
s	downstream step

Sets

CS	set of chromatography steps, = capture, intermediate purification, polishing
R_s	set of resins suitable to chromatography step *s*

Parameters

a, b, c	utilities cost coefficients
A^t	confidence level in chance constraint for initial mass
A_s^y	confidence level in chance constraint for mass after chromatography step *s*
brf	scale-up factor of bioreactor cost
brn	number of bioreactors
cc_{si}	column cost of size *i* at chromatography step *s*, £
cf	scale-up factor of column cost
cy_{sr}	product yield of resin *r* at chromatography step *s*
cyd_s^{lo}	lower limit of triangular distribution of yield deviation at chromatography step *s*
cyd_s^{mo}	mode of triangular distribution of yield deviation at chromatography step *s*
cyd_s^{up}	upper limit of triangular distribution of yield deviation at chromatography step *s*
$dbrc_m$	bioreactor cost at discrete point *m*, £
$dbrv_m$	bioreactor volume at discrete point *m*, L
dem	annual demand, g
dm_{si}	diameter of column size *i* at chromatography step *s*, L
gef	general equipment factor
gu	general utility unit cost, £/L
$inihcp$	initial HCP level, ng/mg
$lang$	Lang factor
lrv	HCP log reduction value
$maxbn$	maximum number of batches
$maxbrv$	maximum bioreactor volume
$mepc$	media price, £/L
ncy_s	product yield of non-chromatography step *s*
$oe\lambda$	other equipment cost ratio to the bioreactor cost
q	maximum digit number in the binary representation of number of batches, $\lceil \log_2 maxbn \rceil$
$refbrc$	reference cost of a bioreactor, £
$refbrv$	reference volume of a bioreactor, L
$refcc$	reference cost of a chromatography column, £
$refdm$	reference diameter of a chromatography column, cm
$tarhcp$	target HCP level, ng/mg
$titre$	upstream product titre, g/L
$titre^{lo}$	lower limit of triangular distribution of upstream product titre, g/L
$titre^{mo}$	mode of triangular distribution of upstream product titre, g/L
$titre^{up}$	upper limit of triangular distribution of upstream product titre, g/L
α	bioreactor working volume ratio
γ	penalty of unsatisfied demand, £/g
θ	media overfill allowance

σ batch success rate
Φ triangular cumulative distribution function of uncertain titre
$\bar{\Phi}_s$ triangular cumulative distribution function of uncertain resin yield deviation

Continuous Variables
ABV annual buffer volume, L
AP annual product output, g
BRC bioreactor cost, £
BRV single bioreactor volume, L
COG annual cost of goods, £
FCI fixed capital investment, £
GUC general utility cost, £
M_0 initial product mass entering downstream processes per batch, g
M_s initial product mass per batch after step *s*, g
MEC media cost, £
OBJ objective
PV_0 initial product volume entering downstream processes per batch, L
UC utilities cost, £
USD unsatisfied demand, g
Λ_m SOS2 variable for piecewise linearization of bioreactor cost

Binary Variables
U_{sr} 1 if resin *r* is selected at chromatography step *s*; 0 otherwise
X_{si} 1 if column size *i* is selected at chromatography step *s*; 0 otherwise
Z_n 1 if the *n*th digit of the binary representation of variable BN is equal to 1; 0 otherwise

Integer Variables
BN number of completed batches
CN_{si} number of columns of size *i* at chromatography step *s*

Auxiliary Variables
$\overline{UM}_{s-1,r}$ $\equiv U_{sr} \cdot M_{s-1}$
$\overline{ZBRV}_n$ $\equiv Z_n \cdot BRV$
$\overline{ZM}_{sn}$ $\equiv Z_n \cdot M_s$

Literature Cited

(1) Ecker, D. M., Jones, S. D., & Levine, H. L. (2015). The therapeutic monoclonal antibody market, *MAbs*, 7, 9-14.
(2) Low, D., O'Leary, R., & Pujar, N. S. (2007). Future of antibody purification. *Journal Chromatography B*, 848, 48-63.
(3) Lakhdar, K., Zhou, Y. H., Savery, J., Titchener-Hooker, N. J., & Papageorgiou, L. G. (2005). Medium term planning of biopharmaceutical manufacture using mathematical programming. *Biotechnology Progress*, 21, 1478–1489.

