

A NOVEL MACHINE LEARNING-BASED DIGITAL TWIN ARCHITECTURE FOR REAL-TIME MONITORING, INTEGRATION AND CONTROL OF THE BREEDING BLANKET

A. Cammi, C. Intorini, C. Tripodo, A. del Nevo, M. Eboli, R. Marinari, L. Savoldi

Energy Department, Nuclear Engineering Division, Politecnico di Milano
ENEA Brasimone Research Center
MAHTEP Group, Dipartimento Energia "Galileo Ferraris", Politecnico di Torino

September 15, 2026

INTRODUCTION: SCOPE

- ▶ **Challenge:** Traditional high-fidelity fusion simulations are too computationally expensive for real-time reactor feedback.
- ▶ **Need:** Devices like ITER and DEMO require fast feedback and continuous monitoring through Digital Twin Environments (DTE).
- ▶ **Problem:** Fusion research currently lacks an integrated, end-to-end data-driven framework, often focusing on isolated components.

Build and prove the feasibility of a SHRED-based DT for the WCLL Breeding Blanket with reconstruction, prediction, forecast capabilities, to be integrated within larger DTE frameworks.

INTRODUCTION: E-TASK INTERLINKS

- ▶ **Modular DTE Framework:** This proposal advocates for a modular DTE where each component (e.g., Breeding Blanket) has a dedicated, high-performing DT.
- ▶ **Multi-Physics Integration:** The project aligns with the Data-Driven Predictive Modelling and Integrated Physics/Engineering Simulation Framework key areas.
- ▶ **Collaborative Modelling:** SHRED-based DTs will be integrated with plasma and balance-of-plant surrogate models, and interfaced with existing platforms like JINTRAC.

Build and prove the feasibility of a SHRED-based DT for the WCLL Breeding Blanket with reconstruction, prediction, forecast capabilities, to be integrated within larger DTE frameworks.

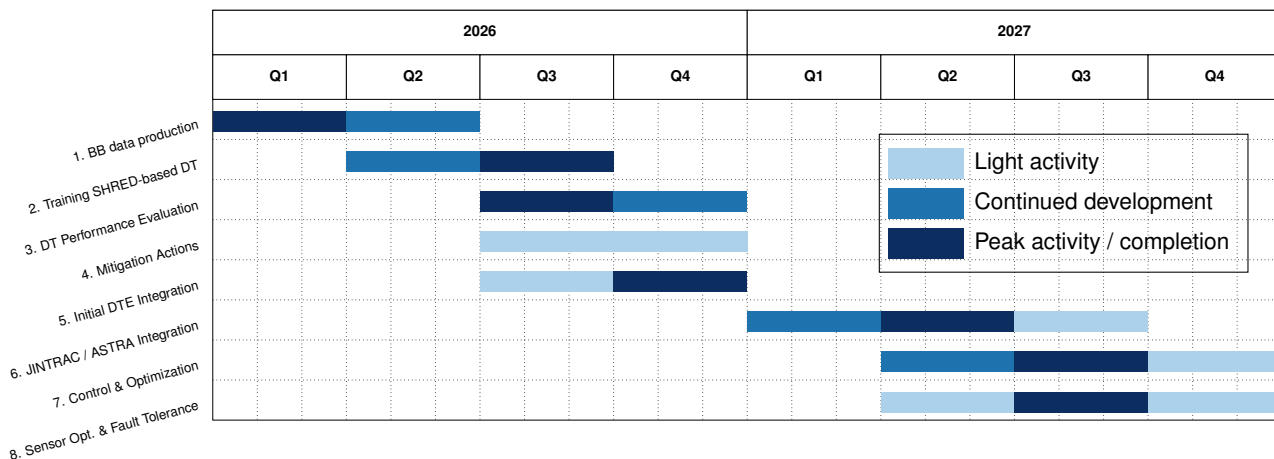
INTRODUCTION: TIMELINE

2026	WP1: Trained SHRED-based DT of the WCLL Breeding Blanket, using existing ENEA BB data for training, capable of operating in reconstruction, prediction and ensemble mode.	Deliverable: First scientific report, first main commit of the code with documentation and a representative benchmark dataset.
	WP2.1: Initial integration of the DT into a Python-based DTE, creating surrogate models for the plasma and the BOP for demonstration.	Deliverable: Second scientific report, second main commit of the code.
2027	WP2.2: Integrate the DT within existing modelling platforms (ASTRA/JINTRAC), ensuring compatibility and interoperability.	Deliverable: Third scientific report, third main commit of the code.
	WP3: DT-based control and optimisation framework, with closed-loop control mechanisms in the DT, assessing sensor optimisation and fault tolerance.	Deliverable: Fourth scientific report, fourth main commit of the code.
	Final Deliverable: Final report summarising the activity and the main results, final commit of the code with complete documentation and tutorial cases for reproducibility.	

