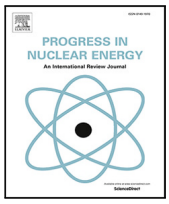
Progress in Nuclear Energy 189 (2025) 105928

Contents lists available at ScienceDirect

Progress in Nuclear Energy

journal homepage: www.elsevier.com/locate/pnucene



# Robust state estimation from partial out-core measurements with Shallow Recurrent Decoder for nuclear reactors

Stefano Riva [a], Carolina Introini [a], Antonio Cammi [a,c,*], J. Nathan Kutz [b]

[a] Politecnico di Milano, Energy Department, CeSNEF - Nuclear Engineering Division, Milan, 20156, Italy
[b] Department of Applied Mathematics and Electrical and Computer Engineering, University of Washington, Seattle, 98195, WA, United States of America
[c] Emirates Nuclear Technology Center (ENTC), Department of Mechanical and Nuclear Engineering, Khalifa University, Abu Dhabi, 127788, United Arab Emirates

ARTICLE INFO

Dataset link: https://github.com/ERMETE-Lab/NuSHRED

*Keywords:*
State reconstruction
Partial observations
Nuclear reactors
Machine learning
SHRED
Monitoring and uncertainty quantification

ABSTRACT

Reliable, real-time state estimation in nuclear reactors is of critical importance for monitoring, control and safety. It further empowers the development of digital twins that are sufficiently accurate for real-world deployment. As nuclear engineering systems are typically characterised by extreme environments, their in-core sensing is a challenging task, even more so in Generation-IV reactor concepts, which feature molten salt or liquid metals as thermal carriers. The emergence of data-driven methods allows for new techniques for accurate and robust estimation of the full state space vector characterising the reactor (mainly composed by neutron fluxes and the thermal-hydraulics fields). These techniques can combine different sources of information, including computational proxy models and local noisy measurements on the system, in order to robustly estimate the state. This work leverages the *Shallow Recurrent Decoder* (SHRED) architecture to estimate the entire state vector of a reactor from three, out-of-core time-series neutron flux measurements alone. Specifically, the Molten Salt Fast Reactor, in the geometry of the EVOL (Evaluation and Viability of Liquid Fuel Fast Reactor System) project, is demonstrated as a test case, with neutron flux measurements alone allowing for reconstruction of the 20 coupled field variables of the dynamics. This approach can further quantify the uncertainty associated with the state estimation due to its considerably low training cost on compressed data. The accurate reconstruction of every characteristic field in real-time makes this approach suitable for monitoring and control purposes in the framework of a reactor digital twin.

## 1. Introduction

To achieve the goals set by the COP28, energy transition must take a central role. In a world that sees an increasing demand for clean and sustainable energy, *new nuclear* technologies present a crucial and complementary solution to renewable energy sources. Nuclear power is a dependable energy source, with high availability and fixed costs. As such, nuclear energy has, jointly with renewable energy, a major role to play in the clean energy transition. Innovative reactor concepts are currently being designed by considering their possible integration with other sources of energy and by following four core tenets as identified by the Generation-IV International Forum (Generation IV International Forum, 2014): (i) efficient use of the fuel, (ii) reduced waste production, (iii) economical competitiveness, and (iv) meeting stringent standards of safety and proliferation resistance.

The development of innovative reactor concepts poses several challenges in terms of design, monitoring and safety. For example, these reactors could, in principle, work with spent fuel from traditional reactors; two out of the six proposed Generation-IV concepts operate with liquid metals as thermal carriers; whereas, one of them, the Molten Salt Fast Reactor (MSFR), adopts a liquid molten salt fuel homogeneously mixed with the coolant (Brovchenko et al., 2013). Moreover, they are usually designed to work in a fast neutron spectrum (Duderstadt and Hamilton, 1976). This condition makes the in-core environment more hostile due to the higher radiation field, with greater damage (and therefore lower useful lifetime) to structural materials and to in-core sensors. To address the challenge of modelling emerging reactor concepts, we leverage a *Shallow Recurrent Decoder* (SHRED) architecture to robustly map out-of-core measurements to full state estimates in the reactor, even fields unmeasured by the sensors. This provides a viable pathway for the construction of a digital twin capable of accurate monitoring of the fission process for safety critical applications.

The production of energy from nuclear reactors has always been subject to rigorous safety criteria at both design and operation levels. This topic has become even more relevant as nuclear reactors are planned to be included in hybrid energy grids along with renewable

* Corresponding author.
*E-mail address:* antonio.cammi@polimi.it (A. Cammi).

https://doi.org/10.1016/j.pnucene.2025.105928
Received 26 November 2024; Received in revised form 27 June 2025; Accepted 7 July 2025
Available online 26 July 2025

sources (Agency, 2023; Ruth et al., 2014). The intermittent nature of these renewable sources require the nuclear reactor to reliably operate in load-following mode, which requires being able to quickly assess the new state of the system following a rapid change in power, whilst ensuring safety constraints at all times. This requires having access in real-time to all the dominant quantities of interest describing the physical behaviour of the system (e.g., temperature, neutron flux, power density and coolant velocity). Fast, accurate and reliable *digital twin* (DT) models of the physical plant (Grieves, 2019) are consequently of growing interest in the nuclear engineering community (Mohanty and Vilim, 2021; Mohanty and Listwan, 2021; Gong et al., 2022a).

A key objective associated with DTs consists of developing an inverse model capable of inferring the full state (characterised by all the different fields of interest) from point (sensor) measurements (Argaud et al., 2018; Gong et al., 2023, 2022b). These inverse models can be constructed by combining mathematical modelling and experimental information to provide a full state estimation in a quick and reliable way. These approaches are critical for real-time control, monitoring of the nuclear reactor, and ensuring safety standards. Furthermore, fast models allow for performing sensitivity analysis and uncertainty quantification in reasonable computational times. Indeed, having a reliable tool that provides state estimation, identifies the onset of potential accidental scenarios, and characterises the unexpected behaviour of experimentally unobservable quantities is essential for the development of efficient and "autonomous" control systems. In fact, as mentioned above, nuclear reactor cores are characterised by a harsh and hostile environment, which makes in-core sensing a non-trivial task. This challenge holds especially true in the MSFR, which, due to the liquid nature of the core, requires the local measurements to be collected from sensors placed outside the core (Cammi et al., 2024; Riva et al., 2024a). Therefore, a DT of a nuclear reactor must be able to provide in real-time a complete estimation of the whole state of the reactor, integrating together the information of models and measurements while also being able to identify and predict accidental scenarios.

The simulation of mathematical models typically requires the numerical solution of parameter-dependent Partial Differential Equations (PDEs), which is not feasible for real-time applications (Kapteyn et al., 2020). Only the recent developments in Reduced Order Modelling (ROM) approaches (Lassila et al., 2014; Rozza et al., 2020) have opened new possibilities towards reliable and efficient DTs for engineering systems. These techniques aim to produce a reduced representation of the solution manifold and provide an approximation that is both sufficiently accurate and obtained with reasonably low computational costs. Among ROM techniques, non-intrusive methods are more suited for the combination of models and measurements. Moreover, with respect to intrusive approaches (Quarteroni et al., 2015), they do not require the knowledge of the governing PDEs. They have been widely used in the literature to combine models with sensors with notable examples including the Gappy Proper Orthogonal Decomposition (Brunton and Kutz, 2022), Generalised Empirical Interpolation Method (Maday et al., 2015) or the Parameterised-Background Data-Weak formulation (Maday et al., 2014).

Data-driven models have also emerged as the leading paradigm for learning input–output maps (Brunton and Kutz, 2022), including mapping input/boundary conditions (Kang et al., 2022; He et al., 2022) or temporal trajectories of measurements (Williams et al., 2024) to the state space, making them applicable for state estimation purposes in nuclear reactors (Gong et al., 2022a,b). Indeed, the non-intrusive techniques mentioned above include a step which determines the optimal sensor placement to obtain a well-conditioned inverse problem upon which state estimates rely (Binev et al., 2018; Manohar et al., 2018). Within the nuclear engineering framework, a two-step approach has been proposed by Introini et al. (2023b) and Gong et al. (2023): the characteristic parameter, producing some local measurements, is estimated with an inverse optimisation problem and then used to calculate the temporal and parametric trajectories embedded in the reduced coefficients. Even though promising results have been obtained even using out-core measurements (Cammi et al., 2024), optimisation problems are often computationally expensive even when solved in a reduced coordinate system; in fact, this was the main computational bottleneck of the work done by the authors in Introini et al. (2023b).

The *SHallow REcurrent Decoder* is a neural network architecture that, following a training and learning process, maps the trajectories of sensor measurements to a latent space, thus encoding the dynamics of the high-dimensional space. In its basic principles, SHRED can be seen as a generalisation of the separation of variables method for solving PDEs (Williams et al., 2024; Ebers et al., 2023; Faraji et al., 2025). This work applies the SHRED architecture to the Molten Salt Fast Reactor considering the EVOL geometry (Brovchenko et al., 2013): in particular, the overall state vector is to be reconstructed using out-core measurements a single field, following Cammi et al. (2024) and Riva et al. (2024a). The system state includes the energy fluxes, the neutron precursors (Duderstadt and Hamilton, 1976), the decay heat groups (Aufiero, 2014) and the thermal-hydraulics fields, i.e., the pressure, the temperature and the velocity; in this work, only one of the energy fluxes is supposed to be observable (this choice is arbitrary and it may depend on the data availability on the plant site), as in general, fluxes measures are easier with respect to others, like the precursors concentration or the velocity. Moreover, the sensors are only placed in the external solid reflector region to avoid in-core placement, an almost impossible task due to the fluid nature of the MSFR. The SHRED architecture is such that the training phase is quite rapid and inexpensive and the required hyperparameter tuning on the net parameters is minimal (Faraji et al., 2025), to the point that it can be easily performed on personal computers (Williams et al., 2024). These features make this architecture well-suited for different physical problems and noisy data: more than one SHRED model can be trained quickly so that different predictions of the temporal dynamics can be assembled to obtain a robust reconstruction, with the additional benefit of providing a mean value of the temporal dynamics and an uncertainty quantification of the estimation itself.

The paper is structured as follows: firstly, in Section 2, the basics of the SHRED architecture are discussed; Section 3 describes the Molten Salt Fast Reactor and the numerical modelling related to the full-order model (i.e., the source of training data); the numerical results are then presented in Section 4; lastly, the main conclusion are drawn in Section 5.

## 2. SHallow REcurrent Decoder

The SHRED (Williams et al., 2024; Ebers et al., 2023; Faraji et al., 2025) architecture is a novel data-driven technique based on the combination of a Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997), a Shallow Decoder Network (SDN) (Erichson et al., 2020) and the compression given by the Singular Value Decomposition (SVD), also known as Proper Orthogonal Decomposition (Brunton and Kutz, 2022; Rozza et al., 2020). The performance of the SHRED architecture was proven in Williams et al. (2024), where given the same physical problems this structure outperformed, in terms of reconstruction accuracy, state-of-the-art linear and non-linear techniques given as low as three sensors. Even more promising, SHRED selects the sensors randomly among the available locations, whereas other non-intrusive methods typically adopt hierarchical greedy algorithms to identify the optimal position of the sensors, a step in the offline phase that often represents one of the computational bottlenecks. Instead, SHRED incorporates the temporal trajectory of the measurements through the LSTM to learn the evolution of the temporal dynamics embedded in the latent space generated by the SVD with the SDN.

