Learning for Power System Dynamics: The Generalization Challenge

Dr. Jochen L. Cremer

Associate Professor

https://www.Jochen-Cremer.com

4th GridFM Workshop
Aachen, Germany

TUDelft DAIE ENER ERGY

https://www.tudelft.nl/ai/delft-ai-energy-lab

Credits & team

Olayiwola Arowolo

Jochen Stiasny

Maosheng Yang

Haiwei Xie

Mert Karaçelebi

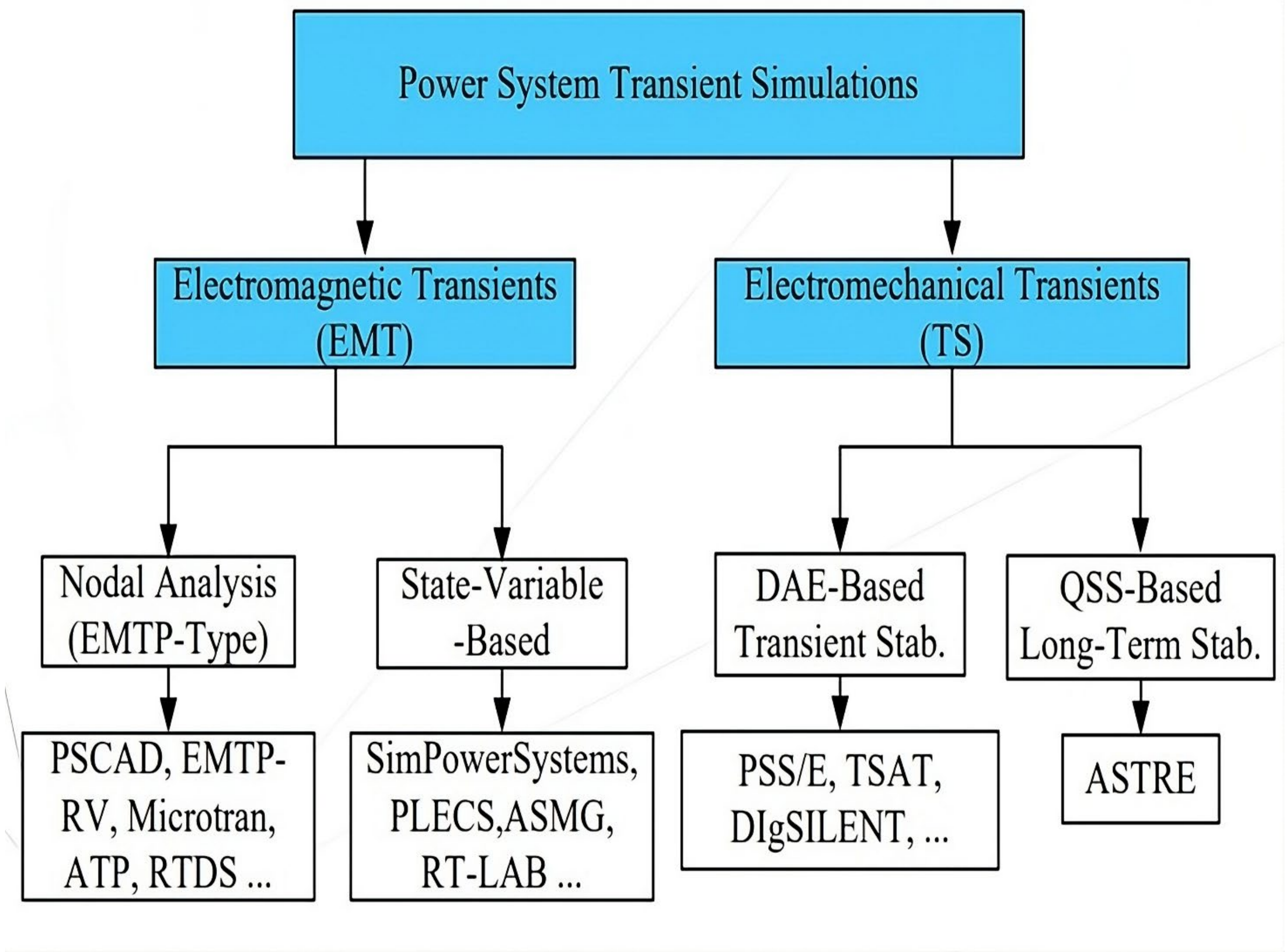
Al-Amin Bugaje