INTRODUCTION: GANTT CHART

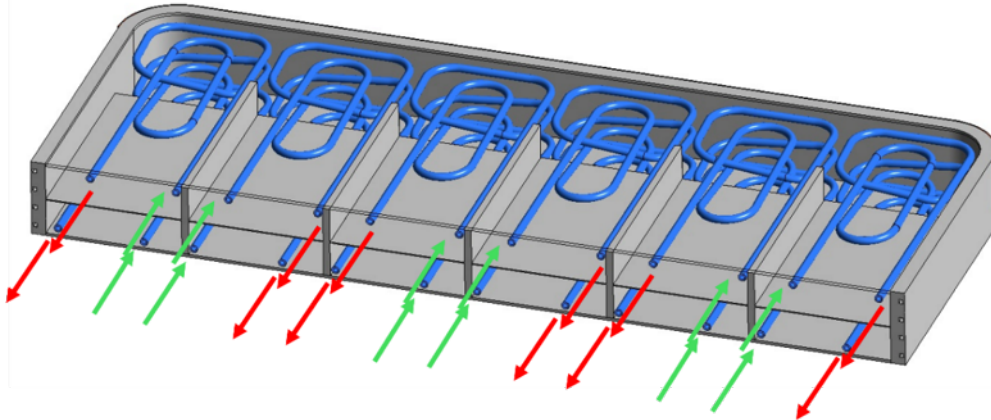
Foreseen Mitigation Actions:

- ▶ **WP1:** Switching architecture or surrogate models if training underperforms, running new simulations if training data are not sufficient.
- ▶ **WP2:** Use of simple lumped models for plasma and BOP for demonstration purposes.
- ▶ **WP3:** If control is not satisfactory, the DT will be used as a simulator and predictor only, focusing on sensor optimisation and fault tolerance.



CURRENT ACTIVITY

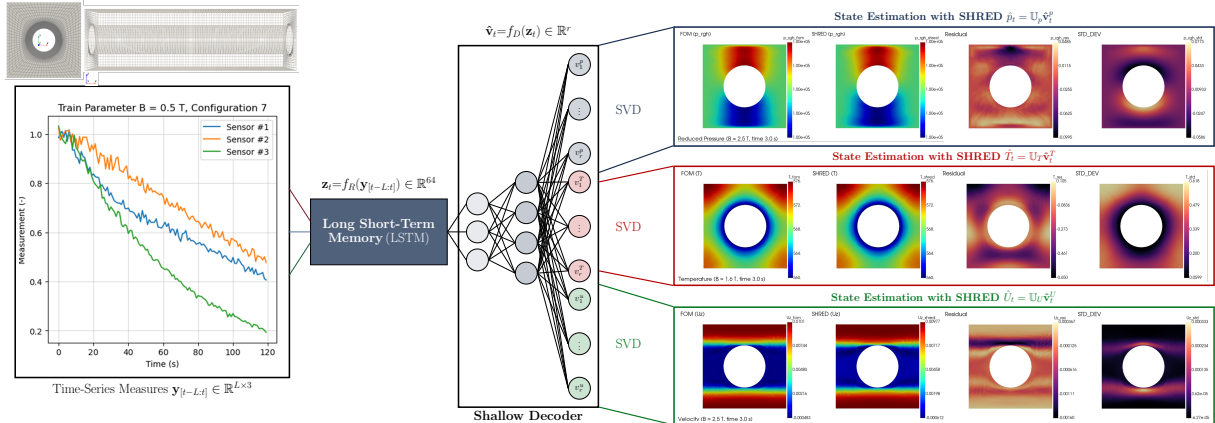
- Create the training (parametric and time-dependent) dataset for the WCLL Breeding Blanket elementary cell (courtesy of ENEA Brasimone) ⇒ **Current bottleneck due to the complexity of the geometry** (but will leverage computational nodes on Pitagora HPC and Japan Plasma Simulator)



- **Mitigation Action** (if needed): coordinate with ENEA and PoliTo to simplify the geometry while still maintaining it as representative of the WCLL BB as possible.