Fig. 1 highlights the architecture of SHRED. As a first step, the neural network is trained to map the measures to a compressed space, such as the low-rank approximation given by the SVD (Brunton and

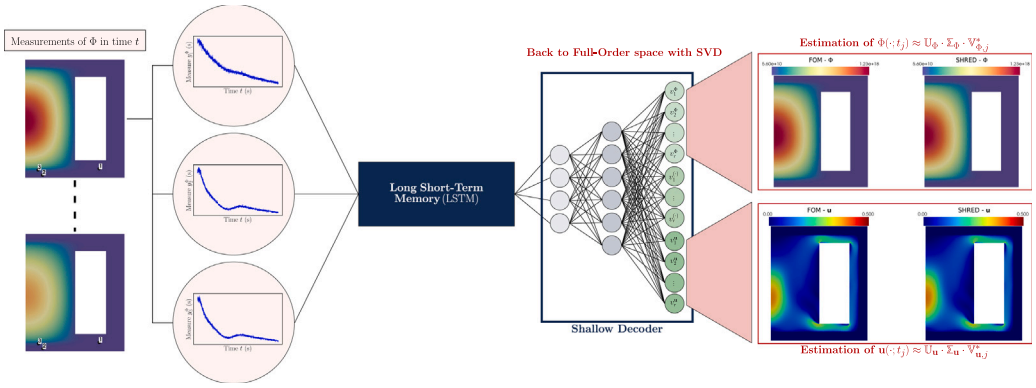



**Fig. 1.** SHRED architecture applied to Nuclear Reactors. Three out-of-core sensors are used to measure a single field variable. The sensor time series is used to construct a latent temporal sequence model which is mapped to compressive representations of all spatio-temporal field variables. The compressive representations can then be mapped to the original state space by the singular value decomposition. The compressive representation allows for laptop level training in minutes even on 20 high-dimensional field variables.

Kutz, 2022). SHRED maps only a few measurements $\mathbf{y}$ of a single observable field to the latent dynamics given by the reduced coefficients $v_n(t)$ of each characteristic quantity of interest (for the specific case of this paper, these quantities are those that describe the physics of the nuclear reactor, such as neutron fluxes, temperature and velocity). This first compression with SVD approximates the full-order space of the snapshots to a low-rank finite dimensional space spanned by the first $r$-rank left singular eigenvectors of the SVD, allowing for a significant reduction in the training cost to the point that the SHRED model can be trained in minutes even on a personal laptop (Faraji et al., 2025). Due to this, it becomes cheap to train different SHRED models for different sensor configurations, i.e., for other random sensor locations, to obtain more than one estimation of the state such that a measure of the uncertainty associated with the prediction can be provided. This way of proceeding can be useful to generate state estimations which are robust against noisy measurements collected on the physical system itself; furthermore, it can be extended to identify possible malfunctions to the sensors, since more than one estimation is given.

In the SHRED architecture, the recurrent neural network (Lipton et al., 2015) is an LSTM (Hochreiter and Schmidhuber, 1997; Yu et al., 2019) network, which models a time sequence of measurements, or trajectory, from a limited number of point sensors in any one of the temporal fields of the nuclear reactor dynamics; the LSTM itself constructs a latent space representation of the dynamics given a time-lagged embedding, which has been shown to be related to Takens embedding theory (Takens, 1981). The latent space then projects the data through an SDN (Erichson et al., 2020) back to the latent space of the fields of interest, to be later de-compressed into the full-order space with the SVD.

As already mentioned, SHRED has a theoretical basis in the PDE theory of separation of variables (Faraji et al., 2025) along with the Takens embedding theorem (Takens, 1981). As already detailed in Williams et al. (Faraji et al., 2025), SHRED exploits the fact that time measurements are equivalent to spatial measurements in a spatio-temporal system. Thus sensor trajectory histories can completely characterise the dynamics regardless of measurement location (unless a region is statistically independent of other regions). Moreover, using the fact that $N$ first order PDEs can be rewritten as an $N$th order PDE, it can be shown that a single field variable encodes all other variables coupled to it through the evolution of its time dynamics. This is the basis of the Takens embedding theorem that guarantees that sensor trajectory information from a single field guarantees a diffeomorphic representation of all other fields. With training, SHRED uniquely determines the diffeomorphism to be fixed to the original spatio-temporal fields. In this sense, the SHRED architecture can be conceived as a generalisation to nonlinear PDEs, whereas a rigorous justification is available for linear PDEs only. As already shown in Williams et al. (2024), Ebers et al. (2023) and Faraji et al. (2025), SHRED provides important advantages with respect to standard data-driven ROM methods including:

**Table 1**
Architecture of the SHRED networks.

| | Hidden layer 1 | Hidden layer 2 |
|---|---|---|
| LSTM | 64 | 64 |
| SDN | 350 | 400 |

- the ability to use only three sensors (even randomly selected) for reconstructing the entire dynamics (all fields) of a physical system;
- the ability to train on compressed data, spanned by the SVD basis and thus enabling laptop level training in minutes;
- the ability to measure a single field variable (the most convenient) and reconstruct coupled spatio-temporal fields that are not observable;
- minimal hyper-parameter tuning; in fact, the same architecture for the neural network can be adopted for different physical problems without the need to optimise the number of neurons and layers.

As noted, the SHRED architecture is agnostic toward sensor placement (Williams et al., 2024; Faraji et al., 2025); namely, there is no need to determine the optimal sensor configuration (Argaud et al., 2018), setting itself apart from most data-driven methods in which selecting the optimal positions for the sensors (that is, the ones that allow retrieving the most information) is a key part of the training phase, especially for safety-critical industries such as the nuclear one (Cammi et al., 2024; Gong et al., 2024). Furthermore, architectures like SHRED can provide full-state reconstruction with even out-core measurements of a single quantity that might be the simplest (least expensive) to diagnose. In harsh environments like the Molten Salt Fast Reactor, this is crucial for monitoring and safety, especially in the long term.

The SHRED architecture has been implemented in Python using the PyTorch package (Ansel et al., 2024) and adopting the code developed by Williams et al. (2024) and Faraji et al. (2025). Both the LSTM and the SDN network are composed by 2 hidden layers, with the number of neurons reported in Table 1: the structure of the network was optimised within the initial works for SHRED (Williams et al., 2024; Faraji et al., 2025): numerical evidences show that these same hyperparameters can be used for any other application, without the need for further tuning (Tomasetto et al., 2025; Introini et al., 2025).

## 3. Molten salt fast reactor

Among the innovative Generation-IV (Generation IV International Forum, 2014) reactors, the Molten Salt Fast Reactor has the unique

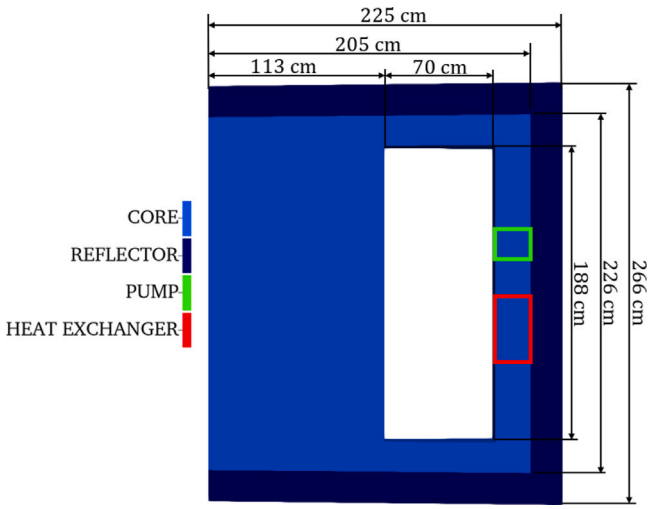



**Fig. 2.** OpenFOAM simulation domain with the main dimensions and the primary loop components. The geometry refers to a 2D axisymmetric wedge of the EVOL geometry of the European MSFR design, including the shown components (molten salt fuel in light blue, Hastelloy reflector in dark blue, primary pump in green, heat exchanger with the intermediate cycle in red). The blank hole represents the solid salt fertile blanket, not simulated in the present model.

feature of a liquid Thorium-based fuel homogeneously mixed with the thermal carrier; in particular, the fuel is a mixture of $^{7}$LiF and $ThF_4$ (77.5–22.5 mol%), having a melting temperature of 838 K. Given a total fuel inventory of 9 $m^3$, this reactor can produce up to 3 $GW_{th}$ thermal (Serp et al., 2014). The liquid nature of the MSFR fuel does not allow in-core sensing, as usually performed in water-cooled reactors, thus restricting measurements to be collected with out-core sensors; additionally, the fast energy spectrum of the reactor (Brovchenko et al., 2013) makes the core environment very hostile (Cammi et al., 2024). Accordingly, the sensor placement will not be performed randomly on the whole domain, but the (random) choice is restricted to the reflector region only, as in Riva et al. (2024a).

As a test case for the SHRED architecture, this work uses the 2D axisymmetric wedge (5°) of the EVOL geometry of the MSFR (Brovchenko et al., 2013), with an important modification: indeed, the present test case includes an external boundary of thickness 20 cm to mimic the presence of the Hastelloy reflector (Riva et al., 2024a): this subdomain is the only region in which, during the training phase, sensor locations can be randomly sampled. The thermo-physical properties of this region are those of nickel-based alloys, with the reflector neutronic group constants computed using the Monte Carlo code OpenMC (Romano et al., 2015), whereas data for the fuel were taken from Aufiero (2014). Thus, the simulation domain $\Omega$ includes two regions with different properties: the liquid core $\Omega_{\text{core}}$ and the solid reflector $\Omega_{\text{refl}}$. Fig. 2 depicts the simulated domain along with its main dimensions, including the primary loop components (pump and heat exchanger). The white cavity represents the location of the fertile blanket, not modelled in the present work.

The adopted numerical solver, developed at Politecnico di Milano, allows performing coupled neutronics and thermal-hydraulics simulations within the OpenFOAM environment (Aufiero, 2014). More in detail, the thermal-hydraulic sub-solver implements the incompressible single-phase version of the Reynolds-Averaged Navier–Stokes (RANS) equations with the Realisable $\kappa-\varepsilon$ turbulence model and the Boussinesq approximation to consider buoyancy effects (under the assumption that thermo-physical properties are constant within the same region); the neutronic sub-solver adopts the multi-group neutron diffusion approximation and includes transport equations for the delayed neutrons and the decay heat precursors. The Doppler and thermal expansion effects for the neutronic feedback coefficients have been modelled using, respectively, a linear and a logarithmic term correcting the reference group constants; furthermore, a momentum source and a heat sink represent the primary loop pump and the heat exchanger, respectively. For the interested reader, a summary of the governing equations, discretised using the Finite Volume Method in OpenFOAM, is reported in Appendix. The mesh consists of 46 424 hexahedral and 266 prismatic cells.