Federica Bellizio

Goran Strbac

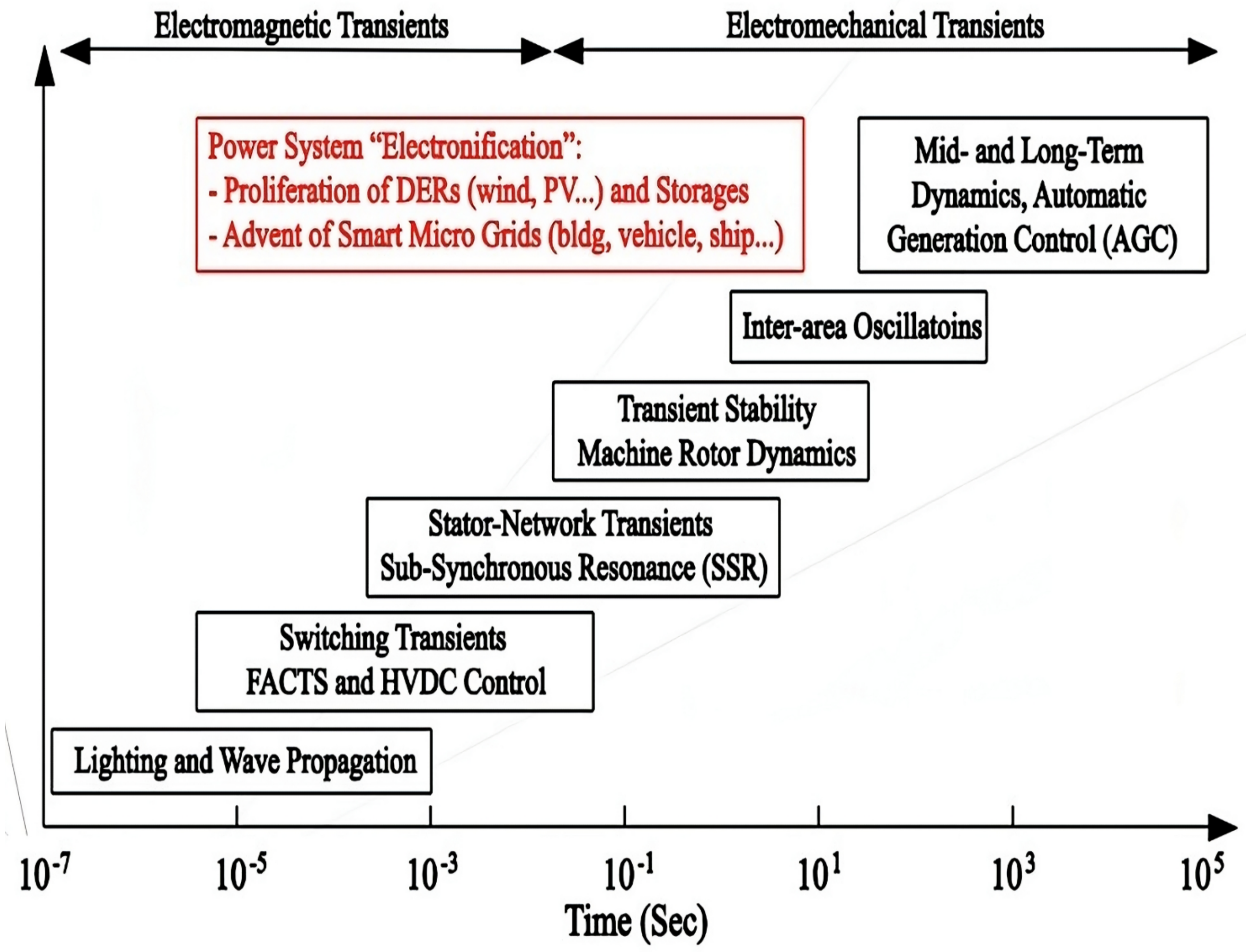
Power System Transient Simulation

- Transient simulation is used to assess the dynamic security of the power system to various contingencies
- Different simulation tools exist; the tool of choice depends on the phenomena being investigated
- Simulation tools are mature; in development for the last 40 years



