An efficient digital twin based on machine learning SVD autoencoder and generalised latent assimilation for nuclear reactor physics

Helin Gong[a,*,1], Sibo Cheng[b,1], Zhang Chen[c], Qing Li[c], César Quilodrán-Casas[b], Dunhui Xiao[d] and Rossella Arcucci[e]

[a]ParisTech Elite Institute of Technology, Shanghai Jiao Tong University, Shanghai, 200240, China
[b]Data Science Institute, Department of computing, Imperial College London, London, SW7 2AZ, UK
[c]Science and Technology on Reactor System Design Technology Laboratory, Nuclear Power Institute of China, Chengdu, 610041, China
[d]Faculty of Sciences and Engineering, Swansea University, Swansea, SA1 8EN, UK
[e]Department of Earth Science & Engineering, Imperial College London, London, SW7 2AZ, UK

ARTICLE INFO

Keywords:
Operational digital twins
Machine learning
Latent Assimilation
SVD-autoencoder
Nuclear reactor physics

ABSTRACT

This paper proposes an approach that combines reduced-order models with machine learning in order to create an digital twin to predict the power distribution over the core during the operation stage. The operational digital twin is designed to solve forward problems given input operation parameters, as well as to solve inverse problems given some observations of the power field. The forward model is non-intrusive and realised using SVD autoencoder reduced order model with the combination of machine learning methods, namely, k-nearest-neighbors and decision trees to build the input-output map. For model parameter estimation, the inverse model is based on a generalised latent assimilation method. The proposed approach is able to make use of the non intrusive reduced order model and the online measurements of the power field. The effectiveness in the sense of accuracy and real-time solver of the digital twin is illustrated through a real engineering problem in nuclear reactor physics - reactor core simulation in the life cycle of HPR1000 affected by input parameters, i.e., control rod inserting step, burnup, power level and inlet temperature of the coolant, which shows potential applications for on-line monitoring purpose.

1. Introduction

The concept and model of digital twins was introduced by Grieves [37]. Since then, digital twins are permeating across nearly all academic, engineering or industrial disciplines, interested readers are referred to the reviews work of [38, 44]. Since the application is quite broad and varied, the definitions, enabling technologies, challenges and opportunities of digital twins are quite different, see [21, 66, 71] for a comprehensive review.

Related works on digital twins in nuclear engineering are recent, for instance, [51, 32, 65]. For example. The work in [32, 31] entrusts digital twins a specific role when dealing with the broader concepts of secure embedded intelligence and integrated state awareness. In [53], the authors mention digital twins focusing on advancing the concept of a nuclear power plant as a Human-Cyber-Physical System. The review article [51] focuses on uncertainty quantification (UQ) and software risk analysis of machine learning (ML) generated digital twin for the nearly autonomous management and control of nuclear power systems for prognostics and diagnostics supporting purposes. A more detailed discussion on digital twin concepts with uncertainty for nuclear power applications can be found in [47]. Additionally, the work [65] focuses on the development of technologies that enable digital-twinning efforts in nuclear proliferation detection. The technical reports [59, 58] released by Argonne National Lab (ANL) describe the development of a digital twin predictive model for Pressure Water Reactor (PWR) components, in a physics-infused and AI/ML-based framework. Varé and Morilhat [73] developed digital twins for models fed by operational online data and to benefit from data analytics to support long term operation of critical nuclear equipment. The work [43] aims to develop flexible modeling

*Corresponding author
gonghelin@sjtu.edu.cn (H. Gong)
ORCID(s): 0000-0002-4094-6795 (H. Gong)
[1]Helin Gong and Sibo Cheng are co-first authors

software package to simulate dynamic processes occurring in Nuclear Power Plant (NPP) power units with VVER-1000 reactors, meanwhile the authors insist that the flexible software package is the kernel of the digital twin of an NPP power unit.

Digital twin is defined as a virtual representation of a physical asset (e.g. a reactor core) enabled through data and simulators amenable to better understanding the behavior of nuclear reactor systems. This concept can be used for real-time prediction, optimization, monitoring, controlling, and improved decision making [64]. In this regard, the objective of a digital twin in nuclear reactor domain can generally be taken as gains in economical efficiency and safety goals in nuclear power applications, particularly in reactor operations. The early application in reactor physics [36] proposes a prototype of reactor physics operational digital twin (RPODT) to predict neutron flux and power distributions in nuclear reactor cores for on-line monitoring purpose. The effectiveness of the proposed digital twin was illustrated through a real engineering problem based on HPR1000 reactor core. In general, there are three main problems associated with building a RPODT:

(i) building a forward solver to efficiently compute approximations to any high-dimensional physical field $u(\boldsymbol{\mu})$ for any given parameter $\boldsymbol{\mu}$;

(ii) building an inverse solver to estimate the field $u(\boldsymbol{\mu})$ from (possibly sparse) data observation Y_{o} when the true parameter $\boldsymbol{\mu}$ is unknown;

(iii) building an inverse solver to estimate the parameter $\boldsymbol{\mu}$ that can give rise to an observed state Y_{o} and the related field $u(\boldsymbol{\mu})$.

The work in [36] suggested a non-intrusive reduced order model with machine learning methods to realize the first goal. Considering the quick response of this forward solver, a naive approach based on Latin Hypercube Sampling (LHS) direct search is used in building the inverse solver to tackle the second and third problems. Consequently, additional work in modelling and analysis is needed to realise a more practical RPODT. This is what we discuss next.

This paper is organised as follows. Section 2 introduces the related work and highlight the contribution of this paper. In Section 3, the forward model with machine learning based singular value decomposition (SVD) autoencoder (AE) and the inverse model with generalised latent assimilation are introduced. In Section 4, we introduce the detail applications to nuclear reactor physics. We first set up the reactor physics operation problems, then show the application of the forward model in Section 4.2 and show in Section 4.3 the application of the inverse model for parameter identification. The computational performance of the proposed RPODT is presented in Section 4.4. At last, some conclusion and perspectives are drawn in Section 5.

2. Related work and contribution

In this section, we present the related work regarding neutron field modelling and simulation tools, reduced-order-modelling and data assimilation for digital twins.

2.1. Neutron field modeling and simulation tools

One of the key components of the RPODT is the simulation models. For the modeling and simulation of a reactor core, the most unique and important physical field one concerns is neutron field and the related core power distribution. The behavior of the neutron field is characterized by physical laws, which are often represented by governing equations in the form of Partial Differential Equations (PDEs), i.e., neutron transport equations, or their approximations [49, 40, 69].

In a classical computational setting, numerical models discrete the governing equations, approximating the solution fields in different ways. The current mature industrial standard codes use different nodal methods to represent the field with node average value followed by pin power reconstruction in each node, such as CORCA-3D [2], COCAGNE [17], SIMULATE-5 [9], ANC9 [14], PARCS [30], DONJON4 [39], DYN3D [10], SCOPE2 [72], Bamboo-Core [77], DIF3D-VARIANT [68] etc. The typical computational time for one simulation of the whole core amounts to 30 seconds as reported in [4].

The recent programs such as CASL [70], NURESIM [15] and NURESIM [20] apply different strategies to develop the next generation high-fidelity simulation tools for NPP. Many of the codes developed are finite element-based with high-fidelity model and even coupled with other codes. For example, under CASL, the core simulator portion of VERA-CS [48], uses the advanced method of characteristics (MOC) to solve the neutron transport equations (MPACT

code), coupled with the sub-channel analysis code CTF and the fuel performance code BISON directly. However, even with High-Performance Computing (HPC), the computational demand of such high-fidelity codes is far too high for a routine industrial application [74]. Thus the scientifically attractive approach remains a perception that it's immediately industrial adoption is currently impractical, due to the required computer power and the often strict time schedules for design and operation usage. However, this does not imply that the high-fidelity scheme is not viable in the long-term, especially in cases where there are no strict time constraints, e.g., to create a detailed understanding of design/operational effects, such as validation of a new type core [48].

In general, these tools quoted above meet the criteria for a digital model, but they are far from the level of constructing an real-time operational digital twin on their own. However, they can still serve a role in support of operational digital twin as a basis using reduced-order modelling (ROM) techniques. Particularly, further advances in ROM may help the investments in high-fidelity simulation give much return in the digital twin framework.

2.2. Reduced order models for digital twins

The reduced order modelling [11] offers new opportunities in reduction of computational effort that neutron simulation codes demand, without significant loss of precision. Therefore ROM is another key technology for real-time operational digital twin development for operation support. We refer to the recent surveys [3, 12, 28] for a thorough introduction to ROM.

Typically ROM is intrusive, requiring full knowledge of the governing equations and the numerical strategy of the full model of the underling physical system. A general way to construct the intrusive ROM is to project the (high-fidelity) full-order model onto the reduced space [12]. In most of nuclear engineering applications, access(probe) to the detail numerical frame of proprietary/commercial codes is generally unavailable. Thus the full model code is always viewed as black-box solver by users, which in turn limits the implementation of traditional ROM. In contrast, non-intrusive ROM aims to construct reduced order model by building the input-output map of the input parameter and the reduced basis through interpolation [55, 75], regression [7, 8, 76], machine-learning [42, 54, 18] or other approaches [57, 60, 13, 19]. Thus non-intrusive ROM is amenable to representation simulation components of the digital twin as primarily real-time.

2.3. Data assimilation with ROM for digital twins

ROM also offers new opportunities for the integration of simulation models and physically observable quantities, which may be in a small-data scenarios. These approaches are also regard as data assimilation with ROM [6], where the data is incorporated into a reduced model, see recent works [41, 4, 24, 29] for detail description of this framework. This framework makes the combination of simulation model with measurement data more efficient to realise real-time operational digital twin. In more detail, simulation models of the digital twin are represented by reduced basis, while the coefficients of the reduced basis and the caused input parameters are determined with observations in an inverse framework. In the offline phase, ROM provides a way to learn the behaviour of the underline physical system, thus allows extracting the principal component information of the system and further more providing a guideline to set the amount and locations of instruments needed for economical consideration [4]. In the online phase, ROM speeds up computations allowing better explorations of the parameter space at an acceptable computational cost to meet the real-time requirement of operational objective.