The transient considered in this work to generate the training dataset for SHRED is the accidental scenario of symmetric and unprotected failure of the primary pump, namely an Unprotected Loss of Fuel Flow (ULOFF). In this scenario, the flow rate of the pump is decreased exponentially (Fig. 3) resulting in a consequent decrease of the velocity magnitude in the reactor. In the first few seconds of the simulation, the reduction in flow rate translates into an increase in the power-to-flow ratio, increasing the average temperature, since there is a strong change in the dynamics from a forced convection problem to a natural circulation one. As the temperature rises in the core, the thermal feedback coefficients of the neutronics (Duderstadt and Hamilton, 1976) come into play, making the overall power decrease exponentially (Fig. 3). The selected time interval of the simulation of the ULOFF transient is 25 s, with a saving time of $\Delta t = 0.05$ s resulting in $N_t = 500$ snapshots. The elapsed CPU time for simulating this accidental scenario is around 10 h. For this first work, the ULOFF transient has been selected as test case to verify the applicability of the architecture to nuclear applications: however, given that the same architecture can be applied to several engineering problems ranging from plasma physics to geosciences, it is given that the SHRED architecture can be applied to any other transient conditions, without loss of generality and without the need of modifying the network. Research on this is currently underway.

The MSFR is a complex nuclear reactor, characterised by several fields that describe the neutron economy and the thermal-hydraulics: in particular, for this case, six group flux in energy $\{\phi_g\}_{g=1}^{6}$, eight groups of delayed neutrons $\{c_k\}_{k=1}^{8}$ and three decay heat groups $\{d_i\}_{i=1}^{3}$ are considered for the neutronic side, to which the thermal-hydraulics triplet, namely pressure, temperature and velocity $(p, T, \mathbf{u})$, must be added. Except for the velocity $\mathbf{u}$, all the others are scalar fields. Overall the full-order state space vector $\mathcal{V}$ is represented by 20 different fields, i.e.

$$\mathcal{V} = \left[\phi_1, \dots, \phi_6, c_1, \dots, c_8, d_1, \dots, d_3, p, T, \mathbf{u}\right] \tag{1}$$

In real engineering systems, it is not always possible to have access to all the quantities of interest, such as those reported above; however, for strongly coupled problems as the MSFR test case, which are inherently multi-physics in nature and for which separation of the various physics is impossible, each field carries some information about other quantities, thus, information on the un-observable fields can be extracted from the observable ones (Gong et al., 2022a; Introini et al., 2023b): for instance, in the present test case temperature and velocity are connected through the buoyancy and the energy balance, and the neutron fluxes strongly depend on the temperature distribution (Aufiero, 2014). It is then legitimate to investigate the indirect reconstruction problem for such cases, assuming only few fields (as low as one) can be directly measured.

### 3.1. SHRED setup for multi-physics systems

The SHRED architecture learns a map between three measurements of an observable field and the reduced representation through the SVD of the full-order state space vector; therefore, it is necessary to obtain time-series data of an observable field and perform the SVD of each characteristic field.

Starting from the latter, each field (generically indicated as $\psi$) is organised in the form of a snapshot matrix $\mathbb{X}_{\psi} \in \mathbb{R}^{\mathcal{N}_h \times N_t}$ such that the $j-$th column represents a snapshot at a given time $t_j$, given $\mathcal{N}_h$ the dimension of the spatial mesh. The snapshots are normalised to the

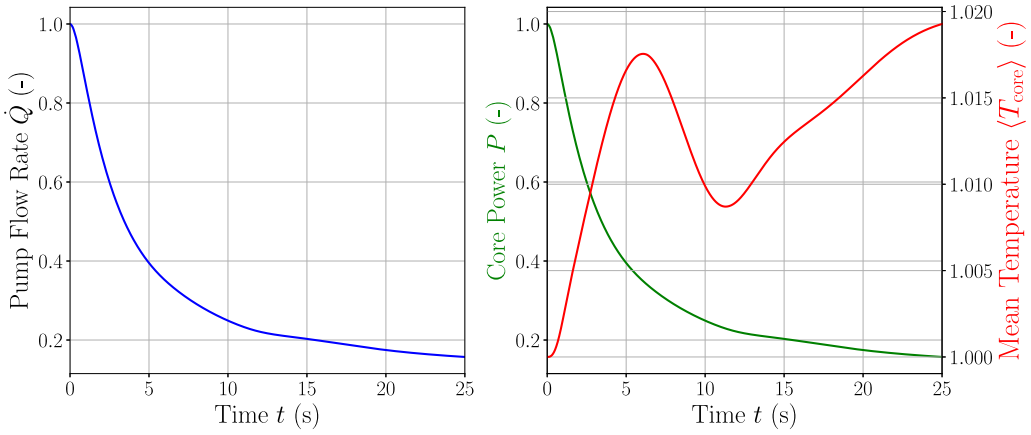



**Fig. 3.** Evolution in time of global core quantities, each normalised with respect to the initial condition, following an unprotected failure of the primary pump and an exponential decrease of the pump flow rate, up to 25 s from the initiating event. On the left, the pump flow rate (input), whereas on the right the core power and the average temperature (outputs).

mean value at the initial time to obtain variables within the same range (as the different fields of the MSFR are characterised by vastly different time scales: as an example, the neutron flux reaches $\sim 10^{19}$ whereas the velocity magnitude is around $\sim 1$), i.e.

$$\mathbb{X}_{\psi,ij} \longleftarrow \frac{\psi(\mathbf{x}_i; t_j)}{\langle \psi(\mathbf{x}_i; 0) \rangle} \tag{2}$$

The Randomised SVD, implemented in *scikit-learn* (Pedregosa et al., 2011), is used to obtain a reduced representation in terms of the first $r$ principal components:

$$\mathbb{U}_{\psi,r} \mathbb{\Sigma}_{\psi,r} \mathbb{V}^*_{\psi,r} \approx \mathbb{X}_{\psi} \tag{3}$$

The columns of the orthogonal matrix $\mathbb{U}_{\psi,r} \in \mathbb{R}^{\mathcal{N}_h \times r}$ represent the basis encoding the spatial information; the diagonal matrix $\mathbb{\Sigma}_{\psi,r} \in \mathbb{R}^{r \times r}$ contains the singular values measuring the importance of each mode; the rows of the orthogonal matrix $\mathbb{V}_{\psi,r} \in \mathbb{R}^{N_t \times r}$ are the reduced coefficient encoding the dynamics of the snapshots. The rank $r$ is taken equal to 10 for all the 20 fields, adopting a truncation based on the singular values (Quarteroni et al., 2015). Fig. 4 shows the decay of the singular values for all the fields within the state vector $\mathcal{V}$. The rank $r$ has been selected by looking at the relative information content $I(r) = 1 - \frac{\sum_{k=1}^{r} \sigma_k^2}{\sum_k \sigma_r^2}$ (Quarteroni et al., 2015), ensuring that for all the considered fields $I(r)$ reaches a value lower than $10^{-4}$. Furthermore, in the plot it is possible to observe the relative importance of each singular value with respect to the first (dominant) one; in particular, the normalised tenth eigenvalue is lower than 0.01 (against a maximum value, corresponding to the first eigenvalue, of 1). This last fact underlines that the amount of information retained by the SVD modes is sufficient to describe all fields with an accuracy of 99% compared to the full-order state. Then, the reduced dynamics at time $t_j$ is collected into a reduced state space vector $\mathbf{v}_j \in \mathbb{R}^{r \cdot 20}$ being the temporal dynamics learnt by the SHRED architecture (Fig. 1).

Focusing on the input of SHRED, the sensor locations can only be sampled in the reflector region (Fig. 2) and as the observable field the fast flux $\phi_1(\mathbf{x}; t)$ has been selected, given $\mathbf{x} \in \Omega$ and time $t \in [0, 25]$ s, which is a sufficiently simple quantity to be measured in real scenarios. This particular choice is not a limitation of the method, as any of the fields can be selected as observed quantity (even though, in the real application, not all fields of interest could be easily measurable). The measurements have been synthetically generated from the OpenFOAM data computed with the model in Section 3, exploiting the pyforce package developed by the authors Riva et al. (2024b) and Cammi et al. (2024) (https://github.com/ERMETE-Lab/ROSE-pyforce); in particular, the sensors are modelled as linear functionals $l_m(\cdot) = l(\cdot; \mathbf{x}_m, s)$ with a Gaussian kernel, as in Maday et al. (2015) and Cammi et al. (2024), each characterised by its the centre of mass $\mathbf{x}_m \in \Omega_{\text{refl}}$ and point spread $s$

$$l_m(\phi_1(\mathbf{x}); \mathbf{x}_m, s) = \int_{\Omega} \phi_1(\mathbf{x}) \cdot K \cdot e^{\frac{-\|\mathbf{x}-\mathbf{x}_m\|_2^2}{2s^2}} \, d\Omega \tag{4}$$

with $K$ defined such that $l_m(1; \mathbf{x}_m, s) = 1$ (Maday et al., 2015). For this application, the point spread has been taken equal to 0.025 as done in Cammi et al. (2024).

As already mentioned, the only available positions for the SHRED model are located in the reflector region $\Omega_{\text{refl}}$: this choice is intentional, as the aim of this work consists of developing a state estimation routine for the control and the monitoring of nuclear reactor and in-core measurements are quite hard to make. Additionally, this work also has the second objective of giving a measure of the uncertainty of the prediction, as the SHRED architecture is quick to train Williams et al. (2024) and Ebers et al. (2023) and it does not require powerful GPUs for the training phase and even personal computers can be used to perform quickly and reliably this task. This implies that several SHRED models can be trained with different random selection of sensors, such that different outputs, i.e. different reduced state space vectors $\mathbf{v}$, are produced, from which the sample mean and the associated standard deviation can be obtained. In this way, the predictions become more robust against random noise and this framework can be generalised to identify possible malfunctions in the reactors; if more than one SHRED model sees an upcoming accidental scenario, proper countermeasures can be undertaken based on more statistically consistent data.

Therefore, in this work $L = 30$ SHRED models are trained using different sensor configurations: Fig. 5 reports all the different positions for each configuration in the reflector region. The sensors have been randomly placed to show the reliability of the SHRED architecture and its capability to provide an accurate state estimation of the whole state space, even when measurements are randomly picked in the domain.