COMPLETED WORK

- The SHRED architecture has been tested on MHD benchmarks and simplified geometries to assess its performance (as well as on fission reactor test cases) $\Rightarrow$ **Very good accuracy all across the board and fast training/deployment times.** $\Rightarrow$ Once the dataset creation bottleneck will be solved, deployment will be much quicker as the SHRED architecture itself for MHD problems is ready to be used (only the input dataset and likely the dataset compression level will need changes).



SHRED ARCHITECTURE

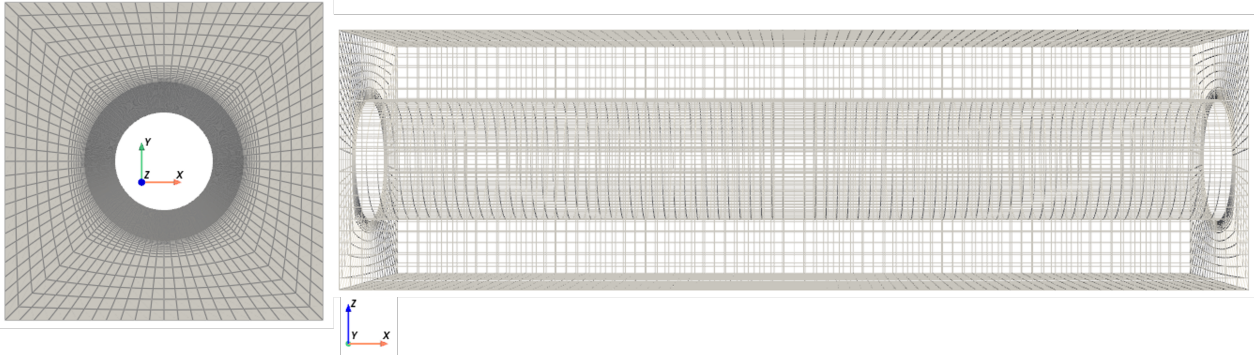
- **TL;DR:** A Long Short-Term Memory (LSTM) network creates a latent temporal model $\mathbf{z}_t$ from some input measurements $\mathbf{y}^T(t)$, which a Shallow Decoder Network (SDN) maps to the reduced POD coefficients $\hat{\mathbf{v}}_t$, which encode the temporal dynamics of the field variables; through SVD projection on the spatial basis functions $\mathbb{U}$, the full-order fields $(\mathbf{u}, T, p)$ are recovered.

Key Advantages

Data Efficient: Can operate with as few as three sensors, located anywhere in the domain.	Resource Efficient: Compressive training allows for laptop-level training and offline operation. Can operate in <i>Ensemble Mode</i> , providing an estimation of uncertainties and enabling faulty sensors detection.	Interpretability: Mathematically based on Takens' embedding theorem and separation of variables.
No Tuning: The same architecture can be used for different applications, provided a suitable dataset.	Physical-Digital Twin Linking: Experimental data from the physical twin act as input to the deployed DT.	Inverse Problem: Can use easy-to-measure quantities in easy-to-reach locations to infer unobservable fields.

DEMONSTRATIVE BENCHMARK: 3D CHANNEL

Scope: Demonstrate the performance of SHRED on a somewhat simple 3D liquid metal channel under an external magnetic field, measuring only temperature data at three random points located near the external boundary.



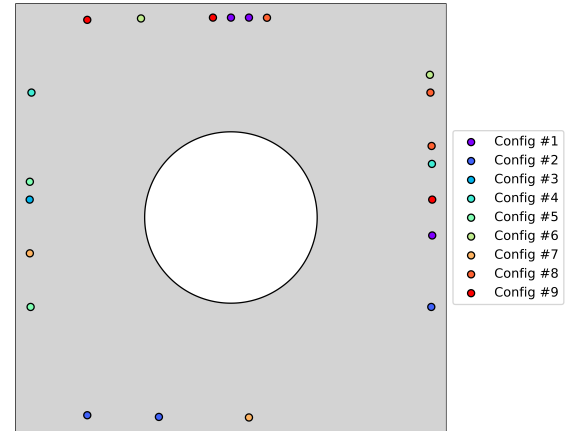
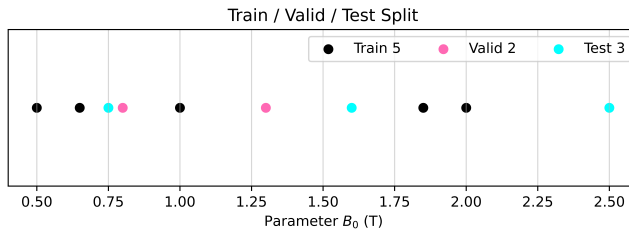
FOM Solver: <https://github.com/ERMETE-Lab/MHD-magnetoHDFoam>

