NENE
International Conference
Nuclear Energy for New Europe

Multi-Physics Model Correction With Data-Driven Reduced Order Modelling

Stefano Riva, Carolina Introini, Antonio Cammi
Politecnico di Milano, Dept. of Energy,
CeSNEF-Nuclear Engineering Division, Nuclear Reactors Group
Via La Masa, 34
20156 Milano, Italy
stefano.riva@polimi.it, carolina.introini@polimi.it, antonio.cammi@polimi.it

ABSTRACT

Nowadays, the state-of-the-art approach in numerical modelling for nuclear reactors is represented by the multi-physics (MP) analysis. This framework enables the investigation of the inter-dependency between different physics characterising a reactor (e.g., neutronics and thermal hydraulics) for a deeper understanding of the phenomena occurring in the system. The coupling can occur in two ways: by developing interfaces between single-physics codes (e.g., Serpent for neutronics and OpenFOAM for thermal-hydraulics) or gathering every physics inside a single environment. The latter path has become state-of-the-art in the nuclear field; however, this approach loses all the previous validations of the single-physics codes; moreover, the computational resources for such complex numerical models are still demanding. Data-driven reduced Order Modelling can play a crucial role by shifting the coupling to the reduced level rather than to the full-order model, thus keeping using the already-validated and widely-used single-physics codes: the data collected from the physical system intrinsically contain multi-physics information and thus, can be used eventually to correct the physics not considered by the model. This work applies this novel approach to a 2D coupled neutronic-thermal case study based on the PWR geometry in the Argonne National Laboratory (ANL) benchmarks; the obtained results are promising in showing the reliability and efficiency of the proposed method and paving the way for a more in-depth investigation in more complex scenarios.

1 INTRODUCTION

For decades, many efforts have been made to develop accurate models for the complete physical description of nuclear reactors and to use them to make predictions and design components. Nuclear systems involve different physics, and coupling feedback effects play a crucial role in the reactor, thus making their mathematical modelling challenging. For years, the state-of-the-art approach in nuclear reactor modelling has been using accurate single physics models (e.g., thermal hydraulics, neutronics or thermo-mechanics), introducing the coupling effects through correlations, typically calibrated from experiments. This choice was driven by the fact that mathematical models, represented by Partial Differential Equations (PDEs), require a numerical solution using computer codes, and the computational resources are usually limited. Thus, a single physics approach was chosen.

However, in recent years the available computational power has increased a lot, making heavier simulations possible; hence, the modelling of nuclear reactors is shifting to Multi-Physics

(MP) modelling [1], i.e., creating a more detailed description of the system itself by coupling different physics together. The numerical solution of these Full Order Models (FOMs) is rather challenging and still requires powerful computers with non-negligible computational time, which can be again critical if multiple evaluations are required (*multi-query* scenarios) or if real-time analyses are needed. In these contexts, using Reduced Order Modelling (ROM) techniques [2] is necessary, as they allow for keeping a desired level of accuracy with reasonable computational times.

Furthermore, in recent years, modelling from data has also become quite popular [3], and the possibility of combining data with some *background* mathematical knowledge has gained interest in several fields within the so-called Data Assimilation (DA) framework [4]. This framework has encountered much interest because it overcomes the main limitations of both models and data: in the former, mathematical models are derived under some assumptions and are affected by uncertainty on the parameters; on the other hand, data may be polluted by random or systematic noise, however, as they are evaluations from the physical system they may contain information on un-modelled physics or non-parametric uncertainty [5]. The typical mathematical setting of DA usually requires multiple evaluations of mathematical models, and hence using ROM methods is required; to this aim, data-driven ROM techniques have been developed, such as the Gappy-POD [6] or the Generalised Empirical Interpolation Method (GEIM) [7]. These methods can be theoretically described within the framework of a Parameterised-Background Data-Weak (PBDW) formulation [5], providing a mathematical base for the use of data to update or correct the prediction of the mathematical models.