The capability of the SHRED architecture to retrieve the full state information from randomly selected sensors is of particular interest for the Nuclear Engineering community, where the problem of finding the optimal sensor placement is still an open issue (Argaud et al., 2018; Cannarile et al., 2018). The proposed framework for Uncertainty Quantification based on the SHRED architecture can be easily integrated with algorithms for optimal sensor placement in the following way: the appropriate algorithm select the most significant locations; then, the SHRED is used to generate robust and reliable predictions by randomly choosing among the candidate locations. This will increase the monitoring capabilities of nuclear reactors, ensuring a safer system. Indeed, in the nuclear power plants industry, the selection of the location of sensors cannot be fully arbitrary, as regulator authorities require to have sensors placed in specific locations to monitor the

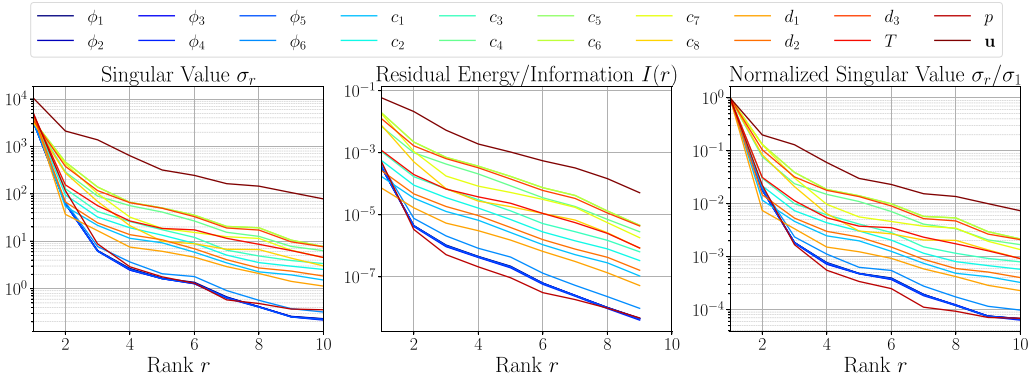



**Fig. 4.** Comparison of the singular values for all the fields in the state vector $\mathcal{V}$: from left to right are plotted the singular values $\sigma_r$, the relative information/energy content $I(r) = 1 - \frac{\sum_{k=1}^{r} \sigma_k^2}{\sum_k \sigma_r^2}$ (Quarteroni et al., 2015) and the relative importance of each singular value with respect to the first one.

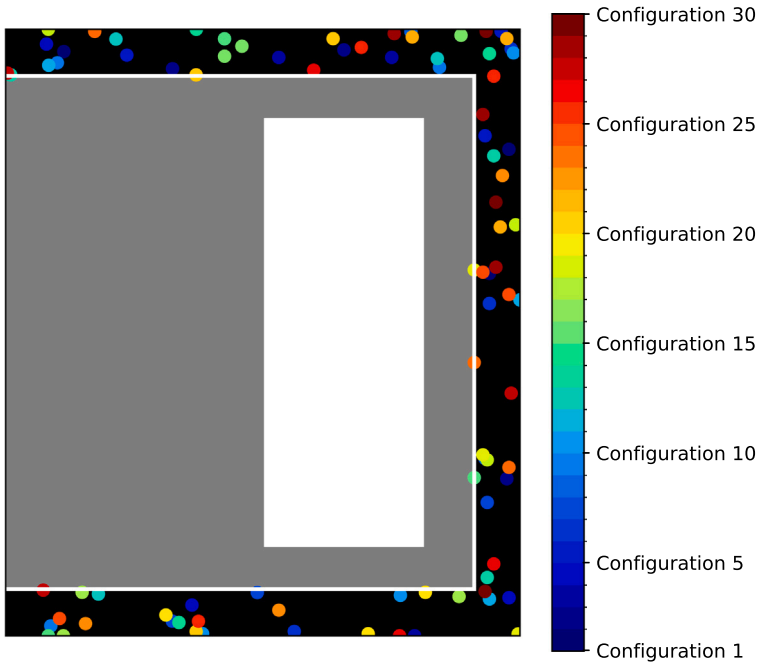



**Fig. 5.** Random sensor configurations adopted in this work, placing them in the reflector region only (black zone); the grey area is the actual core. As expected, there is no discernible pattern in the positioning of the sensors in neither direction.

most sensitive points of the engineering system (for example, where the temperature is expected to reach its maximum). Thus, the coupling between the SHRED architecture and an optimal sensor positioning algorithm will, on the one hand, make use of the potentiality of the SHRED architecture. In contrast, the positioning algorithm will ensure that locations are selected among those of interest from the regulatory and safety point of view.

Once the positions of the sensors have been chosen, the measures can be generated, i.e. the fast flux measurements $\mathbf{y}^k_{\phi_1}(t_j) \in \mathbb{R}^3$ of SHRED configuration $k$. To mimic the real scenario, these have been polluted with random zero-mean Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$

$$y^k_{\phi_1,m}(t_j) = (1+\epsilon) \cdot I^k_m(\phi_1(\mathbf{x}, t_j); \mathbf{x}_m, s) \qquad m = \{1,2,3\} \tag{5}$$

in which the standard deviation of the noise is considered to be 0.05, which is a reasonable value for the out-core of a MSFR (Brovchenko et al., 2013). The data have been split into three subsets (train, validation and test) for training and testing the neural network; then, the $k-$th SHRED model is trained in order to learn the input–output map between the measurements $\mathbf{y}^k_{\phi_1}(t)$ and the reduced state space vector $\mathbf{v}$.

## 4. Numerical results

### 4.1. Learning the dynamics

Since different SHRED models have been trained, different output predictions are retrieved, which can be used to obtain an estimate in terms of mean value and the standard deviation of each reduced coefficient $v^{\psi}_r$. Fig. 6 shows the first 5 SVD coefficients of each field of the state space, comparing the Ground Truth (solid lines), i.e. the Full Order Model (FOM), the SHRED mean value (dashed lines) and the associated confidence interval at 95%. The SHRED output is in almost perfect agreement with the Ground Truth, highlighting that the LSTM and the Shallow Decoder are robust and reliable tools able to learn correctly the dynamics underlying this accidental scenario; the pressure coefficients are the ones characterised by the highest uncertainty; nevertheless, the employment of different SHRED models allows for a mean prediction which is statistically close to the actual value. The errors of each SHRED simulation oscillates slightly compared to the average value, due to the

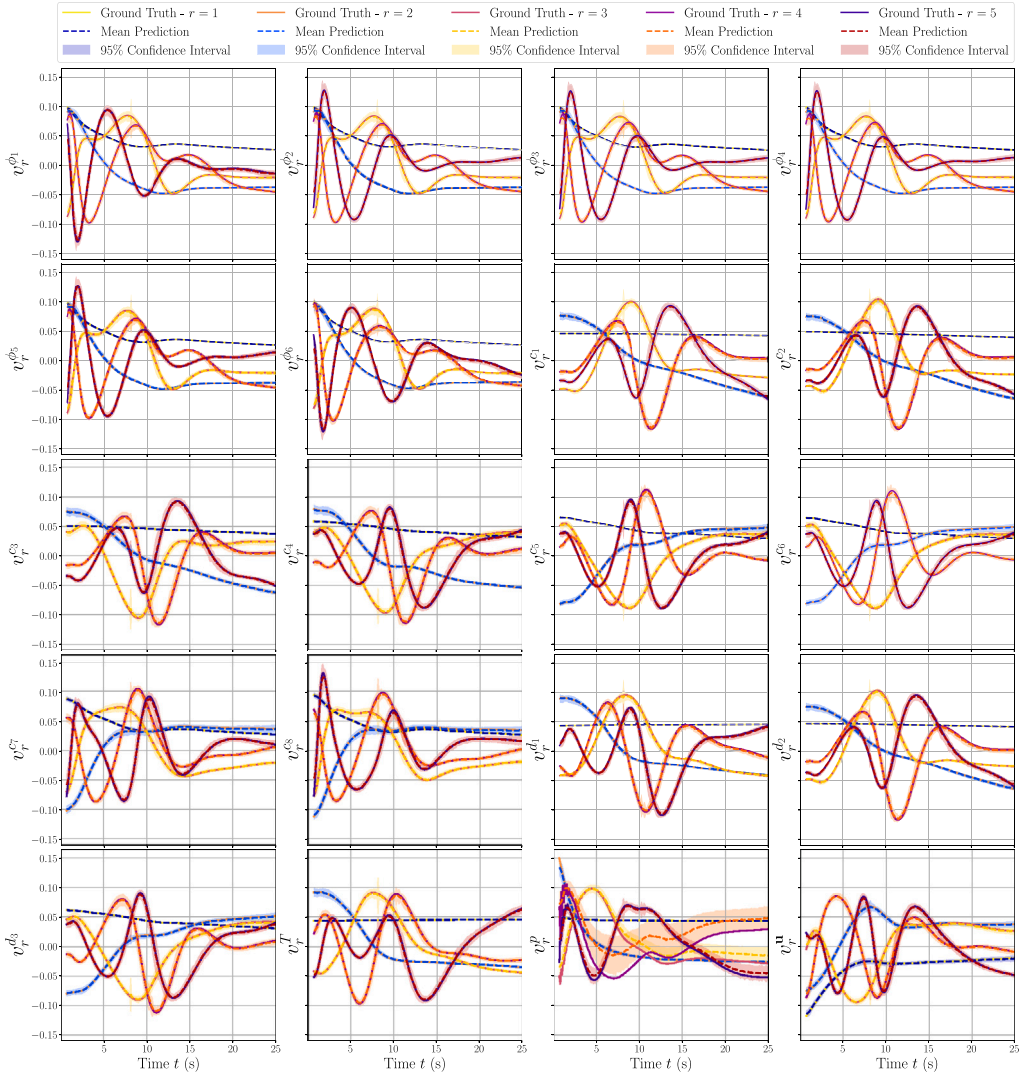



**Fig. 6.** Comparison of the SHRED reconstruction of the reduced state space vector $\mathbf{v}$ with respect to the test dataset for the first 5 SVD coefficients of each field: continuous curves represent the mean of the SHRED models, the dashed lines are the ground truth (descending from the full-order data) and the shaded areas highlight the uncertainty regions for the SHRED models. The SHRED architecture is indeed able to correctly learn the dynamics underlying the considered accidental scenario, and by employing different SHRED models a mean prediction, statistically close to the actual value, can be achieved.

fact that sensors may be placed in points very close to the wall in which the temporal dynamics are less pronounced; however, the ensemble allows for a more robust prediction compared to the single simulation, making the output almost independent on the sensor placement. In terms of computational costs, each SHRED model takes about 1 min of wall-clock time for the training phase on a workstation, and to get a new output, the associated computational cost is barely null these values for the training of SHRED models refers to a workstation with an Intel Core i7-9800X CPU with clock speed 3.80 GHz.

Now that the dynamics have been learnt, it is important to assess how the actual fields are predicted at the full-order level. Within the test set, the snapshots can be reconstructed using Eq. (3), assessing the reliability of the SHRED architecture coupled with the SVD. Let $\hat{\mathbb{X}}^k_{\psi,j} \in \mathbb{R}^{\mathcal{N}_h}$ be the reconstructed $j$th snapshot for SHRED configuration $k$ of the generic field $\psi$ and let $\mathbb{X}_{\psi,j} \in \mathbb{R}^{\mathcal{N}_h}$ be the full-order solution; then, the relative error for the generic field $\psi$ on the test set at time $t_j$ for configuration $k$ can be defined as

$$\varepsilon^{\psi,k}_{2,j} = \frac{\left\| \hat{\mathbb{X}}^k_{\psi,j} - \mathbb{X}_{\psi,j} \right\|_2}{\left\| \mathbb{X}_{\psi,j} \right\|_2} \tag{6}$$

from which the average (in time) relative test error $\varepsilon_2^{\psi,k}$ of field $\psi$ for SHRED configuration $k$ can be computed as

$$\varepsilon_2^{\psi,k} = \frac{1}{N_{t,test}} \sum_{j \in N_{t,test}} \varepsilon^{\psi,k}_{2,j} \tag{7}$$

from which the mean and the standard deviation considering all SHRED configurations can be computed.