EMT versus RMS simulation

- Electromechanical transient (RMS) simulations are for slower dynamics
- Slower dynamics are dominant in conventional power systems
- EMT simulations are for faster dynamics and switching transients
- Fast dynamics from inverter switching are increasingly dominant in low-inertia systems



Transient Simulation Bottleneck

Operating Condition/Contingency

Scenario A

Scenario B

⋮

Scenario C

Simulation tool

PSS/E / PSCAD

Speed bottleneck

Operator/ Decision maker

Security Status

TU Delft | AI Initiative | DAIE ENERGY

Supervised Learning for Dynamics Surrogate

Notation: Power system s, model m, parameter x

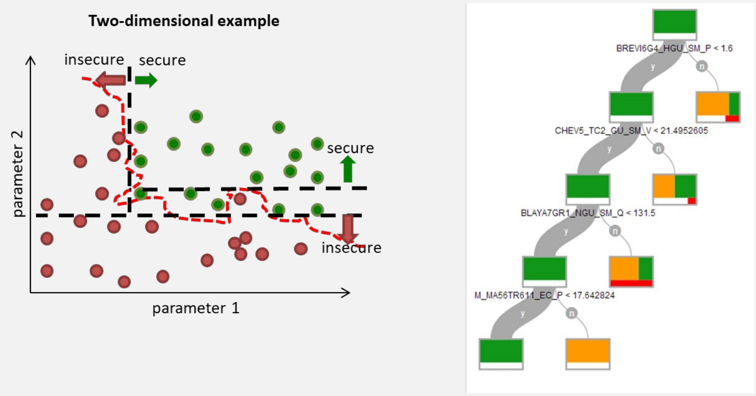
Objective: assess $m(x) \rightarrow y$ **very** fast and **often**

Surrogate approach
1.Generate a training dataset $\Omega^T = \{(x_i, y_i)\}_{i=1}^N$ where $y_i = m(x_i)$ from the full simulator
2.Train surrogate $f(x) \rightarrow \hat{y}$ with supervised loss $\sum_{i \in \Omega^T} \|y_i - \hat{y}_i\|$
3.Use $f(x_j)$ for new j $\notin \Omega^T$

Benefit: speed at inference

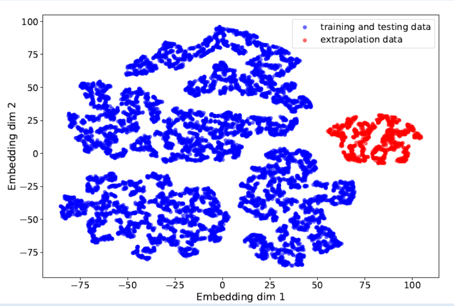
Applications

- Real-time dynamic security assessment ([1,2] and many others)



Challenges

- Out of distribution risks: What if s and m changes? e.g., topology changes
- What if the model is inaccurate $s \neq m$? e.g., inverter-based controls
- Need large, representative training data