This idea can be exploited in the framework of MP modelling: in fact, the single-physics approach could still be used to generate a proper representation in the reduced coordinate system, then real data, which inherently contains MP information, as they come from the real system, are used as an update to introduce the un-modelled physics. In this way, all the previous experience and validation performed on state-of-the-art single-physics codes can be maintained; moreover, single-physics codes are less computationally expensive than the numerical solution of full MP models. Furthermore, in addition to the novel implementation of MP models, brand-new codes must pass a deep phase of verification and validation before being used, for example, for safety-related applications.

This work aims at analysing this possibility using GEIM and PBDW methods to correct inaccurate single-physics models using data, adopting as a benchmark a 2D coupled neutronic-thermal case study based on the PWR geometry in the Argonne National Laboratory (ANL) [8]. The paper is organised as follows: in Section 2 a brief discussion of the data-driven ROM methods used is given, including a presentation of model correction for MP problems; Section 3 is devoted to the numerical test case and the results; in the end, the main conclusions are drawn in Section 4.

2 DATA-DRIVEN REDUCED ORDER MODELLING

All model order reduction techniques are characterised by an offline-online decomposition [2]: firstly, the FOM is solved several times to generate a set of training solutions (called **snapshots** characterised by $\mathcal{N}_h$ degrees of freedom), that are used to generate a set of basis functions (which represent the spatial dependence of the solutions and their physical interpretation may vary according to the adopted algorithm), which spans a reduced space onto which the model is represented; then, in the online phase, a Reduced Order Model[1] is solved repetitively in a fast and accurate way thanks to its much lower complexity. This framework is referred to as the reduced basis approach, in which the state of the system u (i.e., the solution of the FOM)

[1]This can be a "small" system of Ordinary Differential Equations or even a linear system, depending on the adopted technique. Regardless, their dimension M is much lower than $\mathcal{N}_h$.

is approximated with a linear combination of a set of basis function $\{\psi_n\}_{n=1}^N$, i.e.

$$u(\mathbf{x};\boldsymbol{\mu}) \simeq \sum_{n=1}^{N} \alpha_n(\boldsymbol{\mu}) \cdot \psi_n(\mathbf{x}), \qquad \mathbf{x} \in \Omega \subset \mathbb{R}^d \text{ and } \boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^p \tag{1}$$

given $\boldsymbol{\mu}$ the parameters and $\{\alpha_n(\boldsymbol{\mu})\}_{n=1}^N$ the modal or reduced coefficients, weighting the importance of each spatial basis function.

The main focus of this work is on data-driven ROM, a novel set of techniques able to combine the knowledge of the mathematical model and the real evaluations of the system. In this way, compared to standard ROM methods the state prediction is not bounded by the model accuracy, and it can be improved, even during operation, by integrating measurements of the physical fields of interest. These methods add an important step in the offline phase mentioned above, related to the search for the optimal positioning of sensors to maximise the amount of information extracted from the system and thus to better include unanticipated uncertainty and un-modelled physics [5].

2.1 Generalised Empirical Interpolation Method

The Generalised Empirical Interpolation Method (GEIM) is a data-driven ROM technique proposed in [7], based on a greedy algorithm for the selection of basis functions and sensors. In particular, hierarchical spaces are built to approximate a given function u with a suitable interpolant

$$u(\mathbf{x};\boldsymbol{\mu}) \simeq \mathcal{I}_M[u](\mathbf{x};\boldsymbol{\mu}) = \sum_{m=1}^{M} \beta_m(\boldsymbol{\mu}) \cdot q_m(\mathbf{x}) \quad \text{s.t.} \quad \{v_m(u) = v_m\left(\mathcal{I}_M\right)\}_{m=1}^M, \tag{2}$$

where the spatial behaviour is given by the **magic functions** $\{q_m(\mathbf{x}\}_{m=1}^M$ and the parametric dependence by the coefficients $\{\beta_m(\mathbf{x}\}_{m=1}^M$. Each magic function q_m is associated with a **magic sensor** $v_m(\cdot;\mathbf{x}_m,\, s)$, mathematically represented by a linear functional centred in $\mathbf{x}_m \in \Omega$ and with a point-spread[2] $s \in \mathbb{R}^+$, representing the area onto which the sensors "acts" to collect the measure (as sensors have a physical dimension, and thus they don't collect true point-wise information).

