A Novel Machine Learning-Based Digital Twin Architecture for Real-Time Monitoring, Integration and Control of the Breeding Blanket

A. Cammi [1] C. Introini [1] C. Tripodo [1] A. del Nevo [2] M. Eboli [2] R. Marinari [2] L. Savoldi [3]

[1]Energy Department, Nuclear Engineering Division, Politecnico di Milano, 20156, Milano (IT).
[2]ENEA Brasimone Research Center, Bacino del Brasimone 1, 40032 Camugnano (IT)
[3]MAHTEP Group, Dipartimento Energia "Galileo Ferraris", Politecnico di Torino (IT)
Contacts: antonio.cammi@polimi.it, carolina.introini@polimi.it

INTRODUCTION

- **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.

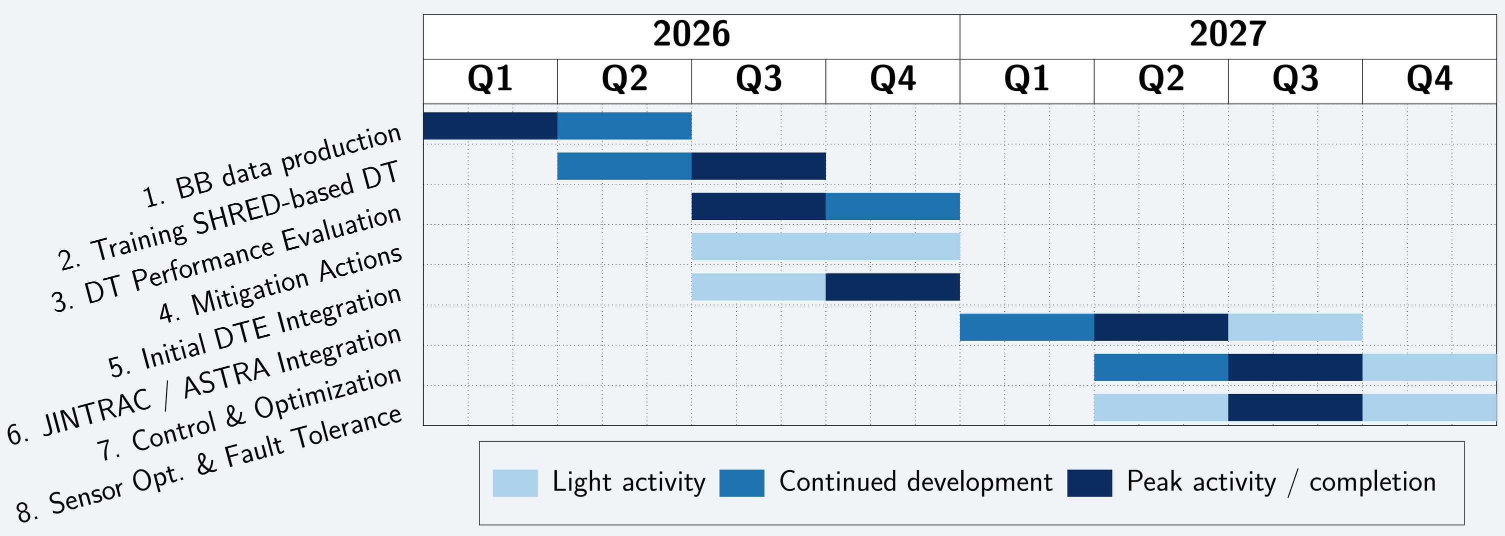
Build and prove the feasibility of a SHRED-based DT for the WCLL Breeding Blanket with reconstruction, prediction, forecast capabilities, to be integrated within larger DTE frameworks.

E-TASC 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.