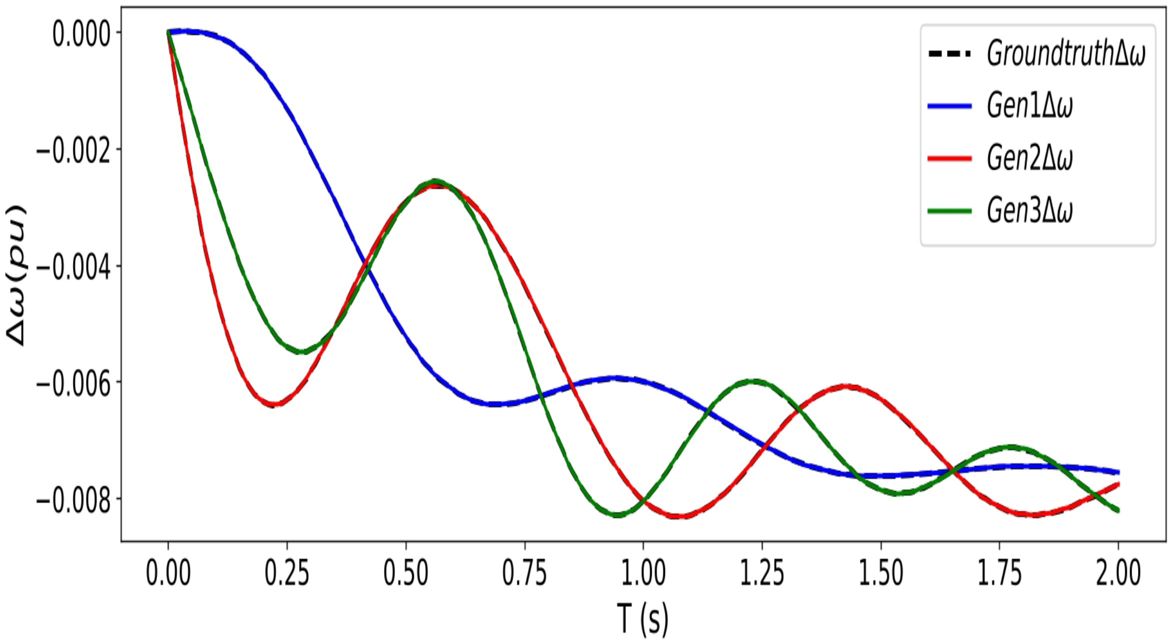
TU Delft | AI Initiative | DAI ENERGY

[1] Wehenkel, Louis, and Mania Pavella. "Decision trees and transient stability of electric power systems." *Automatica* 27.1 (1991): 115-134.
[2] Cremer, Jochen L., Ioannis Konstantelos, and Goran Strbac. "From optimization-based machine learning to interpretable security rules for operation." *IEEE Transactions on Power Systems* 34.5 (2019): 3826-3836.
[3] Olayiwola Arowolo, Jochen Stiasny, Jochen L. Cremer, "Exploring Extrapolation of Machine Learning Models for Power System Time Domain Simulation", *Bulk Power System Dynamics and Control Symposium, 2025*

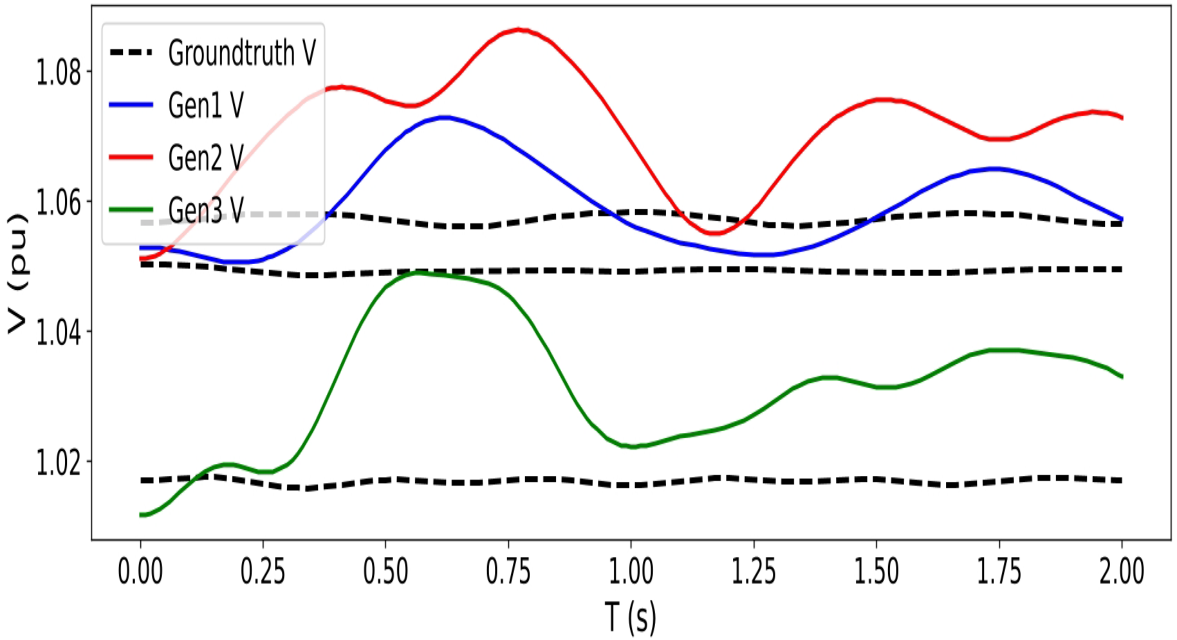
Extrapolation of ML Models in Transient Simulation