Fig. 7 shows a bar plot of the average relative test error $\varepsilon_2^{\psi}$ and the associated standard deviation for each field, measured in the $l_2-$norm. It can be observed that the information coming from the measurements of the fast flux $\phi_1$ is sufficient to obtain the same error level on the other group fluxes, and the error itself is kept around the measurement noise level (5%); whereas the delayed groups (both precursors $c_k$ and decay heat $d_i$) are reconstructed with a low relative energy norm, as well as the temperature and the pressure: this means that the coupling between these fields and the observed one (the fast flux) is strong enough to provide a good reconstruction with sufficiently high accuracy. Moreover, it can be observed that the sensor positions do not affect the SHRED performance: in fact, the uncertainty for the relative test error is small, highlighting that the predictions of each SHRED model are very similar to one another and hence its accuracy. The field showing the highest

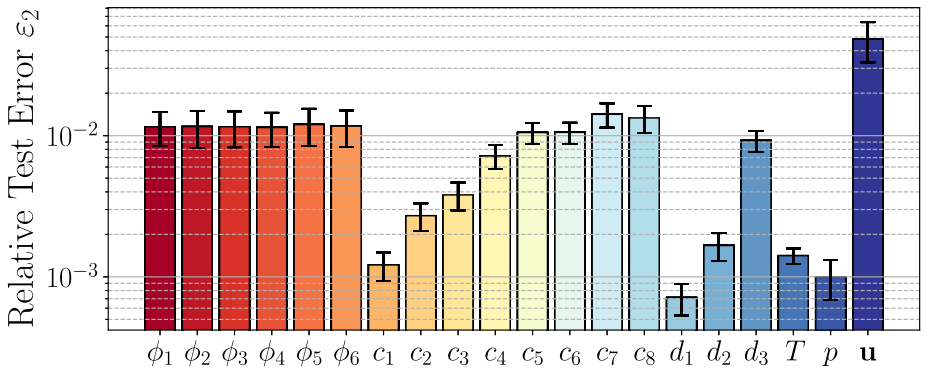



**Fig. 7.** Average (with respect to the different configurations) relative reconstruction errors $\varepsilon_2 = \langle \varepsilon_2^k \rangle$ between the SHRED reconstruction and the FOM for the test set, along with the standard deviation pertaining to each variable of interest, measured in the $l_2$-norm. The velocity (vector) field proves to be the hardest one to reconstruct; still, the information coming from a single field ($\phi_1$) is sufficient to reconstruct the other fields with good accuracy.

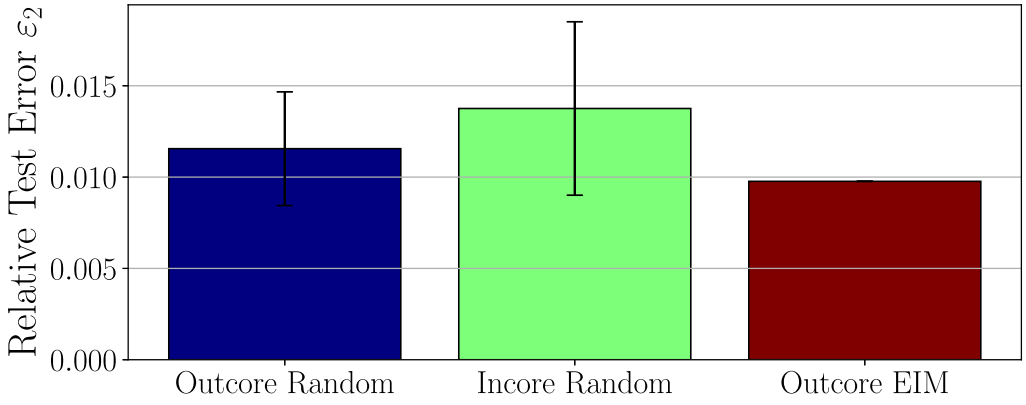



**Fig. 8.** Comparison of the relative errors of the measured field $\phi_1$ using SHRED, with 3 different sensor configurations: randomly selected outcore (blue) and incore (green) and the first 3 sensor given by EIM (red) (Maday et al., 2008). Results are very similar, indicating how SHRED is agnostic to sensor positioning and hierarchy.

error is the velocity **u**, and indeed this is notoriously the most complex field to be reconstructed due to its marked local behaviour: indeed, the singular values of **u** (Fig. 4 are the slowest decaying ones, meaning that, fixing a rank value common for all fields, more information will be discarded for the velocity field; likely, this field will require more SVD modes compared to the other ones, and thus a higher rank value, to catch all the relevant information, including that of the smaller scales. In this case, for the sake of simplicity, the rank $r$ has been taken the same for all the different fields; this assumption can, however, be easily relaxed, allowing for different ranks $r$ according to the singular values decay of the particular field of interest. It is worth mentioning how this limit is not related to the SHRED architecture itself, but rather to the preliminary data compression step.

### 4.2. Sensitivity assessment on sensors

Furthermore, a sensitivity analysis of the sensor placement constraints and the number of sensors has been performed. From a nuclear regulatory standpoint, placing sensor randomly could be hard to accept: therefore, it is worth comparing the overall output of SHRED with other reduced-order techniques that instead select sensors hierarchically, such as the Empirical Interpolation Method (EIM) (Maday et al., 2008) or its generalised version (GEIM) (Maday et al., 2015). Fig. 8 compares the performance of the SHRED architecture in terms of the average relative error for the measured field $\phi_1$ considering 3 sensors either randomly placed in the out-core region, in the core itself or selected as the first 3 sensors (in order of importance) given by the Empirical Interpolation Method (EIM) (Maday et al., 2008): for the two SHRED cases, the results come from the ensemble of 30 different configurations with the given constraint on allowed positions. These configurations provide very similar results, and little improvements are coming from EIM sensors, highlight how the SHRED architecture is agnostic to sensor positioning and almost agnostic to sensor hierarchy (Williams et al., 2024; Faraji et al., 2025). Having as input the measures collected at the EIM locations increases the importance of the location itself and this may explain why the EIM error is a bit lower. Conversely, the incore SHRED shows a slightly higher error than the outcore configuration, due to the noise level: higher noise levels are expected within the core, both due to the higher fluxes but also due to the more complex instrumentation electronics. In particular, referring to Eq. (5), the noise depends on the measurement value: as the fluxes are higher within the core, also the noise will be higher, leading to a slightly worse performance of the incore SHRED compared to the outcore one.

The SHRED architecture with 3 randomly placed sensors in the reflector region has also been compared in Fig. 9 to the case in which 10 measures are used as input for the neural network and to the case in which the Tikhonov Regularised GEIM (TR-GEIM) (Introini et al., 2023a), a state-of-the-art non-intrusive state estimation technique adopted in nuclear reactors, has been used for selecting the sensor positions. This figure shows that SHRED is independent of the number of measurements used, highlighting the fact that the dynamics is correctly embedded in 3 sensors; moreover, the comparison with TR-GEIM underlines its robustness with respect to noise, being SHRED able to reconstruct not only the measured field, but also the un-observable quantities. Additionally, as TR-GEIM selects sensor locations based on an hierarchical algorithm, the fact that SHRED shows comparable, and even better, performances confirms how the SHRED architecture can go beyond the logic of optimising the sensor positioning, thus making it possible to select the best possible location based on economic and accessibility considerations.

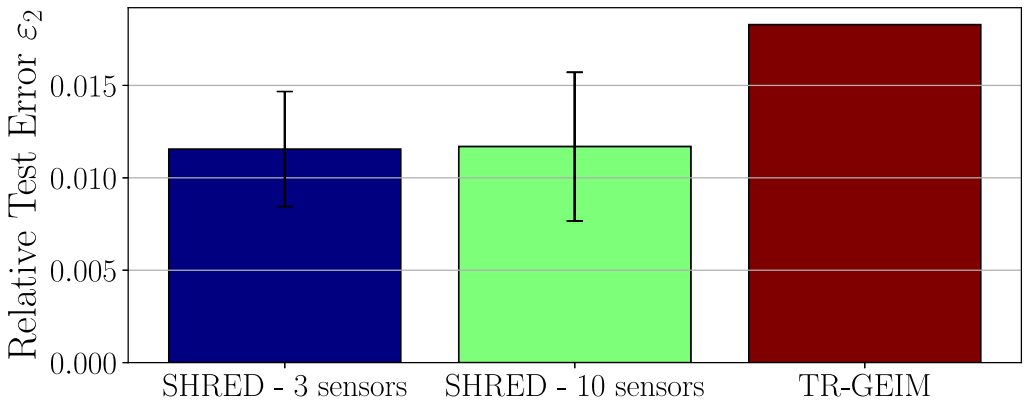



**Fig. 9.** Comparison of the relative errors of the measured field $\phi_1$ using SHRED, with 3 (blue) and 10 (green) random sensors and the TR-GEIM (Introini et al., 2023a) algorithm (red), indicating how SHRED is independent of the number of available sensors, robust against noise, and comparable in performance with a state-of-the-art technique for sensor positioning in nuclear reactors.

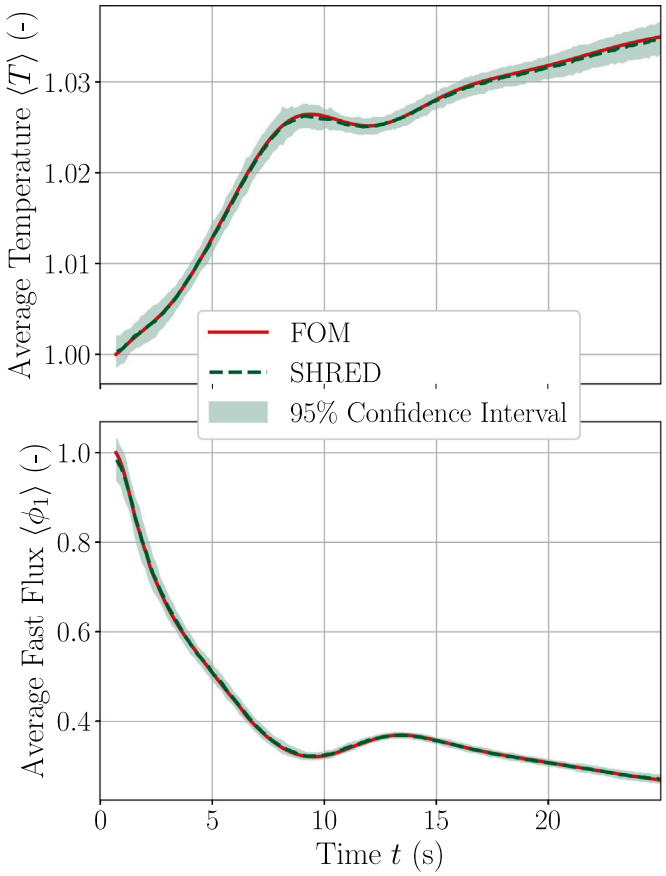



**Fig. 10.** Comparison of the full-order and the SHRED prediction for the average temperature of whole domain (including reflector) and average fast flux, normalised to the initial value and with confidence interval 95%, showing almost perfect agreement.

### 4.3. SHRED reconstruction

This framework can be very useful to obtain estimates of integral quantities typically used to monitor and control the nuclear reactor, such as the average (or maximum) temperature $\langle T \rangle$ and the core power (directly related to the average neutron flux). Fig. 10 reports the average temperature and the average fast flux, normalised to the initial value of the FOM and the mean SHRED reconstruction and the related uncertainty given by the different sensor configurations. The agreement is almost perfect for both quantities, showing that the SHRED is very accurate in estimating integral quantities describing on average what is occurring inside the nuclear reactor.