2.4. Previous work

Our previous work [36] constructed a prototype of RPODT, that combines reduced-order models with machine learning to predict neutron flux and power distributions in a real-time level. The digital twin is not only able to solve forward problems given an input parameter, but also to solve inverse problems given some extra measurements. In the offline phase, the proper orthogonal decomposition (POD) is used to extract the reduced models, while the machine learning techniques, namely, K-nearest neighbours (KNN) and Decision tree (DT) are used to learn the map of the input-parameters and coefficients of the reduced basis. Thereafter the high-fidelity fields are able to be solved in a non-intrusive ROM way. In the online phase, the real time input parameters are used to rapidly reconstruct the neutron field with the digital twin. Further more, the input parameters can be adapted based on the real time measurements in an inverse framework. The effectiveness of the framework is illustrated based on HPR1000 [50] reactor core modelled by two-group neutron diffusion equations affected by the general input parameters of the reactor, such as control rod inserting step, burnup, power level and temperature of the coolant, etc. However, the investments in efficient forward and inverse methods in the digital twin may need further advances. For example, building a forward model with non-linear machine learning is worth trying, in order to catch the non-linear dependence of the parameter. In building the

inverse model, the naive approach shall be replaced with more meaningful methods with theoretical guidance, and the robustness with respect to observation noise shall also be investigated.

2.5. Contributions of this work

In the previous work of [36], the general POD is used to construct the reduced model in a linear framework. Several works [45, 62, 63] have highlighted the connection between AE and SVD (or POD), additionally, by choosing nonlinear activation functions, AE with SVD paves a way to perform nonlinear dimensionality reduction. Particularly, recent works [62, 63] provide a novel hybrid SVD AE in developing the reduced-order model to solve eigenvalue problems of neutron diffusion equations, and its efficiency is illustrated numerically via a one dimension and two dimension benchmark problems. Though the SVD AE was implemented in an intrusive way as a projection-based ROM in [62], it is more suitable for implementation in a non-intrusive way directly. One of our main contribution in this work it to adapt the SVD AE to build a non-intrusive ROM which is able to couple with any industry code in building the real-time operational digital twin. The non-intrusive property is achieved again by KNN or DT to map the input parameter to the latent space i.e., the coefficients of the reduced basis. As we will show later, the SVD AE provides more accurate result than linear POD does.

Further more, because of the high-efficiency of the machine learning based ROM forward model, when constructing the inverse model to infer the input parameter with online sensor data, we proposed a relatively naive approach in [36] to approximate the model parameters from an ensemble of samplings generated using LHS around the initial guess. Meanwhile, the latest progress of data assimilation in latent space [61, 1, 16, 27, 52] (also called latent assimilation) with machine learning make it possible to further simplify our inverse problem instead of the naive sampling approach. Thus our second contribution in this work it to adopt the generalised latent assimilation to solve the inverse problem, which ensures the high efficiency of the operation digital twin. Further more, the robust of the inverse algorithm with respect to observation noise is also tested, which shows a satisfactory result for engineering usage.

3. Digital twin architecture and methodology

A typical architecture of RPODT involves forward and inverse models. The forward model is a non-intrusive reduced order model, which permits to solve the forward problem for a given input parameter in a second or even less to meet real time properties. Particularly, we propose a machine learning based SVD AE to build the forward model in this work. Meanwhile, the CORCA-3D code, referred to as the high-fidelity model, is used to provide high-fidelity solutions for training the reduced order model. The inverse model is realized in a generalised latent assimilation framework, with a combination of the non-intrusive reduced order model and online observations from a nuclear reactor core, to identify which input parameter should comprise the current state.

3.1. Forward model: machine learning based SVD AE

This section briefly describes the key stages of constructing a non-intrusive reduced order model, including preparing high-fidelity snapshots, computing the reduced basis and SVD AE, learning the map from input parameter space to reduced space with machine learning, and constructing a projection-based reduced-order model.

3.1.1. High-fidelity snapshots

Consider a reactor core that maps an input parameter $\boldsymbol{\mu}$ onto the power distribution field u, i.e., the power field u varies in space, and depends on the input parameter of the system. Thus we have $u : \Omega \times \mathcal{D} \to \mathbb{R}$, with the input parameter domain $\mathcal{D} \subset \mathbb{R}^p$ of dimension p and the spatial domain $\Omega \subset \mathbb{R}^d$ which represents the reactor core of dimension d. In our practical nuclear engineering applications, the power field $u(\boldsymbol{r}, \boldsymbol{\mu})$ is solved with the proprietary CORCA-3D code package[2]. The solution manifold $\mathcal{M} := \{u(\boldsymbol{r}, \boldsymbol{\mu}) |\ \boldsymbol{r} \in \Omega, \boldsymbol{\mu} \in \mathcal{D}\}$ is composed of solutions of all possible input parameters from $\mathcal{D}$. For brevity, in the left part of this work, we omit $\boldsymbol{r}$ and denote by $u(\boldsymbol{\mu})$ the parameter-dependent field unless otherwise specified.

3.1.2. Computing the reduced basis

The ROM aims to provide an efficient low-dimensional representation of the snapshots. Such a representation can be achieved using reduced basis [56, 4, 24], particularly the POD basis [67, 34], which is typically computed using discrete sampled training data, i.e., snapshots sampled from $\mathcal{M}$. Consider a collection of N snapshots $\mathcal{S} = \{\boldsymbol{u}(\boldsymbol{\mu}_k)\}_{k=1}^N$ corresponding to a discrete parameter set $\mathcal{D}_N = \{\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_N\} \subseteq \mathcal{D}$, where the M-dimensional vector $\boldsymbol{u} \in \mathbb{R}^M$ is the

discrete version of the power field u. Define the snapshot matrix $\boldsymbol{S} \in \mathbb{R}^{M\times N}$ which contains the snapshots $\boldsymbol{u}(\boldsymbol{\mu}_k)$ as its columns. The correlation matrix $\mathbf{C}$ is then-formed by computing the inner product between each pair of snapshots, i.e.,

$$\mathbf{C}_{i,j} = \frac{1}{N}(\boldsymbol{u}_i, \boldsymbol{u}_j), \quad \forall 1 \leq i, j \leq N. \tag{1}$$

The eigenvalues $\{\lambda_i\}$ and the corresponding eigenvectors $\{\boldsymbol{v}_i\}$ of $\mathbf{C}$ are then computed. The j-th POD basis $\boldsymbol{q}_j$ is given by a linear combination of snapshots

$$\boldsymbol{q}_j = \sum_{i=1}^{N} \boldsymbol{v}_j^i \phi_i, \tag{2}$$

where v_j^i denotes the i-th element of the j-th eigenvector. The magnitude of the j-th eigenvalue λ_j describes the relative importance of the j-th POD basis. Assemble the first n POD basis in a matrix $\boldsymbol{V}_n = [\boldsymbol{q}_1, ..., \boldsymbol{q}_n] \in \mathbb{R}^{M\times n}$. Among all orthonormal bases of size n, the POD basis minimizes the least squares error of the reconstruction of snapshot matrix $\boldsymbol{S}$,

$$\min_{\boldsymbol{V}_n} \|\boldsymbol{S} - \boldsymbol{V}_n \boldsymbol{V}_n^T \boldsymbol{S}\|_F^2 = \sum_{k=n+1}^{N} \lambda_k \, . \tag{3}$$

Thus the POD basis provides an efficient low-dimensional representation of the field in the manifold $\mathcal{M}$. For any $\boldsymbol{u}(\boldsymbol{\mu}) \in \mathcal{M}$, the n-dimensional representation of $\boldsymbol{u}(\boldsymbol{\mu})$ is approximated by

$$\boldsymbol{u}(\boldsymbol{\mu})_n = \sum_{i=1}^{n} \alpha_i(\boldsymbol{\mu}) \boldsymbol{q}_i \, . \tag{4}$$

3.1.3. SVD autoencoder

Traditional POD with equation 4 provides an linear expansion of $u(\boldsymbol{\mu})$ with POD basis. Autoencoders provide a nonlinear model reduction alternative to POD/SVD [45], that may capture features of the high-fidelity snapshots more efficiently. An autoencoder is an unsupervised learning technique in which we leverage a special type of feed-forward neural network for the task of representation learning. In particular, we build a neural network architecture which forces a reduced representation of the original input data by imposing a bottleneck .

As shown in Figure 1, the core of a SVD AE is an autoencoder, which is composed of two networks: the encoder and the decoder, the former maps an input data to the latent variables, and the latter maps from the latent variables to the output data. If we denote the input data by $\boldsymbol{x} \in \mathbb{R}^M$, the output data by $\hat{\boldsymbol{x}} \in \mathbb{R}^M$ and the latent variables by $\tilde{\boldsymbol{x}} \in \mathbb{R}^n$, then the autoencoder can be represented as a composition of and encoder and decoder functions,

$$\hat{\boldsymbol{x}} = f^{\mathrm{AE}}(\boldsymbol{x}, \boldsymbol{w}^{\mathrm{AE}}) = f^{\mathrm{DEC}}(\tilde{\boldsymbol{x}}, \boldsymbol{w}^{\mathrm{DEC}}) = f^{\mathrm{DEC}}(f^{\mathrm{ENC}}(\boldsymbol{x}, \boldsymbol{w}^{\mathrm{ENC}}), \boldsymbol{w}^{\mathrm{DEC}}), \tag{5}$$

where f^{AE}, f^{ENC}, and f^{DEC} represent the autoencoder, encoder, and decoder, with associated weights $\boldsymbol{w}^{\mathrm{AE}}$, $\boldsymbol{w}^{\mathrm{ENC}}$ and $\boldsymbol{w}^{\mathrm{DEC}}$ respectively. The encoder and the decoder are then trained jointly with the mean square error (MSE) as the loss function

$$J(\boldsymbol{w}^{\mathrm{ENC}}, \boldsymbol{w}^{\mathrm{DEC}}) = \frac{1}{N_{\mathrm{train}}^{\mathrm{AE}}} \sum_{j=1}^{N_{\mathrm{train}}^{\mathrm{AE}}} \|\boldsymbol{x}_j - \hat{\boldsymbol{x}}_j\|^2, \tag{6}$$

where $N_{\mathrm{train}}^{\mathrm{AE}}$ represents the size of the AE training dataset. Figure 1 gives an example of the architecture of SVD AE with a fully connected autoencoder with three encoder layers and three decoder layers. With this network, a twelve-dimensional high-fidelity field $\boldsymbol{u}$ is first reduced into a six-dimenaional variable $\boldsymbol{x}$ in the form of reduced basis coefficients $\boldsymbol{\alpha}$, then this six-dimenaional variable is compressed into one latent variable $\tilde{\boldsymbol{x}}$ and then expands back to the original six-dimenaional variable $\hat{\boldsymbol{x}}$, with the six-dimenaional variable as coefficients $\hat{\boldsymbol{\alpha}}$ of the six reduced basis $\{\boldsymbol{q}_1, ..., \boldsymbol{q}_6\}$, the field $\hat{\boldsymbol{u}}$ can be reconstructed in the form of $\hat{\boldsymbol{u}} = \sum_{i=1}^{6} \alpha_i \boldsymbol{q}_i$.