(4) Lakhdar, K., Farid, S. S., Titchener-Hooker, N. J., & Papageorgiou, L. G. (2006). Medium term planning of biopharmaceutical manufacture with uncertain fermentation titers. *Biotechnology Progress*, 22, 1630−1636.
(5) Lakhdar, K., Savery, J., Papageorgiou, L. G., & Farid, S.S. (2007). Multiobjective long term planning of biopharmaceutical manufacturing facilities. *Biotechnology Progress*, 23(6), 1383-1393.
(6) Kabra, S., Shaik, M. A., & Rathore, A. R. (2013). Multi-period scheduling of a multi-stage multi-product bio-pharmaceutical process. *Computers & Chemical Engineering*, 57, 95−103.
(7) Liu, S., Yahia, A., & Papageorgiou, L. G. (2014). Optimal production and maintenance planning of biopharmaceutical manufacturing under performance decay. *Industrial & Engineering Chemistry Research*, 53, 17075-17091.
(8) Siganporia, C. C., Ghosh, S., Daszkowski, T., Papageorgiou, L. G., & Farid, S. S. (2014). Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. *Biotechnology Progress*, 30, 594-606.
(9) Vasquez-Alvarez, E., Lienqueo, M. E., & Pinto, J. M. (2001). Optimal synthesis of protein purification processes. *Biotechnology Progress*, 17, 685-696.
(10) Simeonidis, E., Pinto, J. M., Lienqueo, M. E., Tsoka, S., & Papageorgiou, L. G. (2005). MINLP models for the synthesis of optimal peptide tags and downstream protein processing. *Biotechnology Progress*, 21, 875-884.
(11) Natali, J. M., Pinto, J. M., & Papageorgiou, L. G. (2009). Efficient MILP formulations for the simultaneous optimal peptide tag design and downstream processing synthesis. *AIChE Journal*, 55, 2303-2317.
(12) Polykarpou, E. M., Dalby, P. A., & Papageorgiou, L. G. (2011). Optimal synthesis of chromatographic trains for downstream protein processing. *Biotechnology Progress*, 27, 1653-1660.
(13) Simaria, A. S., Turner, R., & Farid, S. S. (2012). A multi-level meta-heuristic algorithm for the optimisation of antibody purification processes. *Biochemical Engineering Journal*, 69, 144-154.
(14) Allmendinger, R., Simaria, A. S., & Farid, S. S. (2012). Efficient discovery of chromatography equipment sizing strategies for antibody purification processes using evolutionary computing. Parallel Problem Solving in Nature - PPSN XII. *Lecture Notes in Computer Science*, vol. 7492, pp. 468-477.
(15) Allmendinger, R., Simaria, A. S., & Farid, S. S. (2014). Closed-loop optimization of chromatography column sizing strategies in biopharmaceutical manufacture. *Journal of Chemical Technology & Biotechnology*, 89, 1481–1490.
(16) Allmendinger, R., Simaria, A.S., Farid, S.S. (2014). Multiobjective evolutionary optimization in antibody purification process design. *Biochemical Engineering Journal*, 91, 250–264.
(17) Liu, S., Simaria, A. S., Farid, S. S., & Papageorgiou, L. G. (2013). Mixed integer optimisation of antibody purification processes. In: Kraslawski, A., & Turunen, I. (eds.) Proceedings of the 23rd European Symposium on Computer Aided Process Engineering, *Computer Aided Chemical Engineering*, 32, 157-162. Amsterdam: Elsevier.
(18) Liu, S., Simaria, A. S., Farid, S. S., & Papageorgiou, L. G. (2013). Designing cost-effective biopharmaceutical facilities using mixed-integer optimization. *Biotechnology Progress*, 29, 1472–1483.

(19) Liu, S., Simaria, A. S., Farid, S. S., & Papageorgiou, L. G. (2014). Optimising chromatography strategies of antibody purification processes by mixed integer fractional programming techniques. *Computers & Chemical Engineering*, 68, 151-164.
(20) Liu, S., Simaria, A. S., Farid, S. S., & Papageorgiou, L. G. (2015). Mathematical programming approaches for downstream processing optimisation of biopharmaceuticals. *Chemical Engineering Research & Design*, 94, 18-31.
(21) Brunet, R., Guillén-Gosálbez, G., Pérez-Correa, J. R., Caballero, J. A., & Jiménez, L. (2012). Hybrid simulation-optimization based approach for the optimal design of single-product biotechnological processes. *Computers & Chemical Engineering*, 37, 125-135.
(22) Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science*, 6, 73-79.
(23) Petvok, S. B., & Maranas, C. D. (1997). Multiperiod planning and scheduling of multiproduct batch plants under demand uncertainty. *Industrial & Engineering Chemistry Research*, 36, 4864-4881.
(24) Gupta, A., Maranas C. D., & McDonald, C. M. (2000). Midterm supply chain planning under demand uncertainty: customer demand satisfaction and inventory management. *Computers & Chemical Engineering*, 24, 2613-2621.
(25) Liu, L., Huang, G.H., Liu, Y., Fuller, G.A., & Zeng, G.M. (2003). A fuzzy-stochastic robust programming model for regional air quality management under uncertainty. *Engineering Optimization*, 35, 177-199.
(26) Li, P., Arellano-Garcia, H., & Wozny, G. (2008). Chance constrained programming approach to process optimization under uncertainty. *Computers & Chemical Engineering*, 32, 25-45.
(27) Mitra, K. (2009). Multiobjective optimization of an industrial grinding operation under uncertainty. *Chemical Engineering Science*, 63, 5043-5056.
(28) Xu, M., & Zhuan, X. (2013). Optimal planning for wind power capacity in an electric power system. *Renewable Energy*, 53, 280-286.
(29) Levy, N. E., Valente, K. N., Choe, L. H., Lee, K. H., & Lenhoff, A. M. (2014). Identification and characterization of host cell protein product-associated impurities in monoclonal antibody bioprocessing. *Biotechnology & Bioengineering*, 111, 904–912.
(30) Stonier, A., Pain, D., Westlake, A., Hutchinson, N., Thornhill, N. F., & Farid, S. S. (2013). Integration of stochastic simulation with multivariate analysis: Short-term facility fit prediction. *Biotechnology Progress*, 29, 368-377.
(31) Farid, S. S., Washbrook, J., & Titchener-Hooker, N. J. (2005). Decision-support tool for assessing biomanufacturing strategies under uncertainty: stainless steel versus disposable equipment for clinical trial material preparation. *Biotechnology Progress*, 21, 486-497.
(32) Brooke, A., Kendrick, D., Meeraus, A., & Raman, R. (2014). *GAMS – A User's Guide*. Washington, D.C.: GAMS Development Corporation.
(33) Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of Monte Carlo Methods*. New Jersy: John Wiley & Sons.