PLANNED WORK

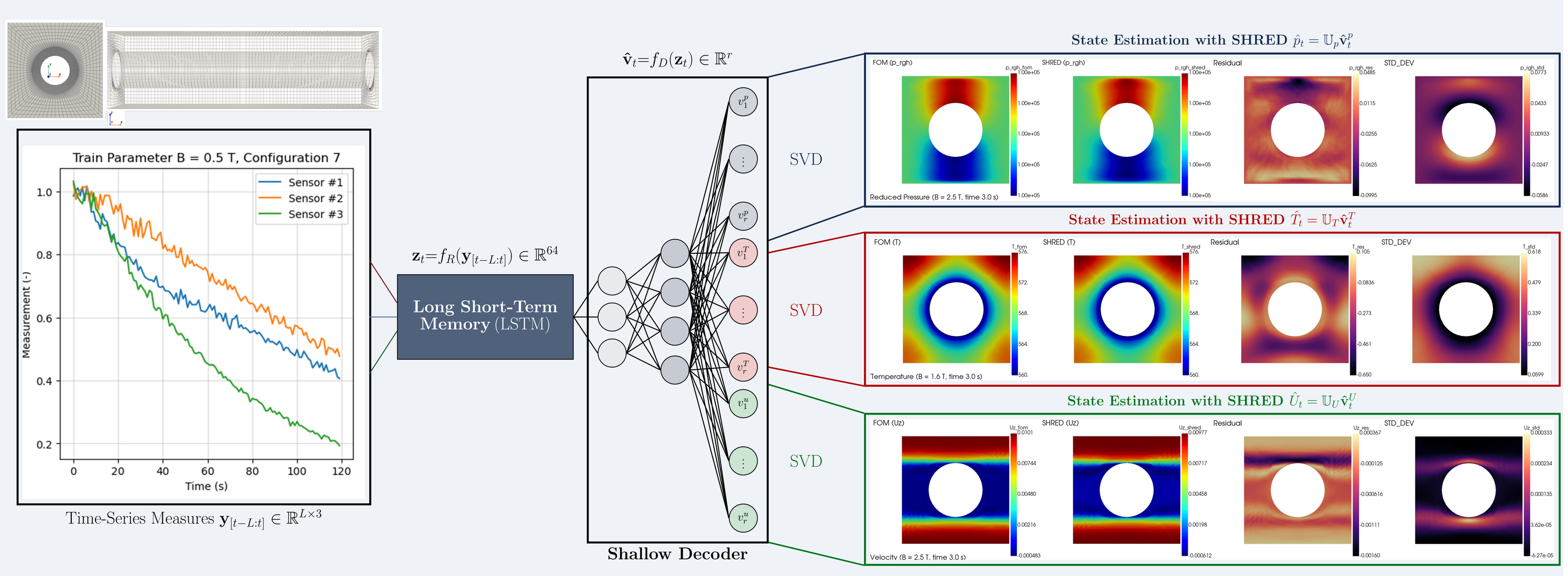
2026	**WP1**: Trained SHRED-based DT of the WCLL Breeding Blanket, using existing ENEA BB data for training, capable of operating in reconstruction, prediction and ensemble mode.	**Deliverable**: First scientific report, first main commit of the code with documentation and a representative benchmark 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.	



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.