The combination of POD and AE (also known as SVD AE) was introduced in the recent works of [62, 63] for constructing a reduced-order model of neutron diffusion eigenvalue problems. This method consists of two stages of dimension reduction. In the first stage, we apply the POD to obtain the principle components of the associated field; in the second stage, we use the first n principle components' coefficients (corresponding to the first n largest eigenvalues) as input data to train a dense autoencoder with fully connected neural networks to further reduce the problem dimension by learning the nonlinearity of the field. As pointed out by [62], since the POD considerably reduces the size of the input data (compared to using high-fidelity data) in AE, applying fully connected neural networks layers is computationally affordable without concerning over-parameterization. Thus the input data of SVD AE is the first n POD coefficients $\boldsymbol{x} = \boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)^T \in \mathbb{R}^n$, and the latent variables $\tilde{\boldsymbol{x}}$ is of dimension r, with $r < n$. In this work, we use the α-leaky rectified linear unit (α-LReLU) as the activation function of neural networks, where $\alpha = 0.2$ as suggested in [33]. The detail structure of the SVD AE used in this work is summarised in Table 1.

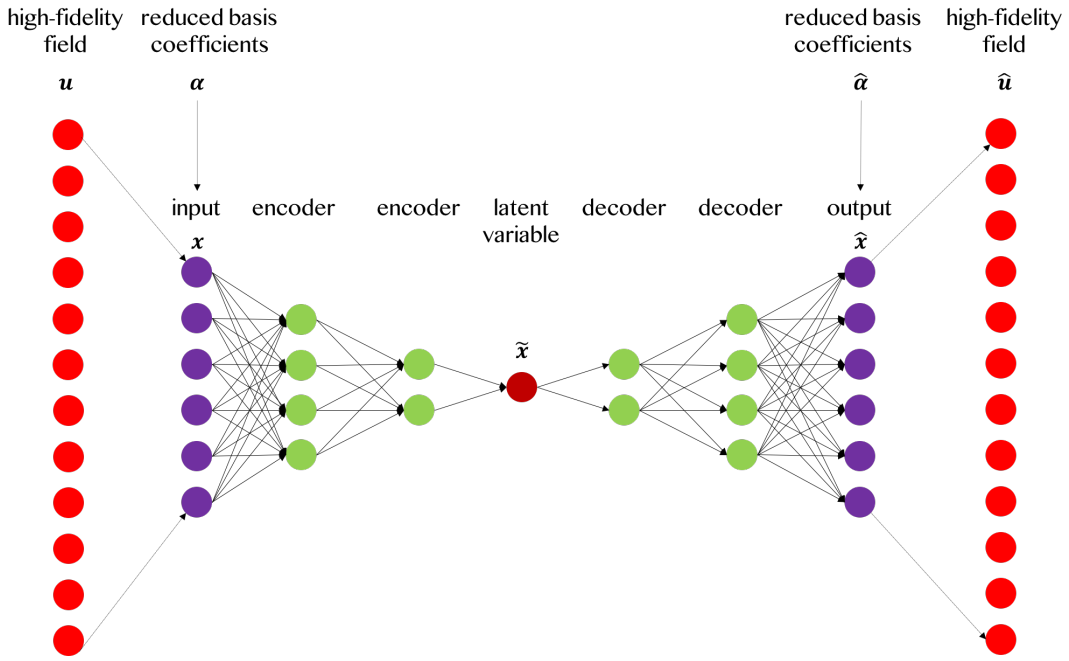


Figure 1: An example of the architecture of SVD AE with a fully connected autoencoder with three encoder layers and three decoder layers. With this network, a twelve-dimensional high-fidelity field $\boldsymbol{u}$ is first reduced into a six-dimenaional variable $\boldsymbol{x}$ in the form of reduced basis coefficients $\boldsymbol{\alpha}$, then this six-dimenaional variable is compressed into one latent variable $\tilde{\boldsymbol{x}}$ and then expands back to the original six-dimenaional variable $\hat{\boldsymbol{x}}$, with the six-dimenaional variable as coefficients $\hat{\boldsymbol{\alpha}}$ of the six reduced basis $\{\boldsymbol{q}_1, ..., \boldsymbol{q}_6\}$, the field $\hat{\boldsymbol{u}}$ can be reconstructed in the form of $\hat{\boldsymbol{u}} = \sum_{i=1}^{6} \alpha_i \boldsymbol{q}_i$

3.1.4. Non-intrusive reduced order model

Once the reduced space is constructed by POD or SVD AE, the final work to build a non-intrusive ROM is to construct a map from the input parameter vector $\boldsymbol{\mu}$ to the coefficients of the reduced space. As noted in [36], DT and KNN are more fit for engineering applications with limited data [46], that is exactly the circumstances we are interested in this work. We follow the work [36] and use DT and KNN to learn the input-output map, and we further denote the non-intrusive machine learning forward model by $f^{\mathrm{ML}}(\boldsymbol{\mu})$. For SVD AE, in the train stage, the input data is the first n POD coefficients $\boldsymbol{\alpha}(\boldsymbol{\mu})$ and the output data is the latent variable $\tilde{\boldsymbol{x}}$, thus we have

$$\tilde{\boldsymbol{x}}(\boldsymbol{\mu}) = f^{\mathrm{ENC}}(\boldsymbol{\alpha}(\boldsymbol{\mu})). \tag{7}$$

Once $\tilde{\boldsymbol{x}}$ is solved by $f^{\mathrm{ENC}}(\boldsymbol{\alpha}(\boldsymbol{\mu}))$ for a training set of $\boldsymbol{\mu}$, we can train the map from $\boldsymbol{\mu}$ to $\tilde{\boldsymbol{x}}$ with DT or KNN algorithm,

$$\tilde{\boldsymbol{x}}(\boldsymbol{\mu}) = f^{\mathrm{ML}}(\boldsymbol{\mu}). \tag{8}$$

Table 1
Neural network structure of the SVD AE

Layer (type)	Output Shape	Activation
Encoder		
Input	500	
Dense (128)	128	0.2-LReLU
Dense (30)	30	0.2-LReLU
Decoder		
Input	30	
Dense 128	128	0.2-LReLU
Dense 799	500	0.2-LReLU

Then we use the POD basis to recover the high-fidelity field $\boldsymbol{u}^{\mathrm{ML}}$ for a given $\boldsymbol{\mu}$. In the end, $\boldsymbol{u}^{\mathrm{ML}}(\boldsymbol{\mu})$ can be represented by $f^{\mathrm{ML}}(\boldsymbol{\mu})$ as follows,

$$\boldsymbol{u}^{\mathrm{ML}}(\boldsymbol{\mu}) = f^{\mathrm{ML}}(\boldsymbol{\mu}) = \boldsymbol{V}_n \hat{\boldsymbol{\alpha}}^{\mathrm{ML}}(\boldsymbol{\mu}) = \boldsymbol{V}_n \hat{\boldsymbol{x}}^{\mathrm{ML}}(\boldsymbol{\mu}) = \boldsymbol{V}_n f^{\mathrm{DEC}}(\tilde{\boldsymbol{x}}^{\mathrm{ML}}(\boldsymbol{\mu})) = \boldsymbol{V}_n f^{\mathrm{DEC}}(f^{\mathrm{ML}}(\boldsymbol{\mu})). \tag{9}$$

Note here that, the setting of the hyper-parameters of DT and KNN are the same as used in [36], and we will show later in the Section 3.3.

3.2. Inverse model with generalised latent assimilation

For the inverse modelling, only partial observations of the entire field are available. Data-driven models such as data assimilation (DA) approaches can be applied [5, 23]. Recently, latent assimilation techniques are introduced in the work of [1, 61, 27] where the DA is performed after having compressed the state and the observation data into the same latent space. However, as noted in the work [25], it is almost infeasible to compress the full state space and observations into a same latent space in a wide range of DA applications, where only a part of the state are observable. Thus in this work, the Generalised Latent Assimilation (GLA) is used, in which we split the observations and the background field into two parts, that only the background field stays in the latent space. In this work, the background state $\boldsymbol{\mu}_b$ is of dimension four and represents the initial guess of the four model parameters. Applying the machine learning forward model f^{ML} and the decoder function f^{DEC}, a background prediction of the full physical space can be obtained,

$$\boldsymbol{x}_b = f^{\mathrm{DEC}}(f^{\mathrm{ML}}(\boldsymbol{\mu}_b)). \tag{10}$$

Denote by $\mathcal{H}^{\mathrm{full}}$ the transformation function which links the full physical field $\boldsymbol{V}_n\boldsymbol{x}$ to observations $\boldsymbol{y}$, i.e.,

$$\boldsymbol{y} = \mathcal{H}^{\mathrm{full}}(\boldsymbol{V}_n\boldsymbol{x}) + \epsilon_y, \tag{11}$$

where ϵ_y is the observation error, supposed to be centered Gaussian in this paper. Applying variational assimilation to estimate the model parameters $\boldsymbol{\mu}$ will lead to the loss function,

$$J = \frac{1}{2}||\boldsymbol{\mu} - \boldsymbol{\mu}_b||^2_{\boldsymbol{B}^{-1}} + \frac{1}{2}||\boldsymbol{y} - \mathcal{H}^{\mathrm{full}}(\boldsymbol{V}_n f^{\mathrm{DEC}}(f^{\mathrm{ML}}(\boldsymbol{\mu})))||^2_{\mathbf{R}^{-1}} \tag{12}$$

where $\mathbf{B}$ and $\mathbf{R}$ denote the prior errors in the state and the observation space, respectively [22]. However, as pointed out by the recent works of [25, 26], the minimisation of equation (12) can be challenging due to the complexity and the non-differentiability of ML functions. Following their ideas, we make use of the GLA algorithms by computing a local polynomial surrogate functions that links $\boldsymbol{\mu}$ and $\boldsymbol{y}$. More precisely, LHS is performed to build a polynomial regression (PR) learning ensemble $\{\boldsymbol{\mu}_b^q\}_{q=1..n_s}$ around the background state within certain range r_s. We then fit the forward model and the decoder fundction by a local polynomial function $\tilde{\mathcal{H}}^p$, i.e.,

$$\tilde{\mathcal{H}}^p = \underset{\mathcal{P}\in P(d_p)}{\operatorname{argmin}}\Big(\sum_{q=1}^{n_s} ||p(\boldsymbol{\mu}_b^q) - \mathcal{H}^{\mathrm{full}}(\boldsymbol{V}_n f^{\mathrm{DEC}}(f^{\mathrm{ML}}(\boldsymbol{\mu}_b))||_2^2\Big)^{1/2}, \tag{13}$$

where $P(d_p)$ denotes the set of polynomial functions of degree d_p and $\{d_p, n_s, r_s\}$ are the hyperparamters of the GLA algorithm. Thus, the cost function J_{GLA} of GLA problems reads:

$$J_{\text{GLA}} = \frac{1}{2}||\boldsymbol{\mu} - \boldsymbol{\mu}_b||^2_{\boldsymbol{B}^{-1}} + \frac{1}{2}||\boldsymbol{y} - \tilde{\mathcal{H}}^p(\boldsymbol{\mu})||^2_{\boldsymbol{R}^{-1}}, \tag{14}$$

$$\boldsymbol{x}_a = \underset{\boldsymbol{x}}{\text{argmin}}\ J_{\text{GLA}}. \tag{15}$$

A flowchart of the GLA is illustrated in Figure 2 and the algorithm is summarised in Algorithm 1.