In the end, Fig. 11 shows some contour plots of the SHRED prediction at the last time step, compared against the FOM and along with the associated residual field $r_\psi$, defined as the absolute difference between the FOM ($\psi$) and the SHRED prediction ($\hat{\psi}$), i.e.

$$r_\psi(\mathbf{x};t) = |\psi(\mathbf{x};t) - \hat{\psi}(\mathbf{x};t)| \tag{8}$$

In addition, the standard deviation, with respect to the SHRED outputs, of the reconstructed fields is show, highlighting the zones with higher variation and thus the dominant structures cut off by the SVD. The SHRED model can provide a correct state estimation of the observable field $\phi_1$ and the un-observable ones, such as temperature $T$ and velocity $\mathbf{u}$. The latter is the most difficult one to reconstruct, nevertheless, the residual field in terms of magnitude is reasonably low. The predictions are physically reasonable with a very good agreement with the full order solutions: as a further confirmation of this, the residuals of the continuity equation (divergence-free condition for the velocity field) have been computed and the results are comparable with the ground-truth, see the complementary Github repository. In the end, some videos of the whole transient can be found at this link.

## 5. Conclusions

This work demonstrates the application of the SHallow REcurrent Decoder network to state estimation in advanced nuclear reactors, considering the test case of the Generation-IV Molten Salt Fast Reactor with only access to three out-core sensor measurements of the fast flux. In particular, the EVOL geometry was selected as the case study, considering the Unprotected Loss of Fuel Flow accidental scenario in which the external pumps lose power exponentially. The SHRED architecture is used to provide an estimation of the entire state space vector, which includes the energy fluxes, the neutron precursors, the decay heat groups and the thermal-hydraulics fields for a total of 20 coupled variables, using as input the out-core sensor measurements of the fast flux only, that is, collected with sensors placed in the external solid reflector region to mimic the actual reactor in which it will be impossible to have in-core sensors due to the liquid nature of the fuel.

In particular, several SHRED models have been trained considering different sensor configurations (in the reflector only), and hence different input measures: in this way, in addition to the full state estimation,

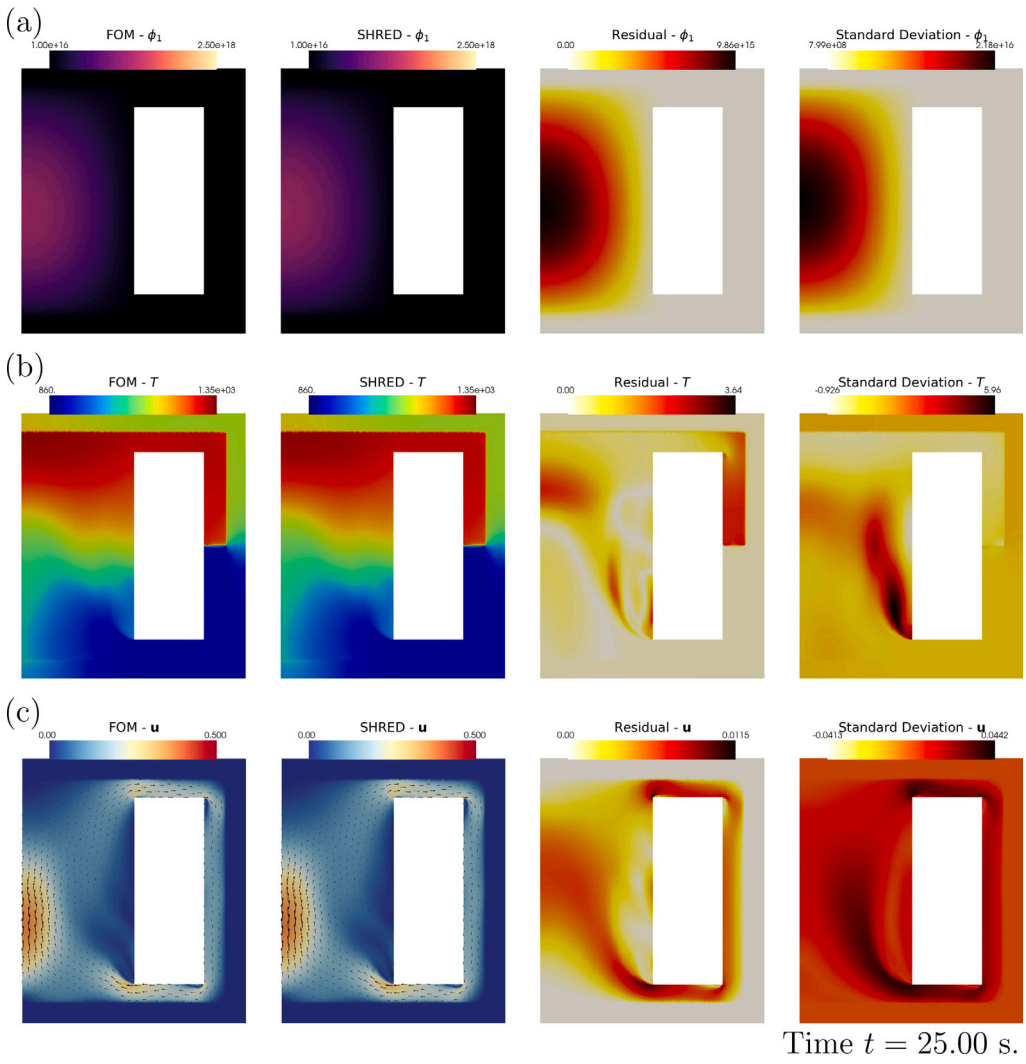



**Fig. 11.** Contour plots near the end of the transient (for the last time instant in the test set) of the observed field $\phi_1$ (a), the temperature $T$ (b) and the velocity $\mathbf{u}$ (c). From left to right, there is the full-order solution, the mean of the different SHRED models, the associated residual field and the standard deviation of the different SHRED models. The SHRED model provides a correct state estimation both of the observable and the un-observable ones; the right-most column, showing the standard deviation, allows to see the dominant structures cut off by the SVD, highlighting where the estimation is poorer.

a quantification of the associated uncertainty can be obtained, making the prediction more robust and reliable against noisy measurements. The results presented in this work are promising and show how the SHRED architecture can be used to reconstruct the state of a nuclear reactor starting from partial out-core observations.

As a first application of the SHRED architecture to nuclear cases, this paper considers a single accidental transient for the reconstruction; in the future, this same methodology will be used to identify the different scenarios in a classification paradigm, adding another fundamental algorithm for the development of digital twins of nuclear reactors. Moreover, this technique will be extended to possibly correct/update mathematical models from the local observation or to combine different levels of fidelity to improve the accuracy, both in the design and the operation phase. Finally, this work focuses on using the SHRED architecture in reconstruction mode from noisy measurements, that is, without considering forecast of the future state of the system: it will matter of future studies to investigate the use of the SHRED architecture in prediction and forecasting regime, leveraging the most recent extensions of the architecture itself.

## Code

The code and data (compressed) are available at: https://github.com/ERMETE-Lab/NuSHRED.

## CRediT authorship contribution statement

**Stefano Riva:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis. **Carolina Introini:** Writing – review & editing, Supervision, Software, Formal analysis. **Antonio Cammi:** Writing – review & editing, Supervision. **J. Nathan Kutz:** Writing – review & editing, Visualization, Supervision, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: J. Nathan Kutz reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have

no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.


## Acknowledgements

The work of JNK was supported in part by the US National Science Foundation (NSF) AI Institute in Dynamical Systems (dynamicsai.org), grant 2112085.


## Appendix. Governing equations of the molten salt fast reactor

The Molten Salt Fast Reactor (Generation IV International Forum, 2014) is an innovative reactor concept in which the fuel is liquid and homogeneously mixed with the molten salt thermal carrier; this configuration offers significant advantages, for example from the point of view of fuel reprocessing and closure of the fuel cycle, but also poses a lot of unique challenges in terms of monitoring and safety. Moreover, due to the strict coupling between the different physics, even more so than other nuclear reactor configurations, the mathematical modelling of such system requires suitable tools to integrate different physics together, starting from the neutronics and thermal-hydraulics (Aufiero, 2014).

The governing equations of the MSFR consist of the Navier–Stokes equations with the energy balance, including turbulence modelling with RANS and the Boussinesq approximation for density variations, the multi-group neutron diffusion and the precursors advection-diffusion: for completeness, this section reports the governing equations and briefly discusses their coupling.

The Navier–Stokes equations for the MSFR reads (Versteeg and Malalasekera, 2007):

$$\nabla \cdot \mathbf{u} = 0 \tag{A.1a}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} = +\nabla\cdot[\nu_{\text{eff}}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)] - \frac{1}{\rho}\nabla p + [1-\beta_T(T-T_0)]\mathbf{g} \tag{A.1b}$$

$$\frac{\partial T}{\partial t} + \nabla\cdot(\mathbf{u}T) = \nabla\cdot(\alpha_{\text{eff}}\nabla T) + \frac{q'''}{\rho c_p} \tag{A.1c}$$

given $\mathbf{u}$ as the velocity, $\rho$ as the density, $p$ as the pressure, $\nu_{\text{eff}}$ as the effective kinematic viscosity (accounting for turbulence modelling), $T$ as the temperature, $\beta_T$ as the thermal expansion, $\mathbf{g}$ as the gravity acceleration, $\alpha_{\text{eff}}$ as the thermal diffusivity, $q'''$ as the power density, which includes both neutron fission and delayed heat sources, and $c_p$ as the specific heat capacity. This set of equations shows the strong coupling between the two different physics; in particular, the source term of the energy equation directly depends on the neutron fluxes, i.e.

$$q''' = (1-\beta_h)\sum_g \bar{E}_{f,g}\Sigma_{f,g}\phi_g + \sum_i \lambda_{h,i}d_i \tag{A.2}$$

with $\phi_g$ being the neutron flux of energy group $g$, $\beta_h$ being the fraction of delayed heat source, $\bar{E}_{f,g}$ being the average energy released by fission in group $g$, $\Sigma_{f,g}$ being the fission cross section (Duderstadt and Hamilton, 1976) of group $g$, $\lambda_{h,i}$ being the decay constant of delayed decay group $d_i$.