DEMONSTRATIVE BENCHMARK: 3D CHANNEL

Dataset Size: $N_p = 10 \times N_t = 120 \times \mathcal{N}_h = 121176$ (SVD compressed size $r = 100$)

Data Preprocessing:

- ▶ Normalisation (MinMaxScaler)
- ▶ Dataset splitting (train - validation - test)
- ▶ Dataset compression through Singular Value Decomposition ($r = 100 \ll \mathcal{N}_h$)
- ▶ Ensemble mode: nine sensor configurations (three sensor each, randomly selected near the boundary)
- ▶ **Parametric Space μ**



DEMONSTRATIVE BENCHMARK: 3D CHANNEL

- ▶ **SHRED I/O:** Three (noisy) temperature measurements near the boundaries per each SHRED model
→ (Averaged) reduced POD coefficients $\mathbb{V}_{\mu_i}^k = (\mathbb{U}^k)^T \mathbb{S}_{\mu_i}^k \rightarrow$ Projection back to the full-order space and error computation.
- ▶ **Training times** (shared Workstation Intel Core i7-988X CPU @ 3.80 GHz, 32 GB RAM): between 4 and 8 minutes total (23 seconds for SHRED configuration on average).
- ▶ **Training MRE 4.72%** (σ 6.73%), **Test MRE 6.59%** (σ 6.24%).

$\mu = B_y$ (T)	$\hat{\mu}$	$\epsilon\%$
0.5	0.52	-3.99%
0.65	0.63	-2.69%
0.75	0.73	2.88%
0.8	0.79	1.03%
1	0.99	0.75%
1.3	1.47	-13.11%
1.6	1.43	10.83%
1.85	1.74	5.97%
2	1.79	10.48%
2.5	1.9	24.08%

Table. Parameter estimation with SHRED for train (blue), validation (green), test interpolated (red), test extrapolated (purple). **Low error up to 1 T (more data), then worst performance. Extrapolation beyond the training range ($B_y = 2.5$ T) up to 25% error on parameter.**

DEMONSTRATIVE BENCHMARK: 3D CHANNEL

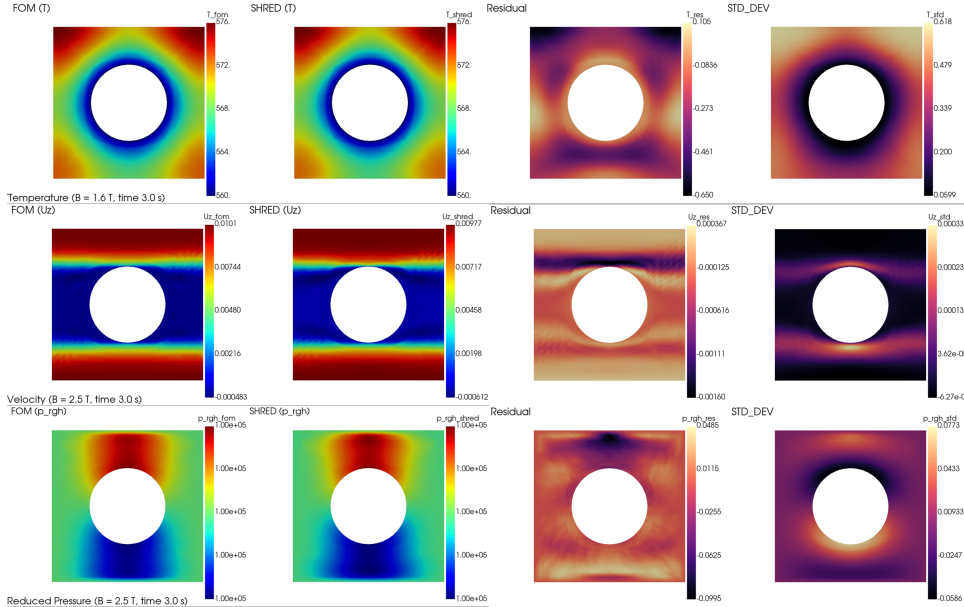


Figure. Reconstruction for the temperature, velocity and pressure field for the test parameter $B = 2.5$ T. First column: reference FOM; second column: SHRED reconstruction; third column: absolute residual; fourth column: SHRED ensemble standard deviation.