During the offline stage, the greedy procedure takes as input the training snapshots and returns a set of magic functions and sensors by minimising at each step the interpolation error. The magic sensors are chosen among a library Υ, representing the set of possible locations in the domain, to maximise the amount of information extracted from the system. Then, the online phase aims at determining the coefficients using collected data $\mathbf{y} \in \mathbb{R}^M$ from the real system:

$$y_m = v_m(u^{\text{true}}(\mathbf{x});\, \mathbf{x}_m,\, s) + \varepsilon_m \qquad \text{with } m = 1, \dots M. \tag{3}$$

For the sake of brevity, in this work data are assumed to be noiseless, i.e. $\varepsilon_m = 0$. Then, the reduced coefficients $\boldsymbol{\beta} \in \mathbb{R}^M$ are the solution of a linear system of dimension $M << \mathcal{N}_h$, arising from the interpolation condition $B\boldsymbol{\beta} = \mathbf{y}$ with $B_{ij} = v_i(q_j)$, given $i, j = 1, \dots M$.

2.2 Parameterised-Background Data-Weak formulation

The Parameterised-Background Data-Weak (PBDW) formulation [5] is based on a least-squares approximation [4], with the important novelty of taking into account unanticipated uncertainty as modelling errors. The state $u(\mathbf{x};\boldsymbol{\mu})$ is approximated through a linear combination

[2] These properties comes from the fact that experimental sensors can be modelled as linear functionals with a Gaussian kernel [9].

of two contributions: the background knowledge z_N and the update η_M

$$u(\mathbf{x};\boldsymbol{\mu}) \simeq z_N(\mathbf{x};\boldsymbol{\mu}) + \eta_M(\mathbf{x};\boldsymbol{\mu}) = \sum_{n=1}^{N} \alpha_n(\boldsymbol{\mu}) \cdot \zeta_n(\mathbf{x}) + \sum_{m=1}^{M} \theta_m(\boldsymbol{\mu}) \cdot g_m(\mathbf{x}), \tag{4}$$

in which $\{\zeta_n\}_{n=1}^N$ is the basis of the reduced space of dimension N approximating the information of the mathematical model and $\{g_m\}_{m=1}^M$ is the basis of the update space, having dimension M, representing un-modelled physics. The reduced space can be built using the most suitable algorithm for the specific problem under consideration: in this work, the Proper Orthogonal Decomposition (POD) [2] is used. On the other hand, the bases of the update space are related to the sensors: in this work, the SGREEDY procedure [5] is used to create the update space, aiming at minimising the reconstruction error of the method in order to select the optimal configuration of the sensors $\{v_m\}_{m=1}^M$. The basis functions g_m are defined as the Riesz representation of the sensor/linear functional v_m (for the mathematical definition, see [5]).

During the online phase, once the data have been collected, the coefficients $\boldsymbol{\alpha} \in \mathbb{R}^N$ and $\boldsymbol{\theta} \in \mathbb{R}^M$ are computed as the solutions of the following linear system of dimension $(N+M)$

$$\begin{bmatrix} A & K \\ K^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \qquad \text{given} \begin{cases} A_{mm'} = (g_m,\, g_{m'})_{L^2(\Omega)} \\ K_{mn} = (g_m,\, \zeta_n)_{L^2(\Omega)} \end{cases} \tag{5}$$

with $m, m' = 1, \dots M$ and $n = 1, \dots, N$.

2.3 Multi-Physics model correction

Both the GEIM and the PBDW are data-driven ROM techniques able to integrate the information of mathematical models and real evaluations on the system: this important feature can be exploited in nuclear reactor modelling as follows. Supposing to have at disposal two different single-physics solvers for a transient problem, one for the neutronics and one for the thermal-hydraulics, two different *high-fidelity* models are built: one with a full and accurate coupling scheme between them acting as the *truth*, labelled as **FOM**, and another with a weaker and inaccurate coupling, labelled as **aFOM** (approximate FOM). The former is implemented using a segregated coupled approach [1], whereas the latter is implemented assuming that the neutronics code cannot directly communicate with the thermal code and that the latent/coupling information is inserted as a boundary exploiting the POD with Interpolation [10]. The idea of aFOM serves to mimic the de-coupling of single physics codes, which has been the state-of-the-art for nuclear reactor modelling for years, in particular, some coupling effects may be missed to be later updated in the online phase.