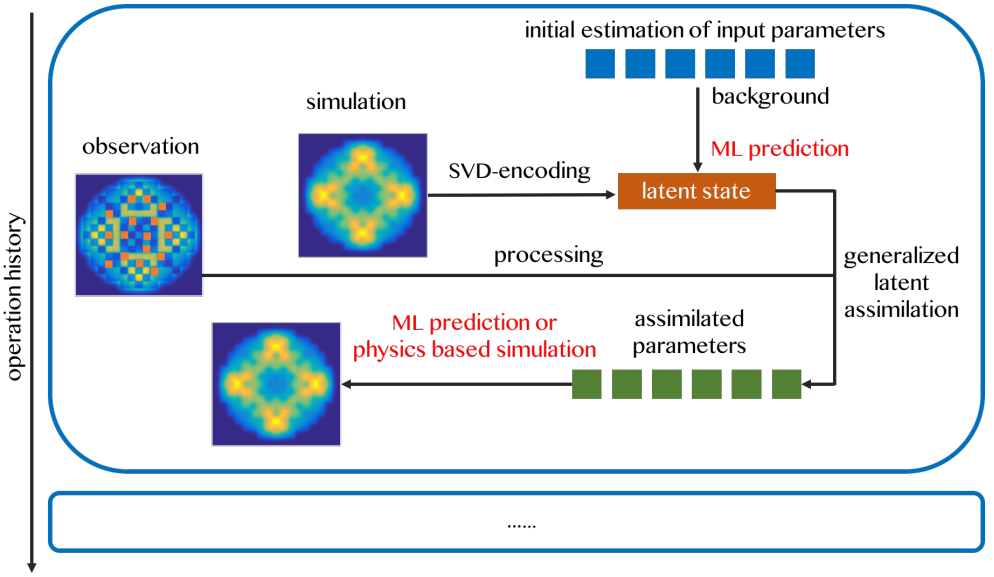


Figure 2: Flowchart of the inverse model for parameter identification using GLA.

Algorithm 1: Three-dimensional variational data assimilation (3Dvar) GLA

Inputs: $\boldsymbol{\mu}_b, \boldsymbol{V}_n, \boldsymbol{B}, \boldsymbol{R}, \boldsymbol{y}, \{f^{\text{ML}}\}$
Paramters: d_p, r_s, n_s

$\{\boldsymbol{\mu}_b^q\}_{q=1..n_s} = \text{LHS Sampling}_{\{d^p, r_s, n_s\}}(\boldsymbol{\mu}_b)$
for $q = 0$ *to* n_s **do**
  $\boldsymbol{y}^q = \mathcal{H}^{\text{full}}(\boldsymbol{V}_n f^{\text{DEC}}(f^{\text{ML}}(\boldsymbol{\mu}_b^q)))$
end
$\tilde{\mathcal{H}}^p = \text{PR}_{\text{train}}\left(\text{input:}\{\boldsymbol{\mu}_b^q\}, \text{output:}\{\boldsymbol{y}^q\}, q = 1..n_s\right)$
$\boldsymbol{\mu}_a = \underset{\boldsymbol{\mu}}{\text{argmin}}\left(\frac{1}{2}||\boldsymbol{\mu} - \boldsymbol{\mu}_b||^2_{\mathbf{B}^{-1}} + \frac{1}{2}||\mathbf{y} - \tilde{\mathcal{H}}^p(\boldsymbol{\mu})||^2_{\mathbf{R}^{-1}}\right)$
outputs: μ_a

3.3. Hyperparameter tuning

In this study, an ensemble of 18479 CORCA-3D simulations is split into three datasets with 14479, 1000 and 3000 samples, used for training, validation and testing, respectively. The forward machine learning models and ROMs are trained on the training set. The exact networks for the SVD AE approach are shown in Table 1. To select the most appropriate model parameters, the hyperparameter tuning is carried out on the validation dataset as shown in Table 2.

Table 2
Hyperparmater grid search space

Model/Hyperparamters	Grid search space	Final set
DT		
split criteria	{ 'gini', 'entropy'}	'gini'
n_{features}	{ 'log2','sqrt'}	'sqrt'
KNN		
k	{ 5,10,20}	5
Metric	$\{L_1, L_2\}$	L_2
GLA		
d_p	2-6	4
n_s	$\{200, 500, 1000, 2000\}$	1000
r_s	$10\% - 100\%$	30%

4. Application to nuclear reactor physics

4.1. Set up of reactor physics operation problems

The realistic engineering background is to reconstruct the power distribution in a reactor core, e.g., during the operation of the PWR HPR1000 [50]. The core is filled with a total of 177 vertical nuclear fuel assemblies, among those 44 are instrumented with self powered neutron detectors (SPNDs) which are used to measure the neutronic activity fields thereafter the power field in the core. In Fig 3, we illustrate a horizontal slice of the HPR1000 core and an axial slice of an assembly with SPNDs implemented in. Because of the symmetry along the x and y axis, only one quarter of the core is given. The fuel assemblies in gray represent the assemblies with control rod, and those marked with D present the assemblies with SPNDs. We refer to [50] for a glance at HPR1000 and the generic neutronic physical model of the reactor. For the sake of completeness, we list here the description of the model with minor modification. This part is original introduced in [35, 36] for the study of data assimilation and then the first work in RPODT.

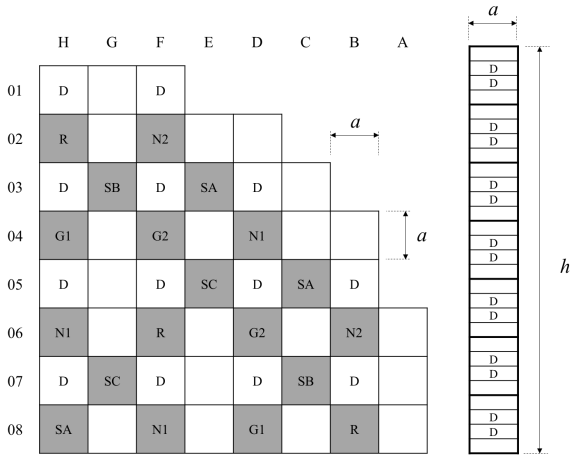


Figure 3: A quarter of the HPR1000 core in radial direction (white square: fuel assembly with SPNDs, gray square: fuel assembly with control rods, D: fuel assembly with neutron detectors) [36].

During the normal operation of HPR1000, two types control rods are used to control the reactor, the first one is compensation rods, which used for coarse control and/or to remove reactivity in relatively large amounts. There are four subtypes compensation rods, i.e., G1,G2,N1,N2. The second one is regulating rods (R rods), which are used for fine adjustments and to maintain desired power or temperature [50]. We consider the power evolution induced by i) the movement of the control rods, ii) the burnup of the nuclear fuel in the whole core, iii) the variation of the power level of the reactor core and iv) the variation of the inlet temperature coolant. The evolution of the power field u is

modelled by two-group diffusion equations as shown in [36], and solved numerically with the advanced node core code package CORCA-3D [2], developed at Nuclear Power Institute of China (NPIC), which is able to solve the 3D few-group diffusion equations, consider the thermal-hydraulic feedback, and the pin-by-pin power reconstruction.

In this work, we treat CORCA-3D as a black box solver with input parameter $\boldsymbol{\mu}$ and output $\boldsymbol{u}$. Specifically, the power field $\boldsymbol{u}$ is limited to depend on the vector of "general" parameters which give the stage of the life cycle:

$$\boldsymbol{\mu} := (S_t, B_u, P_w, T_{in}). \tag{16}$$

- S_t: the inserting steps of control rods, which ranges from 0 to 615. Consider the movement of the compensation rods from all rod clusters out (ARO) to fully inserted (615 steps, considering the overlap steps).
- B_u: the average burnup of the fuel in the whole core. It is a measure of how much energy is extracted from the fuel so it is an increasing function in time. It ranges between 0 (for the beginning of the life cycle) and $B_{u,\max}$ (the end of the life cycle, in this work, we set $B_{u,\max}$=2500 MWd/tU) and its exact evolution depends on the operating history of the reactor.
- P_w: the power level of the reactor core. It ranges between 0.3 and 1 FP (Full Power).
- T_{in}: core coolant temperature at inlet. It ranges between 290 and 300 °C.

Thus, $\boldsymbol{u}$ can be represented by $\boldsymbol{u}(\boldsymbol{\mu}) = \boldsymbol{u}(S_t, B_u, P_w, T_{in})$ implicitly thanks to CORCA-3D. There are 177 fuel assemblies in HPR1000, each assembly is numerically represented using 28 vertical levels. Thus, $\boldsymbol{u}$ is a vector of dimension $M = 4956$ (177×28). The discrete solution set $\boldsymbol{S} := \{\boldsymbol{u}(\boldsymbol{\mu}) | \boldsymbol{\mu} \in \mathcal{D}^s\}$ consists of 18480 solution snapshots with the parameter sampling scheme $\mathcal{D}^s := B_u^s \otimes S_t^s \otimes P_w^s \otimes T_{in}^s$, where $S_t^s = \{0, 1, ..., 615\}$, $B_u^s = \{0, 50, 100, 150, 200, 500, 1000, 1500, 2000, 2500\}$, $P_w^s = \boldsymbol{RU}3(30, 100)$, and $T_{in}^s = \boldsymbol{RU}3(290, 300)$, the operator $\boldsymbol{RU}3(a, b)$ represents a random uniform sampling in the closed set $[a, b]$, $(a < b)$ for three times.