ARCHITECTURE: SHRED



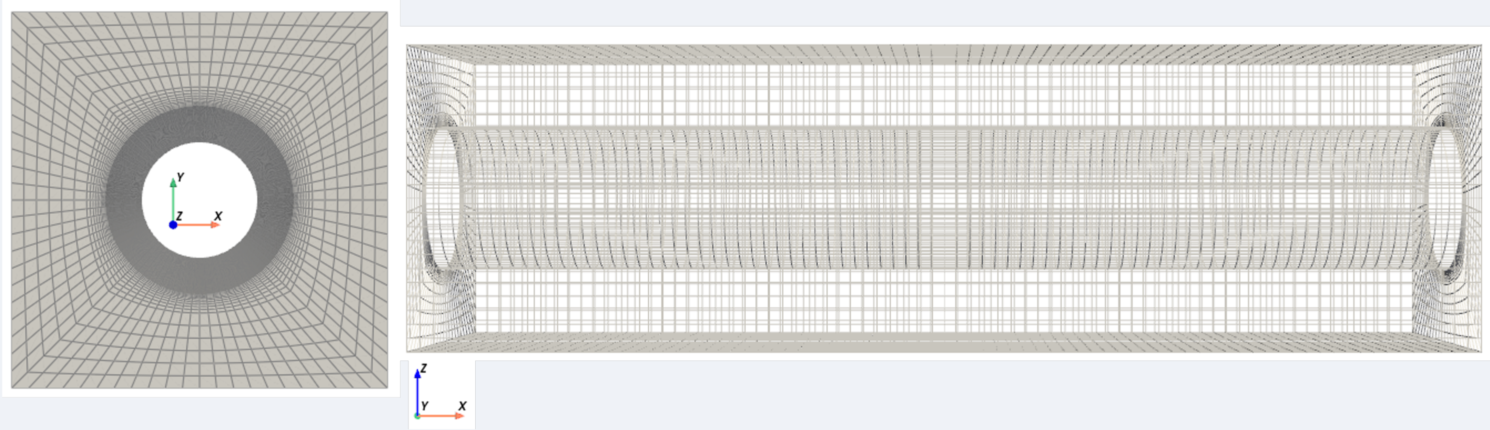
- **Selected ML Architecture**: SHRED (SHallow REcurrent Decoder).
- **TL;DR**: A Long Short-Term Memory (LSTM) network creates a latent temporal model $\mathbf{z}_t$ from some input measurements $\mathbf{y}^T(t)$, which a Shallow Decoder Network (SDN) maps to the reduced POD coefficients $\hat{\mathbf{v}}_t$, 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

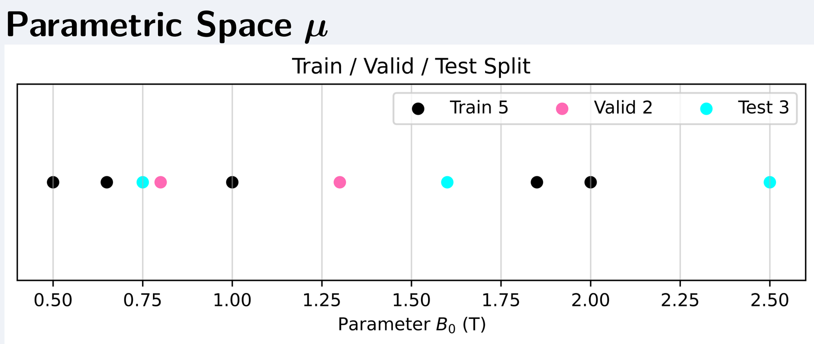
Data Efficient: Can operate with as few as three sensors, located anywhere in the domain.	**Resource Efficient**: Compressive training allows for laptop-level training and offline operation. Can operate in *Ensemble Mode*, providing an estimation of uncertainties and enabling faulty sensors detection.	**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.