The basis functions and the sensors, both for GEIM and PBDW, are generated using the snapshots of the aFOM; then, during the online phase, the data coming from the FOM are used to try to correct the inaccuracy contained in the training data (which comes from the approximated FOM with the weaker coupling).

3 NUMERICAL RESULTS

The possibility of correcting Multi-physics models using data-driven ROM techniques following the methodology discussed above is now investigated on a nuclear test case based on the IAEA 2D PWR benchmark [11, 8]. The FOM and aFOM are implemented in Python using the FEniCSx finite element library; moreover, the ROM techniques have been implemented in the *pyforce* Python package (see Section A), which will be made available in the future.

3.1 Description of the test case

As mentioned above, the test case considered in this work is a 2D geometry mimicking the shape of a PWR. Neutronics is modelled using unsteady 2-groups diffusion equations for the fast and thermal flux ϕ_1 and ϕ_2, along with the precursors' equations for $\{c_j\}_{j=1}^6$:

$$\begin{cases} \dfrac{1}{v_1}\dfrac{\partial \phi_1}{\partial t} - \nabla\cdot\left(D_1\nabla\phi_1\right) + \left(\Sigma_{a,1} + \Sigma_{s,1\to 2} + D_1 B_{z,1}^2\right)\phi_1 = \dfrac{1-\beta}{k_{eff}}\nu\Sigma_{f,2}\phi_2 + \displaystyle\sum_{j=1}^{6}\lambda_j c_j \\ \dfrac{1}{v_2}\dfrac{\partial \phi_2}{\partial t} - \nabla\cdot\left(D_2\nabla\phi_2\right) + \left(\Sigma_{a,2} + D_2 B_{z,2}^2\right)\phi_2 = \Sigma_{s,1\to 2}\phi_1 \\ \quad\dfrac{\partial c_j}{\partial t} = \dfrac{\beta_j}{k_{eff}}\nu\Sigma_{f,2}\phi_2 - \lambda_j c_j \qquad\qquad j = 1,\dots 6 \end{cases} \tag{6}$$

with vacuum and symmetry boundary conditions for the fluxes. On the other hand, the temperature T is governed by the unsteady heat equation

$$C\frac{\partial T}{\partial t} = \nabla\cdot\left(k\nabla T\right) + q''' \tag{7}$$

with convective and symmetry boundary conditions.

These equations are coupled through the temperature dependence of the absorption cross-sections and the diffusion coefficients

$$D_g = D_g^{\text{ref}} + \gamma_D\cdot\ln\frac{T}{T^{\text{ref}}} \qquad\qquad \Sigma_{a,g} = \Sigma_{a,g}^{\text{ref}} + \gamma_a\cdot\ln\frac{T}{T^{\text{ref}}} \qquad\qquad \text{with } g = 1,2 \tag{8}$$

given γ_D and γ_a the feedback coefficients, and through the fission power $q''' = P\cdot\Sigma_{f,2}\phi_2$, provided the normalisation constant P.

This set of equations is solved for the aFOM and the FOM to generate the snapshots necessary for the offline and online phases, respectively. The inaccuracy of the aFOM with respect to the FOM is measured in terms of absolute E and relative ε error, in L^2-norm, and the values are reported in Table 1: the neutronics fields are the most inaccurate ones. The training snapshots of the aFOM have been computed until $t_1 = 1\,$s, whereas the FOM snapshots are collected up to $t_2 = 2\,$s[3] to assess the predictive/extrapolation capabilities of GEIM and PBDW. Hence, the role of $\boldsymbol{\mu}$ is played by the pseudo parameter $t\,(\text{s})$.

Table 1: Absolute E and relative ε error (maximum and average) between the FOM and the aFOM, measured in L^2-norm.