OTHER TEST CASES: 2D MHD FLOW

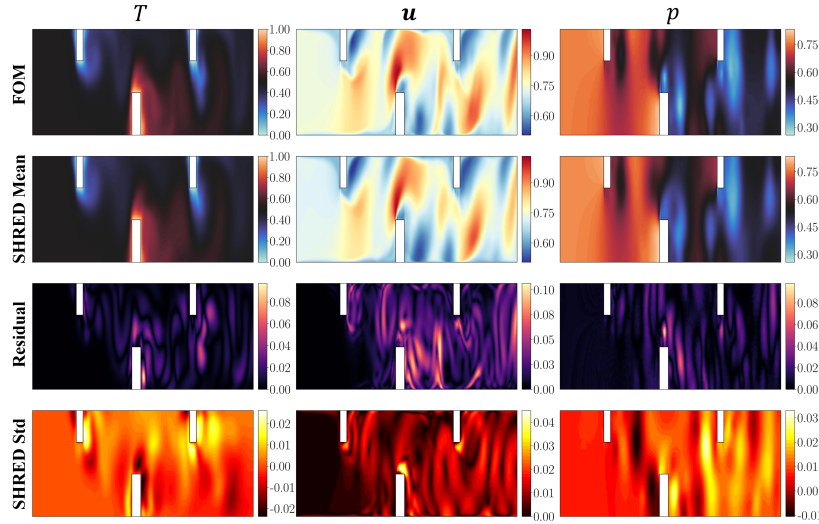


Figure. Results for the temperature (first column), velocity (second column) and pressure (third column) for test $B_0 = 0.06$ T at time $t = 2$ s. The first row displays the reference full-order solution while the second row shows the mean reconstruction obtained by averaging the outputs of the 30 SHRED models. The third row reports the difference between the FOM and the mean SHRED reconstruction while the fourth row shows the standard deviation among the 30 reconstructions.

OTHER TEST CASES: 2D MHD FLOW

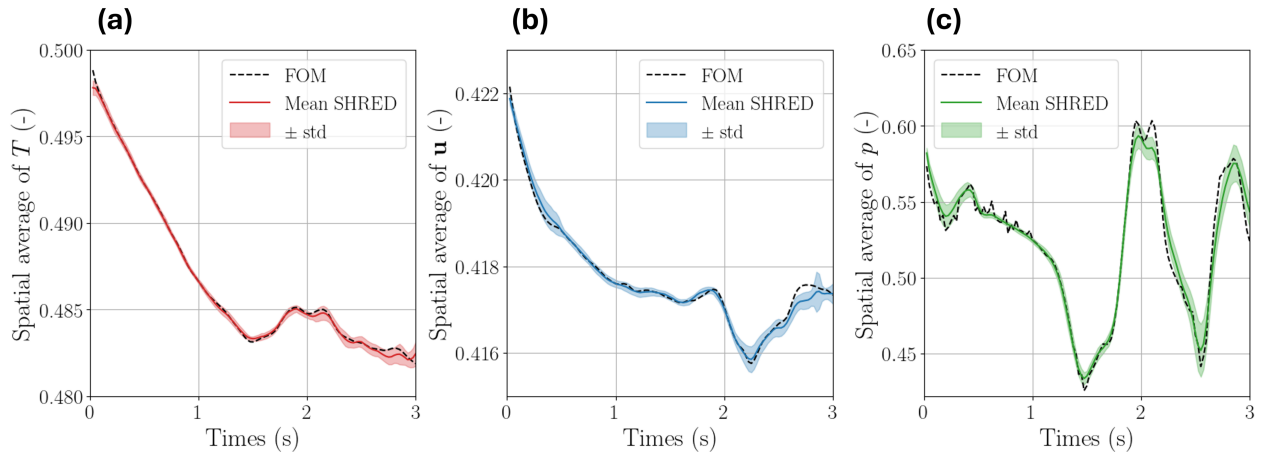


Figure. Temporal evolution of the spatial averages across the geometry for the temperature (a), velocity (b) and pressure (c) fields in the test case with test $B_0 = 0.06$ T. For each field the true profile associated to the full-order solution is compared with the average across the 30 reconstructions obtained, with the related standard deviations. Max error lower than 6%. **These integral estimates can serve as input for the next component DT.**

OTHER TEST CASES: TIME-VARYING MAGNETIC FIELD (3D CHANNEL)