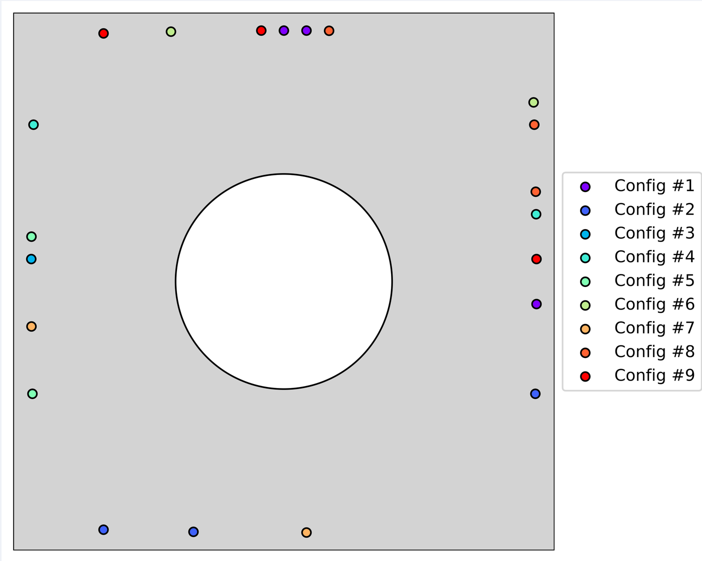
DEMONSTRATIVE BENCHMARK: 3D MHD CHANNEL

- **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`
- **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 ($r = 100 << \mathcal{N}_h$)
  - Ensemble mode: nine sensor configurations (three sensor each, randomly selected near the boundary)
  - **Parametric Space** $\boldsymbol{\mu}$





- **SHRED I/O**: Three (noisy) temperature measurements near the boundaries per each SHRED model → (Averaged) reduced POD coefficients $\mathbb{V}^k_{\mu_i} = (\mathbb{U}^k)^T \mathbb{S}^k_{\mu_i}$ → 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%).

KEY RESULTS

$\boldsymbol{\mu} = B_y$ (T)	$\hat{\boldsymbol{\mu}}$	$\epsilon\%$
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 1. 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.**

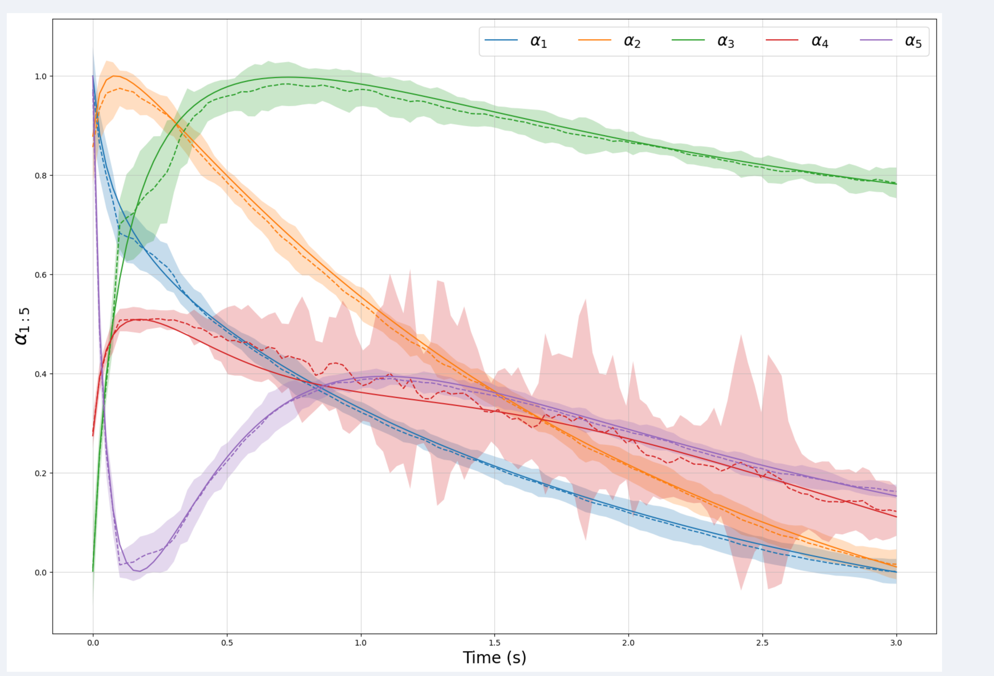


Figure 1. First five POD coefficient for unobservable field U_z, capturing the temporal dynamics, for test parameter $B_y = 2.5$ T. True value is continued line, SHRED average prediction is dashed line, band is estimation uncertainty. **The temporal behaviour is correctly captured (within the uncertainty band) even for the extrapolated case), and even without any direct observations on the velocity field.**