The configuration of the observations in node wise can be found in Figure 3, and the observation vector $\boldsymbol{y}$ (possibly with noise) is of dimension $m = 308$ (44×7) in the whole core. In this work, observations used in the analysis process are not coming from real measurements, but coming from numerical simulations with CORCA-3D. The observations $\boldsymbol{y}$ with combination of prior information on the input parameter $\boldsymbol{\mu}_b$ are used to improve the quality in estimating the input parameter $\boldsymbol{\mu}$, thereafter the power field $\boldsymbol{u}$. In this work, the case with observation noise is also studied.

4.2. Forward model prediction with reduced order model

We first assess the performance of machine learning based reduced order models to solve the forward problem for a given input parameter $\boldsymbol{\mu}$. There are four models considered, the KNN POD model, the KNN SVD AE model, the DT POD model and the DT SVD AE model. In this study, an ensemble of 18479 CORCA-3D simulations is split into three datasets with 14479, 1000 and 3000 samples, used for training, validation and testing, respectively. The forward machine learning models and ROMs are trained on the training set. The detail hyperparameter tuning policy is shown in Table 2.

Figure 4 presents the decay of the first 1000 singular values of the discrete manifold $\boldsymbol{S}$ of 18479 CORCA-3D snapshots. As shown in Table 1, We select the first 500 POD basis and the corresponding coefficients to feed the SVD AE, to make sure the truncated error over the discrete manifold less than 10^{-5}, which seems to improve three orders of magnitude than just select the first 30 POD basis to build the reduced model directly. In this regards, we try to compress a 500 dimensional POD coefficient information into a 30 dimensional latent variable, by further learning the nonlinearity of the manifold with SVD AE.

For a given input parameter $\boldsymbol{\mu}$ in the test dataset, we first introduce the relative L_2 error e^{ML} as shown in Equation 17 to evaluate the quality of each model, where the superscript ML presents the four models quoted above, and $\boldsymbol{u}^{\mathrm{True}}$ presents the field in the test set, which is regarded as the true field.

$$e^{\mathrm{ML}} = \frac{\|\boldsymbol{u}^{\mathrm{ML}} - \boldsymbol{u}^{\mathrm{True}}\|_{L_2}}{\|\boldsymbol{u}^{\mathrm{True}}\|_{L_2}} \tag{17}$$

To evaluate the overall quality of each model, we then introduce the average relative L_2 error E^{ML} over the test dataset as shown in Equation 18, where $N^{\mathrm{test}} = 3000$ in this work.

$$E^{\mathrm{ML}} = \frac{\sum_{i=1}^{N^{\mathrm{test}}} e_i^{\mathrm{ML}}}{N^{\mathrm{test}}} \tag{18}$$

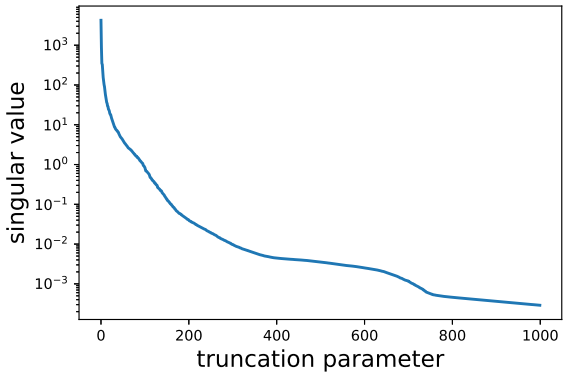
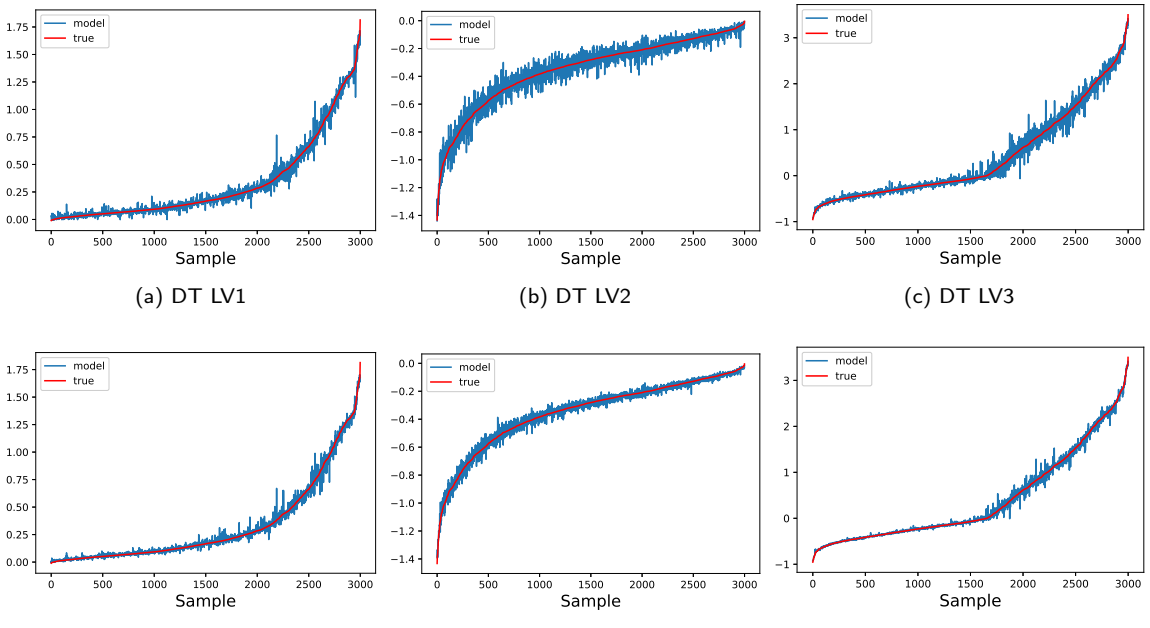


Figure 4: The decay of the first 1000 singular value of the snapshots manifold.

We first show in Figure 5 the behaviour of the first three latent variables for KNN SVD AE and DT SVD AE models in the test dataset. In each figure, the true values are rearranged in ascending order. It is clear that the latent variables learned by KNN stand closer to the true values than DT does, thus KNN based models provide more accuracy results than DT. This could also be confirmed by the results shown in Figure 6.

We illustrate in Figure 6 the average error E^{ML} in the full physical space against the dimension of the latent/reduced space for the four models, i.e., the KNN POD, KNN SVD AE, DT POD and DT SVD AE models. We investigated the performance corresponds to the latent/reduced space of dimension $n = 5, 10, 20, 30$. Obviously, compared with the results of POD based forward machine learning models, the accuracy of SVD AE based forward machine learning models has been significantly improved. Particularly, with only 20 latents variables, the SVD AE convergences to the best accuracy, meanwhile, POD needs more than 30 reduced basis, the latter would results more computing resource.



(a) DT LV1 (b) DT LV2 (c) DT LV3

(d) KNN LV1 (e) KNN LV2 (f) KNN LV3

Figure 5: Prediction results of latent variables (blue) against true values (red) in the test data set.

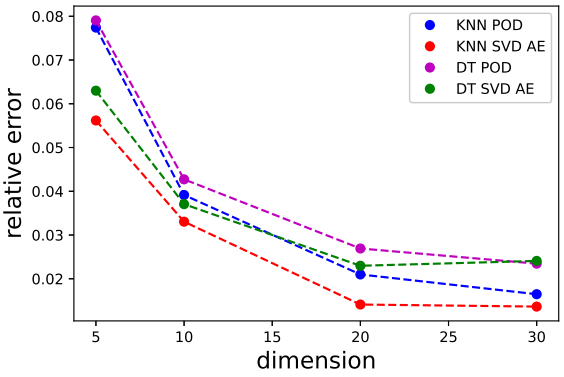


Figure 6: Prediction error in the full physical space against the dimension n of the latent space.

4.3. Inverse model for parameter identification

In this section, we investigate the performance of the inverse model with GLA, given noisy observations $\boldsymbol{y} = \mathcal{H}^{\text{full}}(\boldsymbol{u}) + \epsilon_y$, where ϵ_y is the observation error, supposed to be centered Gaussian distribution of standard derivation σ.

Before we start to analyse the performance of the inverse model, we first give in Figure 7 the importance rate of input variables in the form of *gini* coefficient in the DT forward algorithm. It is clear that the control rod inserting step S_t is the most important parameter that affects the output of the model, followed by the burnup B_u, and the total power level P_w. The parameter that has the least impact on the model is inlet temperature. This finding is consistent with common sense in reactor physics from the engineering point of view.

In Figure 8, we illustrate the absolute prior and post errors of the four input parameters for both DT and KNN methods, for the case of noise level $\sigma = 3\%$. This is again consistent with the result shown in Figure 7, that S_t parameter is more sensitive than other three, thus GLA algorithm improves the estimation of S_t a lot both for DT and KNN methods. Considering the different importance of each input parameter, we use the average relative L_2 error of the reconstructed power field $\boldsymbol{u}$ defined in Equation 19 to quantify the quality of the inverse model.

$$E^{\text{ML}}_{\text{inverse}} = \frac{\sum_{i=1}^{N^{\text{noise}}} e^{\text{ML}}(\boldsymbol{\mu}_a(\boldsymbol{y}_i))}{N^{\text{noise}}}, \tag{19}$$

where $e^{\text{ML}}(\boldsymbol{\mu}_a(\boldsymbol{y}_i))$ is the relative L_2 error defined in Equation 20 for a given observation sample $\boldsymbol{y}_i$.

$$e^{\text{ML}}(\boldsymbol{y}_i) = \frac{\|\boldsymbol{u}^{\text{ML}}(\boldsymbol{\mu}_a(\boldsymbol{y}_i)) - \boldsymbol{u}^{\text{True}}\|_{L_2}}{\|\boldsymbol{u}^{\text{True}}\|_{L_2}} \tag{20}$$

The parameter N^{noise} is set as follows. If the observation noise level $\sigma = 0$, we set $N^{\text{noise}} = 1$, i.e., only one sample is enough. If the noise level $\sigma \neq 0$, we set $N^{\text{noise}} = 50$, i.e., we sample the observation $\boldsymbol{y}$ for 50 times in Gaussian distribution, centered at zero with standard derivation σ.

We then show in Figure 9 the prediction results of latent variables with background in blue line and assimilated in yellow line parameters for the test dataset. The observation error level is set to be $\sigma = 3\%$ for both DT and KNN methods. It is clear that even with noisy observations, our proposed inverse model with GLA is able to make use of observation information, thus improve the estimation of latent variables with respect to a relative bad background (the initial guess). Further more, KNN model provides better estimation than DT does.