Olayiwola
Arowolo

Model performs well for continuous disturbances within the training data distribution



Model fails to extrapolate for OOD discrete disturbances.



TU Delft | AI Initiative | DAI ENERGY

[3] Olayiwola Arowolo, Jochen Stiasny, Jochen Cremer, “Exploring the Extrapolation Performance of Machine Learning Models for Power System Time Domain Simulations”, 12th Bulk Power System Dynamics and Control Symposium and Sustainable Energy, Grids and Networks Journal, 2025

Physics-Informed Learning

Objective: surrogate learning enhanced with physics knowledge from model m

Idea: Incorporate physics residual (e.g. from a PDE or simulator) to guide learning and improve generalization

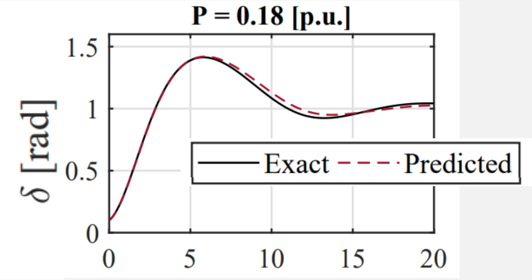
Physics-informed approach

1. Generate offline training dataset $\Omega^T = \{(x_i, y_i)\}_{i=1}^N$ with $y_i = m(x_i)$
2. Train surrogate $f(x) \to \hat{y}$ on composite loss $\sum_{i \in \Omega^T} \|y_i - \hat{y}_i\| + \mathcal{L}_{phys}(f(x_i), m)$
3. Use $f(x_j)$ for new j $\notin \Omega^T$

Benefits: Better generalisation performance with **fewer training samples**

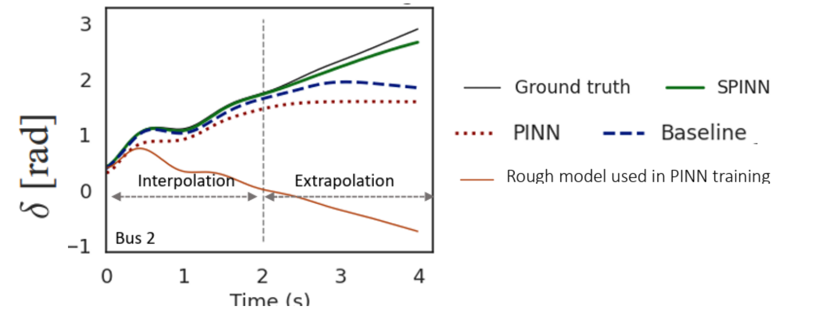
Applications

- Extrapolation in time-domain for dynamic analysis in power systems



Challenges

- Model inaccuracy $s \neq m$
- Changes in s or m
- Data sparsity
- Multi-loss scaling causes training instability
- Scaling issues to many physical loss terms in power systems