Neutronics is governed by the Multi-Group Diffusion Equation (Duderstadt and Hamilton, 1976; Aufiero, 2014):

$$\frac{1}{v_g}\frac{\partial \phi_g}{\partial t} - \nabla\cdot(D_{n,g}\nabla\phi_g) + \Sigma_{r,g}\phi_g = S_g \tag{A.3}$$

with $D_{n,g}$ being the neutron diffusion coefficient of group $g$ and $\Sigma_{r,g}$ being the removal cross-section accounting for all the reactions which decrease the number of neutrons of group $g$, i.e. absorption $\Sigma_{a,g}$ and infra-group scattering $\Sigma_{s,g\to g'}$:

$$\Sigma^0_{r,g} = \left(\Sigma^0_{a,g} + \sum_{g'\neq g}\Sigma^0_{s,g\to g'}\right) \tag{A.4}$$

where the superscript 0 refers to the values at the reference temperature. In fact, the cross section are strongly affected by the neutron energy due to the Doppler broadening effect and the variation in density, i.e.

$$\Sigma_{r,g} = \left(\Sigma^0_{r,g} + A^0_{r,g}\ln\frac{T}{T_0^\Sigma}\right)\cdot\frac{1-\beta_T(T-T_0)}{1-\beta_T(T_0^\Sigma - T_0)} \tag{A.5}$$

In the end, the neutron diffusion equations have a source term $S_g$ accounting for the production of neutrons in the group energy $g$ due to prompt fission, decay of precursors and infra-group scattering:

$$S_g = (1-\beta)\chi_{p,g}S_{p,g} + \chi_{d,g}S_d + S_{s,g} \tag{A.6a}$$

$$S_{p,g} = \sum_{g'}\bar{\nu}_{g'}\Sigma_{f,g'}\phi_{g'} \tag{A.6b}$$

$$S_d = \sum_k \lambda_k c_k \tag{A.6c}$$

$$S_{s,g} = \sum_{g'\neq g}\Sigma_{s,g'\to g}\phi_{g'} \tag{A.6d}$$

with $\chi_{p,g}$ and $\chi_{d,g}$ as the fission spectrum of prompt and delayed neutrons and $\bar{\nu}_{g'}$ as the average number of neutrons produced in group $g'$.

When dealing with transient scenarios, the production of neutrons is split into prompt (due to fission) and delayed (due to decay): in the nuclear reactors world, this latter contribution is described by the group precursors $c_k$ which are governed by an advection–diffusion equation in the salt:

$$\frac{\partial c_k}{\partial t} + \nabla\cdot(\mathbf{u}c_k) = +\nabla\cdot(D_{\text{eff}}\nabla c_k) - \lambda_k c_k + \beta_{d,k}\sum_g \bar{\nu}_g\Sigma_{f,g}\phi_g \tag{A.7}$$

with $D_{\text{eff}}$ the effective diffusion coefficient and $\lambda_k$ the decay constant of delayed neutron precursor group $k$. In the end, the precursors contribute also to the energy Eq. (A.1c), where they are grouped together into the decay heat group $d_i$

$$\frac{\partial d_i}{\partial t} + \nabla\cdot(\mathbf{u}d_i) = +\nabla\cdot(D_{\text{eff}}\nabla d_i) - \lambda_{h,i}d_i + \beta_{h,l}\sum_g \bar{E}_{f,g}\Sigma_{f,g}\phi_g \tag{A.8}$$

## List of symbols

**Acronyms**

| | |
|---|---|
| DT | Digital Twin |
| EVOL | Evaluation and Viability of Liquid Fuel Fast Reactor System |
| FOM | Full Order Model |
| LSTM | Long Short-Term Memory |
| MSFR | Molten Salt Fast Reactor |
| PDE | Partial Differential Equation |
| POD | Proper Orthogonal Decomposition |
| RANS | Reynolds-Averaged Navier–Stokes |
| ROM | Reduced Order Modelling |
| SDN | Shallow Decoder Network |
| SHRED | SHallow REcurrent Decoder |
| SVD | Singular Value Decomposition |
| ULOFF | Unprotected Loss of Fuel Flow |

**Greek Letters**

| | |
|---|---|
| $\alpha_{\text{eff}}$ | Effective thermal diffusivity |

| | |
|---|---|
| $\bar{\nu}_g$ | Average Number of Neutrons produced by fission with energy in group $g$ |
| $\beta$ | Delayed Neutron Fraction |
| $\beta_h$ | Delay Heat Fraction |
| $\beta_T$ | Thermal Expansion Coefficient |
| $\chi_{d,g}$ | Delayed Neutron Spectrum of group $g$ |
| $\chi_{p,g}$ | Prompt Neutron Spectrum of group $g$ |
| $\epsilon$ | Random Noise |
| $\hat{\psi}$ | SHRED reconstruction of a Generic Field |
| $\kappa - \varepsilon$ | Turbulent Kinetic Energy and Turbulent Dissipation Rate |
| $\lambda_{h,i}$ | Decay Heat Constant for group $i$ |
| $\nu_{\text{eff}}$ | Effective kinematic viscosity |
| $\Omega$ | Physical Domain |
| $\phi_g$ | $g$th Neutron group Flux |
| $\psi$ | Generic Field |
| $\rho$ | Density |
| $\Sigma_{a,g}$ | Absorption Cross Section of group $g$ |
| $\Sigma_{f,g}$ | Fission Cross Section of group $g$ |
| $\Sigma_{r,g}$ | Removal Cross Section of group $g$ |
| $\Sigma_{s,g\to g'}$ | Scattering Cross Section from group $g$ to $g'$ |
| $\epsilon_2$ | Relative Error between FOM and SHRED in energy norm |
| $c_k$ | $k$th precursors group |
| $d_i$ | $i$th decay heat group |

**Latin Symbols**

| | |
|---|---|
| $\bar{E}_f$ | Average Energy released by Fission |
| $\Delta t$ | Time Step |
| $\dot{Q}$ | Pump Flow Rate |
| $\hat{\mathbb{X}}_\psi$ | Reconstructed Snapshot matrix with SHRED for generic field $\psi$ |
| $\mathbb{\Sigma}_\psi$ | SVD singular values for generic field $\psi$ |
| $\mathbb{U}_\psi$ | SVD basis for generic field $\psi$ |
| $\mathbb{V}_\psi$ | SVD reduced dynamics for generic field $\psi$ |
| $\mathbb{X}_\psi$ | Snapshot matrix for generic field $\psi$ |
| $\mathcal{N}_h$ | Spatial degrees of freedom |
| $\mathcal{V}$ | Full-Order state space |
| $\mathbf{g}$ | Gravity |
| $\mathbf{u}$ | Velocity vector |
| $\mathbf{v}_j$ | Reduced state space vector at time $t_j$ |
| $\mathbf{x}$ | Space coordinate |
| $\mathbf{y}$ | Measurement vector |
| $A^0_{r,g}$ | Proportionality Constant for Doppler Broadening for group $g$ |
| $c_p$ | Specific Heat Capacity at constant pressure |
| $D_{n,g}$ | Neutron Diffusion coefficient of group $g$ |
| $K$ | Normalisation constant for the sensor |
| $L$ | Number of SHRED models |
| $l_m$ | $m$th functional representing the sensor |
| $N_t$ | Number of time snapshots |
| $P$ | Core Power |
| $p$ | Pressure |
| $q'''$ | Power Density |
| $r$ | Rank of the SVD |
| $r_\psi$ | Residual field for generic field $\psi$ |
| $s$ | Point-spread of the sensor |
| $S_d$ | Delayed Neutron Source |
| $S_g$ | Neutron Source for group $g$ |
| $S_{p,g}$ | Prompt Neutron Source from Fission in group $g$ |
| $S_{s,g}$ | Source from Scattering Neutrons |
| $T$ | Temperature |
| $t$ | Time |
| $v_n$ | Reduced/Modal coefficient of rank $n$ |

## Data availability

The code and data (compressed) are available at: https://github.com/ERMETE-Lab/NuSHRED.

## References


Agency, I.A.E., 2023. Nuclear–Renewable Hybrid Energy Systems. Tech. Rep. NR-T-1.24, IAEA, Vienna, URL https://www.iaea.org/publications/15098/nuclear-renewable-hybrid-energy-systems.

Ansel, J., Yang, E., He, H., Gimelshein, N., Jain, A., Voznesensky, M., Bao, B., Bell, P., Berard, D., Burovski, E., Chauhan, G., Chourdia, A., Constable, W., Desmaison, A., DeVito, Z., Ellison, E., Feng, W., Gong, J., Gschwind, M., Hirsh, B., Huang, S., Kalambarkar, K., Kirsch, L., Lazos, M., Lezcano, M., Liang, Y., Liang, J., Lu, Y., Luk, C.K., Maher, B., Pan, Y., Puhrsch, C., Reso, M., Saroufim, M., Siraichi, M.Y., Suk, H., Zhang, S., Suo, M., Tillet, P., Zhao, X., Wang, E., Zhou, K., Zou, R., Wang, X., Mathews, A., Wen, W., Chanan, G., Wu, P., Chintala, S., 2024. PyTorch 2: Faster machine learning through dynamic Python bytecode transformation and graph compilation. In: Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2. ASPLOS '24, Association for Computing Machinery, New York, NY, USA, pp. 929–947. http://dx.doi.org/10.1145/3620665.3640366.

Argaud, J.-P., Bouriquet, B., de Caso, F., Gong, H., Maday, Y., Mula, O., 2018. Sensor placement in nuclear reactors based on the generalized empirical interpolation method. J. Comput. Phys. 363, 354–370. http://dx.doi.org/10.1016/j.jcp.2018.02.050, URL https://www.sciencedirect.com/science/article/pii/S0021999118301414.

Aufiero, M., 2014. Development of Advanced Simulation Tools for Circulating-Fuel Nuclear Reactors (Ph.D. thesis). Politecnico di Milano, http://dx.doi.org/10.13140/2.1.4455.1044.

Binev, P., Cohen, A., Mula, O., Nichols, J., 2018. Greedy algorithms for optimal measurements selection in state estimation using reduced models. SIAM/ASA J. Uncertain. Quantif. 6 (3), 1101–1126. http://dx.doi.org/10.1137/17M1157635, Publisher: Society for Industrial and Applied Mathematics. URL https://epubs.siam.org/doi/10.1137/17M1157635.

Brovchenko, M., Merle Lucotte, E., Rouch, H., Alcaro, F., Allibert, M., Aufiero, M., Cammi, A., Dulla, S., Feynberg, O., Frima, L., et al., 2013. Optimization of the pre-conceptual design of the MSFR.

Brunton, S.L., Kutz, J.N., 2022. Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control, second ed. Cambridge University Press, USA.

Cammi, A., Riva, S., Introini, C., Loi, L., Padovani, E., 2024. Data-driven model order reduction for sensor positioning and indirect reconstruction with noisy data: Application to a Circulating Fuel Reactor. Nucl. Eng. Des. 421, 113105. http://dx.doi.org/10.1016/j.nucengdes.2024.113105, URL https://www.sciencedirect.com/science/article/pii/S002954932400205X.

Cannarile, F., Baraldi, P., Colombo, P., Zio, E., 2018. A novel method for sensor data validation based on the analysis of wavelet transform scalograms. Int. J. Progn. Heal. Manag. 9 (1), http://dx.doi.org/10.36001/ijphm.2018.v9i1.2670, Number: 1. URL https://papers.phmsociety.org/index.php/ijphm/article/view/2670.

Duderstadt, J.J., Hamilton, L.J., 1976. Nuclear Reactor Analysis. Wiley New York.

Ebers, M.R., Williams, J.P., Steele, K.M., Kutz, J.N., 2023. Leveraging arbitrary mobile sensor trajectories with shallow recurrent decoder networks for full-state reconstruction. arXiv preprint arXiv:2307.11793.

Erichson, N.B., Mathelin, L., Yao, Z., Brunton, S.L., Mahoney, M.W., Kutz, J.N., 2020. Shallow neural networks for fluid flow reconstruction with limited sensors. Proc. R. Soc. A 476 (2238), 20200097. http://dx.doi.org/10.1098/rspa.2020.0097.