The results for the whole test dataset for noise level $\sigma = 4\%$ and $\sigma = 8\%$ for both DT and KNN methods are illustrated in Figure 10. In these figures, both background error and assimilated error are shown for comparison purpose. It is easy to find that with GLA, the error of the reconstructed field improves a lot than only using the background information. In more detail, we present in Figure 11 the error evolution of DT and KNN methods when the observation noise level varies from zero to 0.14. In this case, the background error is fixed to be ~ 0.22. From this figure, from the

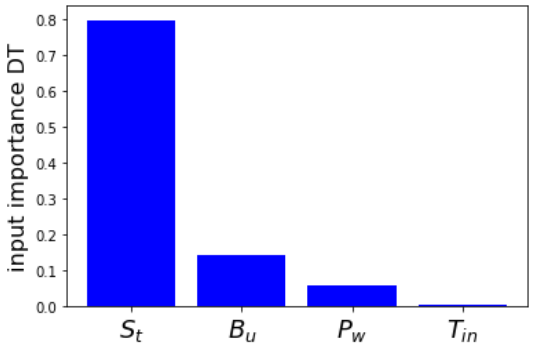
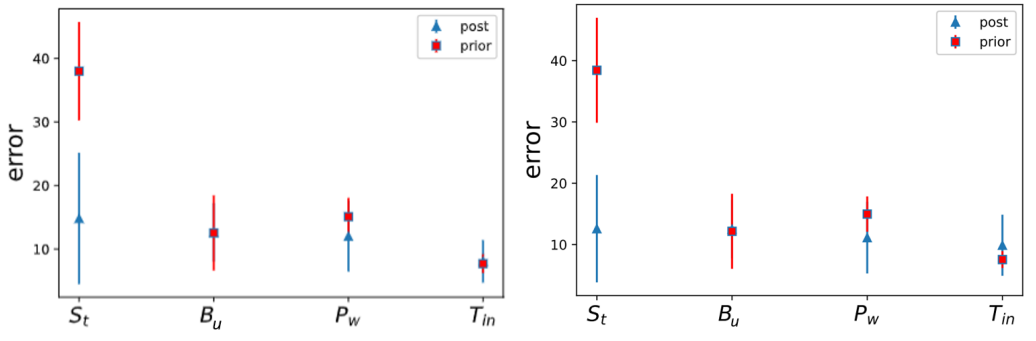


Figure 7: Importance rate (*gini* coefficient) of input variables in the DT forward algorithm.



(a) DT (b) KNN

Figure 8: Prior and Post error of parameter estimation with 3% of observation error.

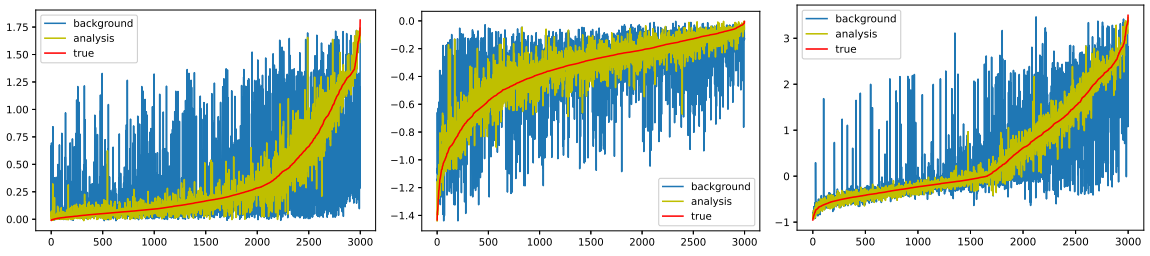
Table 3
Averaged computational time for offline training and online prediction for different ML methods. The dimension of the latent space is fixed to be $n = 30$.

offline training			online evaluation				
SVD AE	KNN	DT	KNN-based	DT-based	GLA (KNN)	GLA (DT)	CORCA-3D
542.07s	0.01s	0.14s	0.076s	0.079s	0.885s	0.923s	$\sim$ 40s

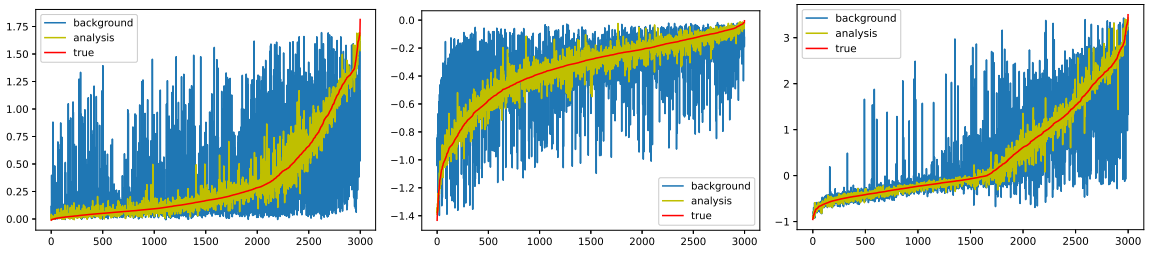
engineering point of view, once the observation is more accurate than the background, the inverse model we proposed is able to provide improved results than just only using the background information.

4.4. Computational performance

The computational work in this paper, including offline training and online predictions are carried out with Intel CPUs (2.30GHz and 26.75Gb RAM) of Google Colab environment. In fact, thanks to the dimension reduction of SVD in the first stage, the training of SVD AE doesn't require the use of GPU which is an important advantage regarding other deep learning (DL)-based AEs. The high-fidelity snapshots are computed with CORCA-3D code package at an industry server at NPIC. The averaged computational time for offline training and online prediction for different ML methods are presented in Table 3. The dimension of the latent space is fixed to be $n = 30$. Particularly, in the online phase, both forward and inverse models can be solved in one second, thus permits the proposed reactor physics operational digital twin to be in real time, which is of practical value from engineering point of view.

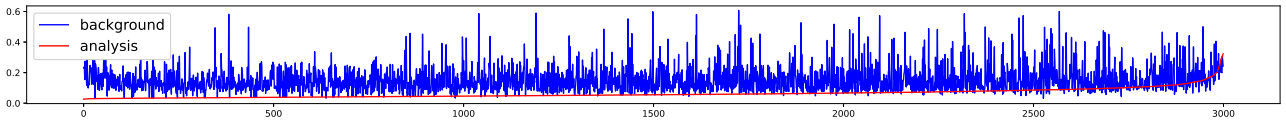


(a) DT LV1 (b) DT LV2 (c) DT LV3

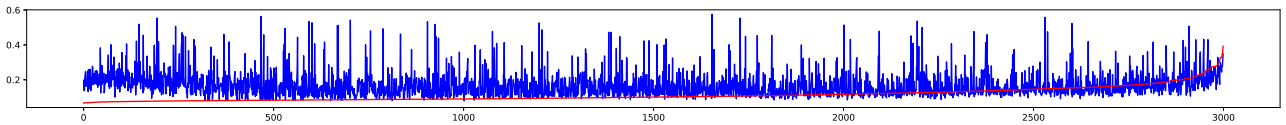


(d) KNN LV1 (e) KNN LV2 (f) KNN LV1

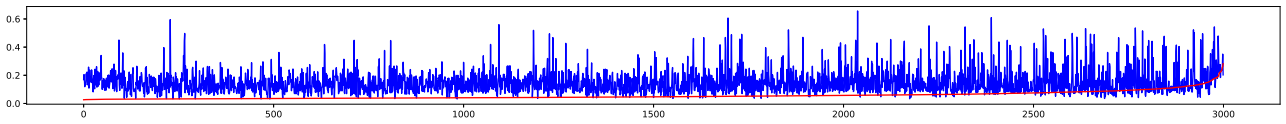
Figure 9: Prediction results of latent variables with background (blue) and assimilated (yellow) parameters in the test data set. The observation error level is set to be 3% for both DT and KNN methods.



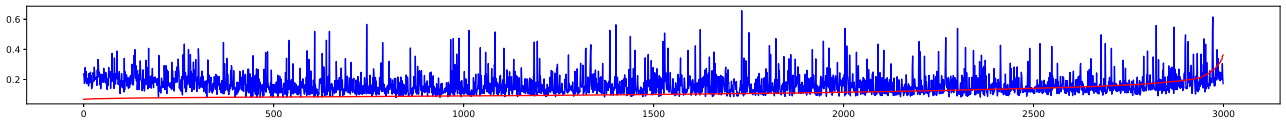
(a) DT observation noise 4%



(b) DT observation noise 8%

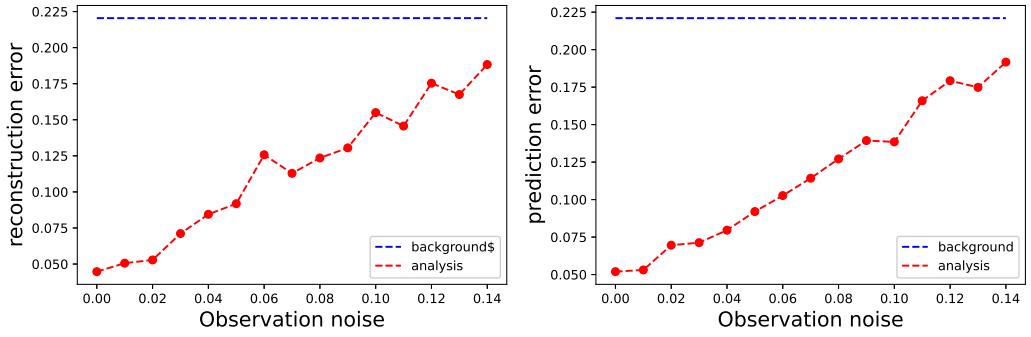


(c) KNN observation noise 4%



(d) KNN observation noise 8%

Figure 10: Relative prediction error in the full physical space using background and assimilated parameters on the test dataset of 3000 samples.



(a) KNN (b) DT

Figure 11: Relative prediction error against the observation noise level.

5. Conclusion

In this paper, a reactor physics operational digital twin is proposed (developed) to predict the power distribution of a nuclear reactor core. The operational digital twin is designed to solve forward problems given operational input parameters, as well as to solve inverse problems given some extra measurements of the power field. Our main contribution is to implement a SVD AE reduced order model to build a non-intrusive forward model with machine learning methods such as KNN and DT, which has a better performance than using general reduced order method such as POD. The second contribution is to apply the generalised latent assimilation method to build the inverse model of the operational digital twin. Benefits from the rapid response of the machine learning based SVD AE forward model, which is far less than solving the forward problem using an industrial code directly. The digital twin is able to evaluate the forward model many times in the iteration process of the GLA based inverse model, thus helps the inverse model to be solved in one second, to realise the real-time property of the operational digital twin.