Field	$\max_{t\in\Xi_{\text{train}}} E$	$\max_{t\in\Xi_{\text{train}}} \varepsilon$	$\langle E\rangle_{t\in\Xi_{\text{train}}}$	$\langle \varepsilon\rangle_{t\in\Xi_{\text{train}}}$
T	752.93	1.1%	177.55	0.26%
ϕ_1	77.43	19.4%	20.39	4.9%
ϕ_2	21.67	19.0%	5.70	4.8%

3.2 State estimation: model correction and prediction

This work focuses on the direct reconstruction and time prediction of the fluxes ϕ_1 and ϕ_2 and the temperature T through GEIM and PBDW using the training snapshots of the aFOM and the "synthetic measurements" of the FOM. The point spread of the sensors is assumed to be $s = 1$ for all the fields, whereas the available positions are all the cells of the numerical mesh.

[3]The particular transient chosen as the case study has been chosen following the value of the benchmarks [8].

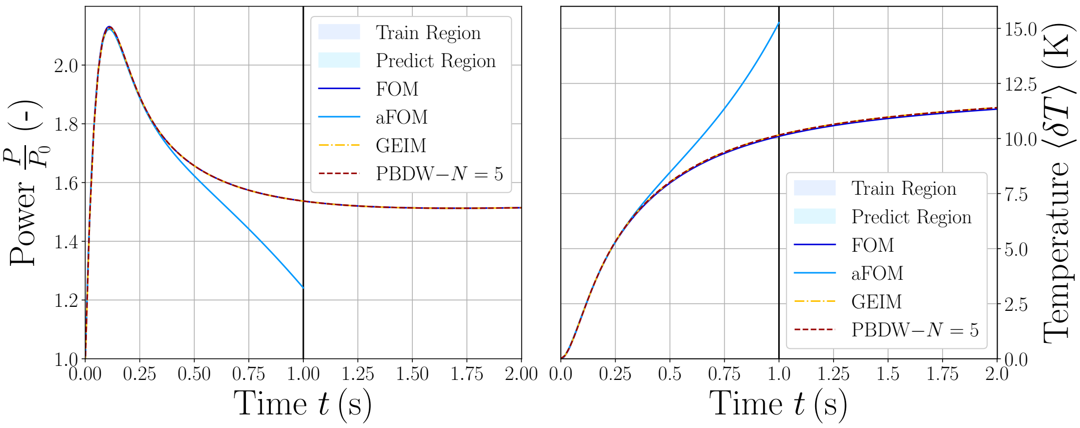

Figure 1: Comparison of the normalised power and average temperature with respect to initial condition for the FOM, aFOM, GEIM and PBDW.

Once the snapshots for the aFOM are obtained from the numerical solver, the offline phase of the GEIM and PBDW can take place. For the PBDW, at first, the POD is performed on each field (ϕ_1, ϕ_2) and T to assess the reducibility of the problem and to generate the basis functions $\{\zeta_n\}_{n=1}^N$ of each field; then, the SGREEDY algorithm is executed to find the sensors positions of the PBDW and thus the basis $\{g_m\}_{m=1}^m$. For the GEIM, the GEIM offline phase is performed on each variable to place the magic sensors $\{v_m\}_{m=1}^M$ and to select the magic functions $\{q_m\}_{m=1}^M$.