- Temporal evolution: $B(t) = A\sin(\omega t + \phi) + C \Rightarrow$ Four-parameter case

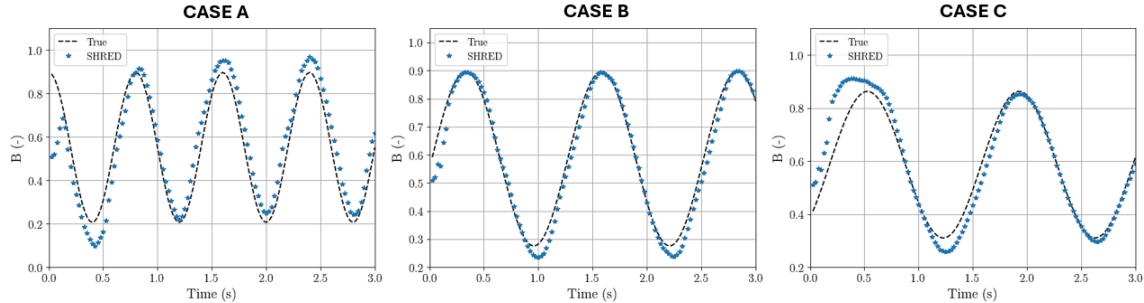


Figure. Temporal profile of the magnetic field reconstructed by SHRED, for three test cases. Max error 18% for case C at $t \sim 0.5$ s; stabilizing at around 2%. **The surrogate model can therefore reconstruct the temporal shape of the input.**

OTHER TEST CASES: TWO-PARAMETER CASE (FLOW PAST TWO CYLINDERS)

- **Parameters:** Magnitude field B and inclination angle α

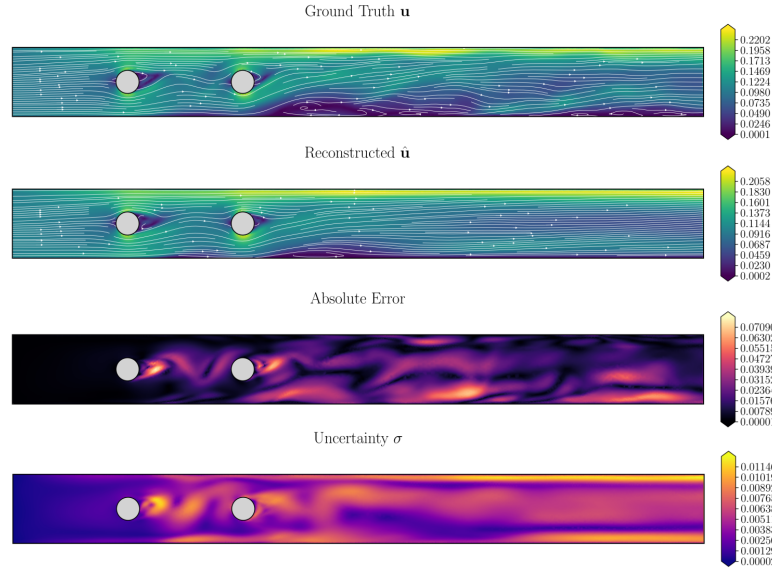


Figure. SHRED normalized reconstruction of the velocity field for test value $|B| = 0.3$ and $\alpha = 30$ at $t = 2.5$ s, together with the absolute error and ensemble uncertainty. The neural network effectively captures the formation of a single side layer on the top, accordingly to the specific inclination angle.

FURTHER WORKS

Other analyses have been carried out on SHRED for MHD applications (see <https://www.politesi.polimi.it/handle/10589/253211>: C. Scardino, *State estimation in magnetohydrodynamics: from interpolation to recurrent networks*, MsC Thesis, 2026 (Supervisors: S. Riva, A. Cammi, C. Introini)

- ▶ Testing SHRED in reconstruction, parametric and prediction mode (SHRED outperforms conventional surrogate techniques and Deep Learning models in reconstruction and parametric mode, comparable to math-based techniques in prediction mode) → **Along with tests on fission reactors, SHRED is confirmed as the best choice to build surrogate models for DTs.**
- ▶ SHRED as operator learning: SHRED approximates the inverse differential operators governing the system, such as the inverse Laplacian for the $T \rightarrow p$ mapping, not only fitting the data but capturing insights into the physics underlying the governing equations → **Compared to Deep Learning techniques, SHRED retains a certain degree of physical interpretability.**