The operational digital twin is tested through a real engineering problem in nuclear reactor physics i.e., reactor core simulation in the life cycle of HPR1000. Numerical results proved the efficiency and accuracy of the forward/inverse model, even with noisy observations, it is still acceptable from the engineering point view. This work is valuable because of its systematical structure, easy implementable and satisfactory accuracy and time cost for the prediction of operational input parameters and the related power field. Thus it successfully supports the implementation of real engineering scale operational digital twin of nuclear power plants. The framework proposed is general, it can be also adapted to other industry domain to build the operational digital twin. The next work is to bring out the uncertainty qualification of the operational digital twin systematically, to push forward the real engineering applications.

Contribution statement

Helin Gong: Conceptualization, Methodology, Nuclear engineering data curation, Writing & Editing, Funding acquisition. Sibo Cheng: Methodology, Machine learning and data assimilation computation, Writing & Editing. Zhang Chen: Conceptualization, Nuclear engineering data curation, Funding acquisition. Qing Li: Supervision, Review. César Quilodrán Casas: Machine learning methodology, Review. Dunhui Xiao: Supervision, Validation. Rossella Arcucci: Supervision, Review, Funding acquisition.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Grant No. 11905216, Grant No. 12175220). This research is partially funded by the Leverhulme Centre for Wildfires, Environment and Society through the Leverhulme Trust, Grant No. RC-2018-023. This work is partially supported by the EP/T000414/1 PREdictive Modelling with QuantIfication of UncERtainty for MultiphasE Systems (PREMIERE).

Acronyms

RPODT reactor physics operational digital twin
UQ uncertainty quantification
ML machine learning
PWR Pressure Water Reactor
ANL Argonne National Lab
NPP Nuclear Power Plant
PDEs Partial Differential Equations
MOC method of characteristics
ROM reduced-order modelling
HPC High-Performance Computing
POD proper orthogonal decomposition
LHS Latin hypercube sampling
SPNDs self powered neutron detectors
NPIC Nuclear Power Institute of China
AE autoencoder
DA data assimilation
3Dvar Three-dimensional variational data assimilation
PR polynomial regression
SVD singular value decomposition
LHS Latin Hypercube Sampling
DL deep learning
KNN K-nearest neighbours
DT Decision tree
GLA Generalised Latent Assimilation

References

[1] Maddalena Amendola, Rossella Arcucci, Laetitia Mottet, Cesar Quilodran Casas, Shiwei Fan, Christopher Pain, Paul Linden, and Yi-Ke Guo. Data assimilation in the latent space of a neural network. *arXiv preprint arXiv:2012.12056*, 2020.

[2] P. An, Y. Ma, P. Xiao, F. Guo, W. Lu, and X. Chai. Development and validation of reactor nuclear design code CORCA-3D. *Nuclear Engineering and Technology*, 51:1721–1728, 2019.

[3] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. *Contemporary mathematics*, 280:193–220, 2001.

[4] J.-P. Argaud, B. Bouriquet, F. de Caso, H. Gong, Y. Maday, and O. Mula. Sensor placement in nuclear reactors based on the Generalized Empirical Interpolation Method. *Journal of Computational Physics*, 363(2018):354–370, 2018.

[5] Simon Arridge, Peter Maass, Ozan Öktem, and Carola-Bibiane Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.

[6] Mark Asch, Marc Bocquet, and Maëlle Nodet. *Data assimilation: methods, algorithms, and applications*. SIAM, 2016.

[7] C. Audouze, F. De Vuyst, and P. B. Nair. Reduced-order modeling of parameterized PDEs using time-space-parameter principal component analysis. *International Journal for Numerical Methods in Engineering*, 80(8), 2009.

[8] Christophe Audouze, Florian De Vuyst, and Prasanth B. Nair. Nonintrusive reduced-order modeling of parametrized time-dependent partial differential equations. *Numerical Methods for Partial Differential Equations*, 29(5):1587–1628, 2013.

[9] T. Bahadir and SÖ Lindahl. Studsvik's next generation nodal code simulate-5. In *In Proceedings of the Advances in Nuclear Fuel Management IV (ANFM 2009), Hilton Head Island, SC, USA, 12–15 April 2009*, 2009.

[10] Baier, Silvio, Gommlich, Andre, Kozmenkov, Yaroslav, Bilodid, Yuri, Holt, and Lars and. The reactor dynamics code dyn3d-models, validation and applications. *Progress in nuclear engergy*, 89(May):170–190, 2016.

[11] Mario Bebendorf, Yvon Maday, and Benjamin Stamm. Comparison of some reduced representation approximations. In *Reduced Order Methods for Modeling and Computational Reduction*, pages 67–100. Springer, 2014.

[12] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review*, 57(4):483–531, 2015.

[13] Peter Benner, Pawan Goyal, Boris Kramer, Benjamin Peherstorfer, and Karen Willcox. Operator inference for non-intrusive model reduction of systems with non-polynomial nonlinear terms. *Computer Methods in Applied Mechanics and Engineering*, 372:113433, 2020.

[14] William A.1 Boyd, Larry T.1 Mayhue, Vincent S.1 Penkrot, and Baocheng1 Zhang. The whitestar development project: Westinghouse's next generation core design simulator and core monitoring software to power the nuclear renaissance. *American Nuclear Society - International Conference on Mathematics, Computational Methods and Reacto*, 2009.

[15] K. Bradley. Neams update. quarterly report for october - december 2011. *Office of Scientific and Technical Information Technical Reports*, 2012.

[16] Caterina Buizza, César Quilodrán Casas, Philip Nadler, Julian Mack, Stefano Marrone, Zainab Titus, Clémence Le Cornec, Evelyn Heylen, Tolga Dur, Luis Baca Ruiz, et al. Data learning: Integrating data assimilation and machine learning. *Journal of Computational Science*, 58: 101525, 2022.

[17] A. Calloo, D. Couyras, F. Févotte, and M. Guillo. Cocagne: EDF new neutronic core code for ANDROMEDE calculation chain. In *Proceedings of International Conference on Mathematics & Computational Methods Applied to Nuclear Science & Engineering (M&C), Jeju, Korea*, 2017.

[18] César Quilodrán Casas, Rossella Arcucci, Pin Wu, Christopher Pain, and Yi-Ke Guo. A reduced order deep data assimilation model. *Physica D: Nonlinear Phenomena*, 412:132615, 2020.

[19] R. Chakir, Y. Maday, and P. Parnaudeau. A non-intrusive reduced basis approach for parametrized heat transfer problems. *Journal of Computational Physics*, 376:617 – 633, 2019. ISSN 0021-9991.

[20] Christian Chauliac, José-Maria Aragonés, Dominique Bestion, Dan Gabriel Cacuci, Nicolas Crouzet, Frank-Peter Weiss, and Martin A. Zimmermann. Nuresim – a european simulation platform for nuclear reactor safety: Multi-scale and multi-physics calculations, sensitivity and uncertainty analysis. *Nuclear Engineering and Design*, 241(9):3416–3426, 2011. ISSN 0029-5493. doi: https://doi.org/10.1016/j.nucengdes.2010.09.040. URL https://www.sciencedirect.com/science/article/pii/S0029549311001713. Seventh European Commission conference on Euratom research and training in reactor systems (Fission Safety 2009).

[21] Yubao Chen. Integrated and intelligent manufacturing: Perspectives and enablers. *Engineering*, 3(5):588–595, 2017.

[22] S. Cheng, J.-P. Argaud, B. Iooss, D. Lucor, and A. Ponçot. Background error covariance iterative updating with invariant observation measures for data assimilation. *Stochastic Environmental Research and Risk Assessment*, 33(11):2033–2051, 2019.

[23] Sibo Cheng, Jean-Philippe Argaud, Bertrand Iooss, Didier Lucor, and Angélique Ponçot. Error covariance tuning in variational data assimilation: application to an operating hydrological model. *Stochastic Environmental Research and Risk Assessment*, 35(5):1019–1038, 2021.

[24] Sibo Cheng, Didier Lucor, and Jean-Philippe Argaud. Observation data compression for variational assimilation of dynamical systems. *Journal of Computational Science*, page 101405, 2021.

[25] Sibo Cheng, Jianhua Chen, Charitos Anastasiou, Panagiota Angeli, Omar K Matar, Yi-Ke Guo, Christopher C Pain, and Rossella Arcucci. Generalised latent assimilation in heterogeneous reduced spaces with machine learning surrogate models. *arXiv preprint arXiv:2204.03497*, 2022.

[26] Sibo Cheng, Yufang Jin, Sandy P Harrison, César Quilodrán-Casas, Iain Colin Prentice, Yi-Ke Guo, and Rossella Arcucci. Parameter flexible wildfire prediction using machine learning techniques: Forward and inverse modelling. *Remote Sensing*, 14(13):3228, 2022.

[27] Sibo Cheng, I Colin Prentice, Yuhan Huang, Yufang Jin, Yi-Ke Guo, and Rossella Arcucci. Data-driven surrogate model with latent data assimilation: Application to wildfire forecasting. *Journal of Computational Physics*, page 111302, 2022.

[28] F. Chinesta, A. Huerta, G. Rozza, and K. Willcox. Model Order Reduction: a survey. Wiley, 2016.

[29] Albert Cohen, Wolfgang Dahmen, and Ron DeVore. State estimation—the role of reduced models. In Tomás Chacón Rebollo, Rosa Donat, and Inmaculada Higueras, editors, *Recent Advances in Industrial and Applied Mathematics*, pages 57–77, Cham, 2022. Springer International Publishing. ISBN 978-3-030-86236-7.

[30] T. J. Downar, D. A. Barber, R. M. Miller, C. H. Lee, and A. P. Ulses. Parcs: Purdue advanced reactor core simulator. *in: Proceeding of the International Meeting on New Frontiers of Nuclear Technology: Reactor Physics, Safety and High-Performance Computing (PHYSOR 2002), Seoul, South-Korea*.

[31] Humberto E Garcia, Steven E Aumeier, and Ahmad Y Al-Rashdan. Integrated state awareness through secure embedded intelligence in nuclear systems: Opportunities and implications. *Nuclear Science and Engineering*, 194(4):249–269, 2020.

[32] Humberto E Garcia, Steven E Aumeier, Ahmad Y Al-Rashdan, and Bri L Rolston. Secure embedded intelligence in nuclear systems: Framework and methods. *Annals of Nuclear Energy*, 140:107261, 2020.

[33] Moritz Geist, Philipp Petersen, Mones Raslan, Reinhold Schneider, and Gitta Kutyniok. Numerical solution of the parametric diffusion equation by deep neural networks. *Journal of Scientific Computing*, 88(1):22, Jun 2021. ISSN 1573-7691. doi: 10.1007/s10915-021-01532-w. URL https://doi.org/10.1007/s10915-021-01532-w.