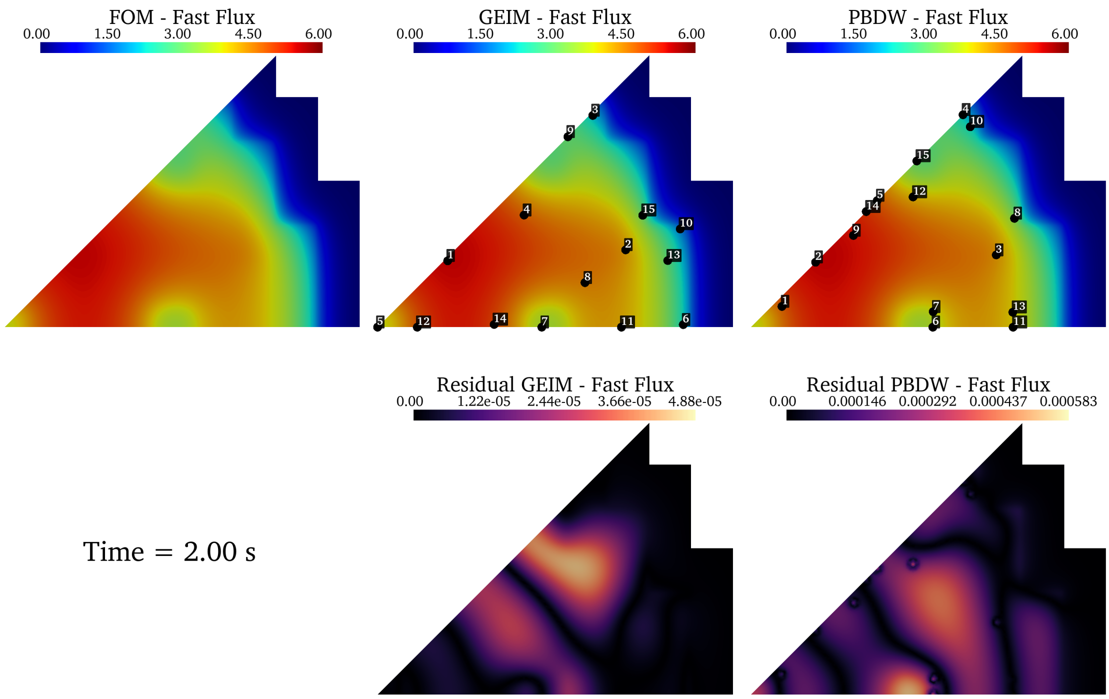

Figure 2: Contour plots of the FOM, GEIM and PBDW (interpolant and residual) at the final time, 2 seconds for the fast flux ϕ_1. The black dots represent the positions of the experimental sensors.

The online phase can finally take place, starting by "collecting the measurements" from the snapshots of the FOM through Equation (3). It has been chosen [4] as the maximum number

[4]These values are taken by looking at the POD eigenvalues and the reconstruction error of the algorithms in the

of available sensors $M = 15$ both for GEIM and PBDW (for the former, this value is equal to the dimension of the reduced space), whereas the dimension of the reduced space of PBDW is set at $N = 5$.

Figure 1 shows the reconstruction of the core power (normalised with respect to the initial condition) and the average temperature difference compared to the initial one using GEIM and PBDW. The aFOM predicts a wrong evolution of both quantities due to the inaccurate coupling between physics, whereas the estimation of GEIM and PBDW is in almost perfect accordance with the *truth* (in this case, the FOM). This result shows that even starting from inaccurate training snapshots, the basis/magic functions contain enough physical information about the coupling phenomena to correctly predict the power and temperature evolution.

In addition to the comparison in terms of power and average temperature, the spatial field reconstructions for the fast flux ϕ_1 of GEIM and PBDW are compared to the *truth* in Figure 2, which also shows the residual field $r[\cdot]$ defined as $r[u](\mathbf{x}) = |u - \mathcal{P}_M u|$, given $\mathcal{P}_M$ the reconstruction operator with M measurements. This field was chosen because it is the one with the highest difference between FOM and aFOM and thus the worst case. The average value of the residual fields is quite low even for the worst reconstructed field, which is proof of the accuracy of the reconstruction by GEIM and PBDW.