Faraji, F., Reza, M., Kutz, J.N., 2025. Shallow recurrent decoder for reduced order modeling of E × B plasma dynamics. Mach. Learn.: Sci. Technol. 6 (2), 025024. http://dx.doi.org/10.1088/2632-2153/adcd20.

Generation IV International Forum, 2014. Technology Roadmap Update for Generation IV Nuclear Energy Systems.

Gong, H., Cheng, S., Chen, Z., Li, Q., 2022a. Data-enabled physics-informed machine learning for reduced-order modeling digital twin: Application to nuclear reactor physics. Nucl. Sci. Eng. 196 (6), 668–693. http://dx.doi.org/10.1080/00295639.2021.2014752, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00295639.2021.2014752.

Gong, H., Cheng, S., Chen, Z., Li, Q., Quilodrán-Casas, C., Xiao, D., Arcucci, R., 2022b. An efficient digital twin based on machine learning SVD autoencoder and generalised latent assimilation for nuclear reactor physics. Ann. Nucl. Energy 179, 109431. http://dx.doi.org/10.1016/j.anucene.2022.109431, URL https://linkinghub.elsevier.com/retrieve/pii/S0306454922004613.

Gong, H.-L., Li, H., Xiao, D., Cheng, S., 2024. Reactor field reconstruction from sparse and movable sensors using Voronoi tessellation-assisted convolutional neural networks. Nucl. Sci. Tech. 35 (5), 43. http://dx.doi.org/10.1007/s41365-024-01400-w.

Gong, H., Zhu, T., Chen, Z., Wan, Y., Li, Q., 2023. Parameter identification and state estimation for nuclear reactor operation digital twin. Ann. Nucl. Energy 180, 109497. http://dx.doi.org/10.1016/j.anucene.2022.109497, URL https://linkinghub.elsevier.com/retrieve/pii/S0306454922005278.

Grieves, M.W., 2019. Virtually intelligent product systems: Digital and physical twins. In: Complex Systems Engineering: Theory and Practice. Progress in Astronautics and Aeronauting, pp. 175–200. http://dx.doi.org/10.2514/5.9781624105654.0175.0200, Ch. 7, arXiv:https://arc.aiaa.org/doi/pdf/10.2514/5.9781624105654.0175.0200. URL https://arc.aiaa.org/doi/abs/10.2514/5.9781624105654.0175.0200.

He, S., Wang, M., Zhang, J., Tian, W., Qiu, S., Su, G., 2022. A deep-learning reduced-order model for thermal hydraulic characteristics rapid estimation of steam generators. Int. J. Heat Mass Transfer 198, 123424. http://dx.doi.org/10.1016/j.ijheatmasstransfer.2022.123424, URL https://www.sciencedirect.com/science/article/pii/S0017931022008936.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. http://dx.doi.org/10.1162/neco.1997.9.8.1735.

Introini, C., Cavalleri, S., Lorenzi, S., Riva, S., Cammi, A., 2023a. Stabilization of Generalized Empirical Interpolation Method (GEIM) in presence of noise: A novel approach based on Tikhonov regularization. Comput. Methods Appl. Mech. Engrg. 404, 115773. http://dx.doi.org/10.1016/j.cma.2022.115773, URL https://www.sciencedirect.com/science/article/pii/S0045782522007290.

Introini, C., Riva, S., Kutz, J.N., Cammi, A., 2025. From models to experiments: Shallow recurrent decoder networks on the DYNASTY experimental facility. http://dx.doi.org/10.48550/arXiv.2503.08907, arXiv:2503.08907 [cs]. URL http://arxiv.org/abs/2503.08907.

Introini, C., Riva, S., Lorenzi, S., Cavalleri, S., Cammi, A., 2023b. Non-intrusive system state reconstruction from indirect measurements: A novel approach based on Hybrid Data Assimilation methods. Ann. Nucl. Energy 182, 109538. http://dx.doi.org/10.1016/j.anucene.2022.109538, URL https://www.sciencedirect.com/science/article/pii/S0306454922005680.

Kang, H., Tian, Z., Chen, G., Li, L., Wang, T., 2022. Application of POD reduced-order algorithm on data-driven modeling of rod bundle. Nucl. Eng. Technol. 54 (1), 36–48. http://dx.doi.org/10.1016/j.net.2021.07.010, URL https://www.sciencedirect.com/science/article/pii/S1738573321004137.

Kapteyn, M.G., Knezevic, D.J., Willcox, K., 2020. Toward predictive digital twins via component-based reduced-order models and interpretable machine learning. In: AIAA Scitech 2020 Forum. pp. 1–19. http://dx.doi.org/10.2514/6.2020-0418, arXiv:https://arc.aiaa.org/doi/pdf/10.2514/6.2020-0418. URL https://arc.aiaa.org/doi/abs/10.2514/6.2020-0418.

Lassila, T., Manzoni, A., Quarteroni, A., Rozza, G., 2014. Model order reduction in fluid dynamics: Challenges and perspectives. In: Reduced Order Methods for Modeling and Computational Reduction. Springer International Publishing, Cham, pp. 235–273. http://dx.doi.org/10.1007/978-3-319-02090-7_9.

Lipton, Z.C., Berkowitz, J., Elkan, C., 2015. A critical review of recurrent neural networks for sequence learning. arXiv preprint arXiv:1506.00019.

Maday, Y., Mula, O., Patera, A.T., Yano, M., 2015. The Generalized Empirical Interpolation Method: Stability theory on Hilbert spaces with an application to the Stokes equation. Comput. Methods Appl. Mech. Engrg. 287, 310–334. http://dx.doi.org/10.1016/j.cma.2015.01.018, Publisher: Elsevier B.V..

Maday, Y., Nguyen, N., Patera, A., Pau, G.S.H., 2008. A general multipurpose interpolation procedure: The magic points. Commun. Pure Appl. Anal. 8, http://dx.doi.org/10.3934/cpaa.2009.8.383.

Maday, Y., Patera, A., Penn, J., Yano, M., 2014. A parameterized-background data-weak approach to variational data assimilation: formulation, analysis, and application to acoustics. Internat. J. Numer. Methods Engrg. 102, http://dx.doi.org/10.1002/nme.4747.

Manohar, K., Brunton, B.W., Kutz, J.N., Brunton, S.L., 2018. Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns. IEEE Control Syst. Mag. 38 (3), 63–86. http://dx.doi.org/10.1109/MCS.2018.2810460.

Mohanty, S., Listwan, J., 2021. 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. Tech. Rep. ANL/LWRS-21/02, 1822853, 171255, Argonne National Laboratory, http://dx.doi.org/10.2172/1822853, URL https://www.osti.gov/servlets/purl/1822853/.

Mohanty, S., Vilim, R., 2021. Physics-Infused AI/ML Based Digital-Twin Framework for Flow-Induced-Vibration Damage Prediction in a Nuclear Reactor Heat Exchanger. Tech. Rep. ANL/NSE-21/8, 1830413, 172195, Argonne National Laboratory, http://dx.doi.org/10.2172/1830413, URL https://www.osti.gov/servlets/purl/1830413/.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830, http://jmlr.org/papers/v12/pedregosa11a.html.

Quarteroni, A., Manzoni, A., Negri, F., 2015. Reduced Basis Methods for Partial Differential Equations: An Introduction. In: UNITEXT, Springer International Publishing, URL https://link.springer.com/book/10.1007/978-3-319-15431-2.

Riva, S., Deanesi, S., Introini, C., Lorenzi, S., Cammi, A., 2024a. Neutron flux reconstruction from out-core sparse measurements using data-driven reduced order modelling. In: Proceedings of the International Conference on Physics of Reactors. PHYSOR 2024, pp. 1632–1641. http://dx.doi.org/10.13182/PHYSOR24-43444, URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85202802833&doi=10.13182%2fPHYSOR24-43444&partnerID=40&md5=1c4a8e77492242e3a8ed39787b40f37d.

Riva, S., Introini, C., Cammi, A., 2024b. Multi-physics model bias correction with data-driven reduced order techniques: Application to nuclear case studies. Appl. Math. Model. 135, 243–268. http://dx.doi.org/10.1016/j.apm.2024.06.040, URL https://www.sciencedirect.com/science/article/pii/S0307904X24003196.

Romano, P.K., Horelik, N.E., Herman, B.R., Nelson, A.G., Forget, B., Smith, K., 2015. OpenMC: A state-of-the-art Monte Carlo code for research and development. Ann. Nucl. Energy 82, 90–97. http://dx.doi.org/10.1016/j.anucene.2014.07.048, Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2013, SNA + MC 2013. Pluri- and Trans-disciplinarity, Towards New Modeling and Numerical Simulation Paradigms. URL https://www.sciencedirect.com/science/article/pii/S030645491400379X.

Rozza, G., Hess, M., Stabile, G., Tezzele, M., Ballarin, F., Gräßle, C., Hinze, M., Volkwein, S., Chinesta, F., Ladeveze, P., Maday, Y., Patera, A., Farhat Char, J., 2020. Model Order Reduction: Volume 2: Snapshot-Based Methods and Algorithms. De Gruyter, http://dx.doi.org/10.1515/9783110671490.

Ruth, M.F., Zinaman, O.R., Antkowiak, M., Boardman, R.D., Cherry, R.S., Bazilian, M.D., 2014. Nuclear-renewable hybrid energy systems: Opportunities, interconnections, and needs. Energy Convers. Manage. 78, 684–694. http://dx.doi.org/10.1016/j.enconman.2013.11.030, URL https://www.sciencedirect.com/science/article/pii/S0196890413007516.

Serp, J., Allibert, M., Beneš, O., Delpech, S., Feynberg, O., Ghetta, V., Heuer, D., Holcomb, D., Ignatiev, V., Kloosterman, J.L., Luzzi, L., Merle-Lucotte, E., Uhlíř, J., Yoshioka, R., Zhimin, D., 2014. The molten salt reactor (MSR) in generation IV: Overview and perspectives. Prog. Nucl. Energy 77, 308–319. http://dx.doi.org/10.1016/j.pnucene.2014.02.014, URL https://www.sciencedirect.com/science/article/pii/S0149197014000456.

Takens, F., 1981. Detecting strange attractors in turbulence. Lecture Notes in Math. 898, 366–381.

Tomasetto, M., Williams, J.P., Braghin, F., Manzoni, A., Kutz, J.N., 2025. Reduced order modeling with shallow recurrent decoder networks. http://dx.doi.org/10.48550/arXiv.2502.10930, arXiv:2502.10930 [cs]. URL http://arxiv.org/abs/2502.10930.

Versteeg, H.K., Malalasekera, W., 2007. An Introduction to Computational Fluid Dynamics: The Finite Volume Method. Pearson Education Limited, URL https://books.google.it/books?id=RvBZ-UMpGzIC.

Williams, J.P., Zahn, O., Kutz, J.N., 2024. Sensing with shallow recurrent decoder networks. Proc. R. Soc. A: Math. Phys. Eng. Sci. 480 (2298), 20240054. http://dx.doi.org/10.1098/rspa.2024.0054, arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.2024.0054. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2024.0054.

Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. 31 (7), 1235–1270. http://dx.doi.org/10.1162/neco_a_01199.