[34] Helin Gong, Yingrui Yu, and Qing Li. Reactor power distribution detection and estimation via a stabilized gappy proper orthogonal decomposition method. *Nuclear Engineering and Design*, 370:110833, 2020. ISSN 0029-5493.

[35] Helin Gong, Zhang Chen, Yvon Maday, and Qing Li. Optimal and fast field reconstruction with reduced basis and limited observations: Application to reactor core online monitoring. *Nuclear Engineering and Design*, 377:111113, 2021. ISSN 0029-5493. doi: https://doi.org/10.1016/j.nucengdes.2021.111113. URL https://www.sciencedirect.com/science/article/pii/S0029549321000650.

[36] Helin Gong, Sibo Cheng, Zhang Chen, and Qing Li. Data-enabled physics-informed machine learning for reduced-order modeling digital twin: Application to nuclear reactor physics. *Nuclear Science and Engineering*, 0(0):1–26, 2022. doi: 10.1080/00295639.2021.2014752. URL https://www.tandfonline.com/doi/abs/10.1080/00295639.2021.2014752.

[37] M. W. Grieves. *Virtually Intelligent Product Systems: Digital and Physical Twins*. Complex Systems Engineering: Theory and Practice, 2019.

[38] Michael Grieves and John Vickers. *Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems*, pages 85–113. Springer International Publishing, Cham, 2017. ISBN 978-3-319-38756-7. doi: 10.1007/978-3-319-38756-7_4. URL https://doi.org/10.1007/978-3-319-38756-7_4.

[39] A. Hébert and D Sekki. Chambon rjÉpdmmq, canada, tech. rep. ige-300. a user guide for donjon version4, 2013. 2009.

[40] Alain Hébert. *Applied reactor physics*. Presses inter Polytechnique, 2009.

[41] Gong Helin. *Data assimilation with reduced basis and noisy measurement: Applications to nuclear reactor cores*. PhD thesis, Sorbonne University, 2018.

[42] J.S. Hesthaven and S. Ubbiali. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363:55 – 78, 2018. ISSN 0021-9991.

[43] Elena Jharko. Digital twin of npps: Simulation systems and verification. In *2021 International Russian Automation Conference (RusAutoCon)*, pages 852–857. IEEE, 2021.

[44] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29:36–52, 2020.

[45] K. Kashima. Nonlinear model reduction by deep autoencoder of noise response data. In *In Proceedings of the 2016 IEEE 55th Conference on Decision and Control (CDC), Las Vegas; 2016:5750-5755.*, 2016.

[46] Hamid Khayyam, Gelayol Golkarnarenji, and Reza N Jazar. Limited data modelling approaches for engineering applications. In *Nonlinear Approaches in Engineering Applications*, pages 345–379. Springer, 2018.

[47] Brendan Kochunas and Xun Huan. Digital twin concepts with uncertainty for nuclear power applications. *Energies*, 14(14):4235, 2021.

[48] Brendan Kochunas, Benjamin Collins, Shane Stimpson, Robert Salko, Daniel Jabaay, Aaron Graham, Yuxuan Liu, Kang Seog Kim, William Wieselquist, Andrew Godfrey, Kevin Clarno, Scott Palmtag, Thomas Downar, and Jess Gehin. Vera core simulator methodology for pressurized water reactor cycle depletion. *Nuclear Science and Engineering*, 185(1):217–231, 2017. doi: 10.13182/NSE16-39. URL https://doi.org/10.13182/NSE16-39.

[49] Elmer E Lewis. *Fundamentals of nuclear reactor physics*. Elsevier, 2008.

[50] X. Li, Q. Liu, Q. Li, L. Chen, X. Liu, S. Wang, Y. Xie, and Z. Chen. 177 Core Nuclear Design for HPR1000. *Nuclear Power Engineering*, 40 S1:8–12, 2019.

[51] Linyu Lin, Han Bao, and Nam Dinh. Uncertainty quantification and software risk analysis for digital twins in the nearly autonomous management and control systems: A review. *Annals of Nuclear Energy*, 160:108362, 2021.

[52] C. Liu, R. Fu, D. Xiao, R. Stefanescu, P. Sharma, C. Zhu, S. Sun, and C. Wang. Enkf data-driven reduced order assimilation system. *Engineering Analysis with Boundary Elements*, 139:46–55, 2022. ISSN 0955-7997. doi: https://doi.org/10.1016/j.enganabound.2022.02.016. URL https://www.sciencedirect.com/science/article/pii/S0955799722000510.

[53] Chao Lu, Jiafei Lyu, Liming Zhang, Aicheng Gong, Yipeng Fan, Jiangpeng Yan, and Xiu Li. Nuclear power plants with artificial intelligence in industry 4.0 era: Top-level design and current applications—a systemic review. *IEEE Access*, 8:194315–194332, 2020.

[54] Hugo FS Lui and William R Wolf. Construction of reduced-order models for fluid flows using deep feedforward neural networks. *Journal of Fluid Mechanics*, 872:963–994, 2019.

[55] H Ly and H Tran. Modeling and control of physical processes using proper orthog- onal decomposition. *J Math Comput Model*, 33:223–236, 2001.

[56] Y. Maday. Reduced basis method for the rapid and reliable solution of partial differential equations. In *in International Congress of Mathematicians. Vol. III, 1255–1270, Eur. Math. Soc., Zürich*. Citeseer, 2006.

[57] Laura Mainini and Karen E. Willcox. Surrogate modeling approach to support real-time structural assessment and decision making. *AIAA J*, 53(6):1612–1626, 2015.

[58] Subhasish Mohanty and Joseph Listwan. Development of digital twin predictive model for pwr components: Updates on multi times series temperature prediction using recurrent neural network, dmw fatigue tests, system level thermal-mechanical-stress analysis. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States), 2021.

[59] Subhasish Mohanty and Richard B Vilim. Physics-infused ai/ml based digital-twin framework for flow-induced-vibration damage prediction in a nuclear reactor heat exchanger. Technical report, Argonne National Lab.(ANL), Argonne, IL (United States), 2021.

[60] Benjamin Peherstorfer and Karen Willcox. Data-driven operator inference for nonintrusive projection-based model reduction. *Computer Methods in Applied Mechanics and Engineering*, 306:196–215, 2016.

[61] Mathis Peyron, Anthony Fillion, Selime Gürol, Victor Marchais, Serge Gratton, Pierre Boudier, and Gael Goret. Latent space data assimilation by using deep learning. *Quarterly Journal of the Royal Meteorological Society*, 147(740):3759–3777, 2021. doi: https://doi.org/10.1002/qj.4153. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4153.

[62] Toby R. F. Phillips, Claire E. Heaney, Paul N. Smith, and Christopher C. Pain. An autoencoder-based reduced-order model for eigenvalue problems with application to neutron diffusion. *International Journal for Numerical Methods in Engineering*, 122(15):3780–3811, 2021. doi: https://doi.org/10.1002/nme.6681. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6681.

[63] César Quilodrán-Casas, Rossella Arcucci, Laetitia Mottet, Yike Guo, and Christopher Pain. Adversarial autoencoders and adversarial LSTM for improved forecasts of urban air pollution simulations. *arXiv preprint arXiv:2104.06297*, 2021.

[64] Adil Rasheed, Omer San, and Trond Kvamsdal. Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access*, 8: 21980–22012, 2020. doi: 10.1109/ACCESS.2020.2970143.

[65] Christopher Ritter, Ross Hays, Jeren Browning, Ryan Stewart, Samuel Bays, Gustavo Reyes, Mark Schanfein, Adam Pluth, Piyush Sabharwall, Ross Kunz, et al. Digital twin to detect nuclear proliferation: A case study. *Journal of Energy Resources Technology*, 144(10):102108, 2022.

[66] Michael Schluse, Marc Priggemeyer, Linus Atorf, and Juergen Rossmann. Experimentable digital twins—streamlining simulation-based systems engineering for industry 4.0. *IEEE Transactions on industrial informatics*, 14(4):1722–1731, 2018.

[67] Lawrence Sirovich. Turbulence and the dynamics of coherent structures. II. Symmetries and transformations. *Quarterly of Applied mathematics*, 45(3):573–582, 1987.

[68] M. Smith, E. Lewis, and Shemon E. Pwr core tracking using a next-generation core calculation code, scope2. *DIF3D-VARIANT 11.0: A Decade of Updates, Argonne National Lab. (ANL), Argonne, IL, 2014.*, 2014.

[69] Weston M Stacey. *Nuclear reactor physics*. John Wiley & Sons, 2018.

[70] R. Szilard, H. Zhang, D. Kothe, and P. Turinsky. The consortium for advanced simulation of light water reactors. *American Physical Society*, 2011.

[71] Fei Tao, Qinglin Qi, Lihui Wang, and AYC Nee. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering*, 5(4):653–661, 2019.

[72] M. Tatsumi, A. Yamamoto, H. Nagano, and K. Sengoku. Pwr core tracking using a next-generation core calculation code, scope2. *in: Proceedings of the International Conference Global Environment and Advanced Nuclear Power (GENES4/ANP2003), Paper*, 2003.

[73] Christophe Varé and Patrick Morilhat. Digital twins, a new step for long term operation of nuclear power plants. In *Engineering Assets and Public Infrastructures in the Age of Digitalization*, pages 96–103. Springer, 2020.

[74] Jun Wang, Xin Li, Chris Allison, and Judy Hohorst. *Nuclear Power Plant Design and Analysis Codes: Development, Validation, and Application*. Woodhead Publishing, 2020.

[75] D. Xiao, F. Fang, C.C. Pain, and I.M. Navon. A parameterized non-intrusive reduced order model and error analysis for general time-dependent nonlinear partial differential equations and its applications. *Computer Methods in Applied Mechanics and Engineering*, 317:868–889, 2017. ISSN 0045-7825.

[76] Dunhui Xiao, Pan Yang, Fangxin Fang, Jiansheng Xiang, Chris C Pain, and Ionel M Navon. Non-intrusive reduced order modelling of fluid–structure interactions. *Computer Methods in Applied Mechanics and Engineering*, 303:35–54, 2016.

[77] Wen Yang, Hongchun Wu, Yunzhao Li, Jiewei Yang, and Liangzhi Cao. Development and verification of pwr-core fuel management calculation code system necp-bamboo: Part ii bamboo-core. *Nuclear Engineering and Design*, 337:279–290, 2018. ISSN 0029-5493. doi: https://doi.org/10.1016/j.nucengdes.2018.07.017. URL https://www.sciencedirect.com/science/article/pii/S002954931830791X.