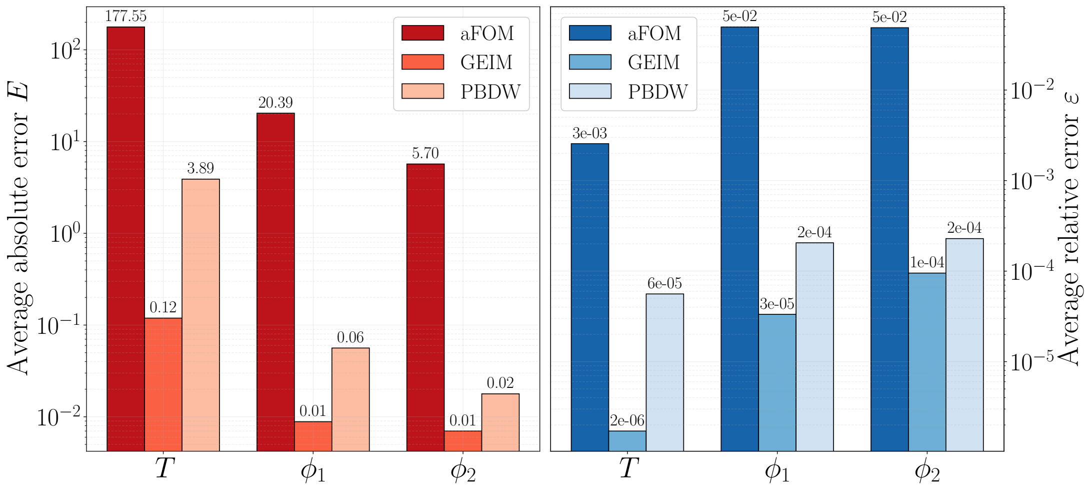

Figure 3: Bar plot of the absolute E and relative ε average error with respect to the *truth* (FOM).

Finally, Figure 3 shows the error between aFOM, GEIM and PBDW compared to the *truth* (i.e., the FOM) measured using the L^2-norm: indeed, the information of the measurements is enough to globally correct the inaccuracy of the background model.

4 CONCLUSIONS

This work analyses the possibility of combining the knowledge of mathematical models and data to update the former, exploiting the new information coming from the latter. In particular, this possibility is studied in the MP framework for nuclear reactors: data-driven ROM techniques can be adopted to save the experience and the validity of single-physics codes: in particular, the prediction of the mathematical model, represented in a reduced coordinate system, is corrected by the data, which contains information on un-modelled physics about the coupling effects between the different variables. In this work, a simple 2D test case is discussed as a preliminary test for this approach, and the results obtained are really promising: in fact, the wrong output of the mathematical model is properly corrected both globally, in terms of power and average temperature, and locally.

offline phase, as in [2, 9], not reported here for sake of brevity.

In the future, further investigations are foreseen considering noisy data, reconstruction of unobservable fields and different modelling strategies for the generation of the snapshots of aFOM.

A DATA AVAILABILITY

The *pyforce* (Python Framework for data-driven model Order Reduction of multi-physiCs problEms) package used in this work is part of the ROSE framework (Reduced Order multi-phySics data-drivEn) developed by the authors. The package will be published on GitHub at https://github.com/ROSE-Polimi/pyforce under the MIT licence.

REFERENCES

[1] C. Demazière. 6 - Neutronic/thermal-hydraulic coupling. In Christophe Demazière, editor, *Modelling of Nuclear Reactor Multi-physics*, pages 311–336. Academic Press, 2020.

[2] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. *Model Order Reduction in Fluid Dynamics: Challenges and Perspectives*, pages 235–273. Springer International Publishing, Cham, 2014.

[3] S. L. Brunton and N. J. Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, USA, 2nd edition, 2022.

[4] A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *WIREs Climate Change*, 9(5):e535, 2018.

[5] Y. Maday, A. Patera, J. Penn, and M. Yano. A parameterized-background data-weak approach to variational data assimilation: formulation, analysis, and application to acoustics. *International Journal for Numerical Methods in Engineering*, 102, 2014.

[6] K. Willcox. Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition. *Computers & Fluids*, 35(2):208–226, 2006.

[7] Y. Maday and O. Mula. A Generalized Empirical Interpolation Method: Application of Reduced Basis Techniques to Data Assimilation. *Springer INdAM Series*, pages 221–235, 2013.

[8] Computational Benchmark Problem Committee for the Mathematics and Computation Division of the American Nuclear Society. Argonne Code Center: Benchmark problem book. Technical report, Argonne National Laboratory, 1977.

[9] C. Introini, S. Cavalleri, S. Lorenzi, S. Riva, and A. Cammi. Stabilization of Generalized Empirical Interpolation Method (GEIM) in presence of noise: A novel approach based on Tikhonov regularization. *Computer Methods in Applied Mechanics and Engineering*, 404:115773, 2023.

[10] N. Demo, M. Tezzele, A. Mola, and G. Rozza. A complete data-driven framework for the efficient solution of parametric shape design and optimisation in naval engineering problems. In *MARINE 2019: VIII International Conference on Computational Methods in Marine Engineering*, 05 2019.

[11] G. Theler, A. Clausse, and F. J. Bonetto. Solution of the 2D IAEA PWR benchmark with the neutronic code MILONGA. In *Actas de la Reunion Anual de la Asociacion Argentina de Tecnologıa Nuclear*, 2011.