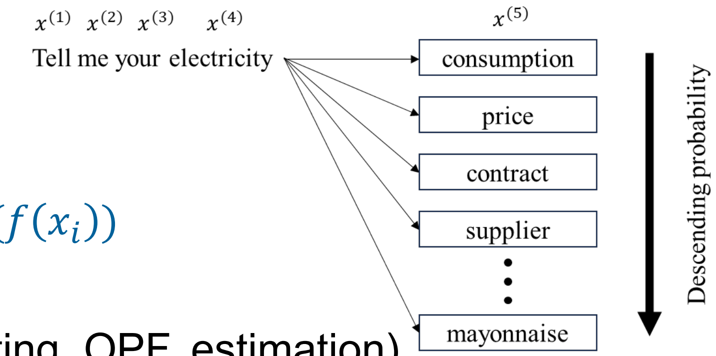


[4] Misyris, George S., Andreas Venzke, and Spyros Chatzivasileiadis. "Physics-informed neural networks for power systems." In *2020 IEEE power & energy society general meeting (PESGM)*, pp. 1-5. IEEE, 2020.
[5] Jochen Stiasny, Spyros Chatzivasileiadis, Physics-informed neural networks for time-domain simulations: Accuracy, computational cost, and flexibility, Electric Power Systems Research, Volume 224, 2023
[6] Xie, Haiwei, Federica Bellizio, Jochen L. Cremer, and Goran Strbac. "Regularised Learning with Selected Physics for Power System Dynamics." In *2023 IEEE Belgrade PowerTech*, pp. 1-7. IEEE, 2023.

Self-Supervised Learning

Objective: Learn a **useful internal representation** from unlabeled data by solving a **pretext task** — no human-labeled or simulator-labeled outputs required.

Idea: instead of training on (x_i, y_i) train on auto-generated pseudo-labels or tasks constructed from structure x_i



Approach

1. Generate many inputs $\Omega^T = \{(x_i)\}_{i=1}^N$
2. Define self-supervised pretext loss $\mathcal{L}_{pretext}(f(x_i))$
3. Train encoder $\sum_{i\in\Omega^T} \mathcal{L}_{pretext}(f(x_i))$
4. Use $f(x)$ for downstream *task* (e.g. forecasting, OPF, estimation)

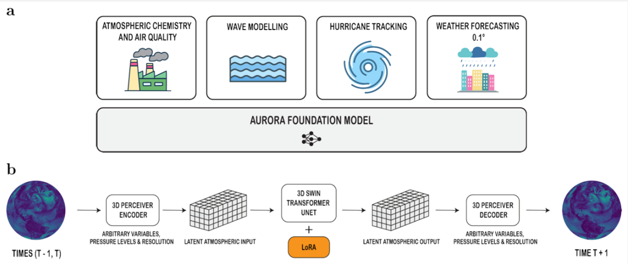
Benefits: Good initialization when little data, good transfer to downstream tasks

Challenges

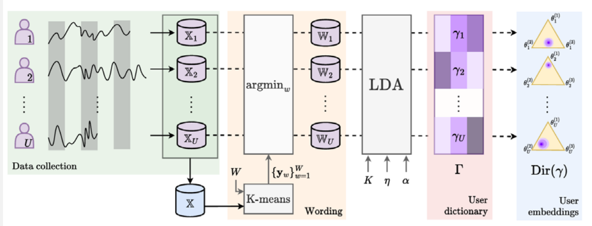
- Design pretext loss and model architectures with broad set of tasks, grid conditions, topologies
- Generate large data sets
- …

Applications

- Natural Language Processing
- Weather foundational models
- Earth system foundational models [9]



Load forecasting of users [10]



Grid foundation models (GFM) [11]

TU Delft | AI Initiative | DAI ENERGY

[9] Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Vaughan, A., Brandstetter, J., ... & Perdikaris, P. (2024). A foundation model for the earth system. *arXiv preprint arXiv:2405.13063*.
[10] Bölat, Kutay, and Simon Tindemans. "GUIDE-VAE: Advancing Data Generation with User Information and Pattern Dictionaries." *arXiv preprint arXiv:2411.03936* (2024).
[11] Hamann, H. F., Gjorgiev, B., Brunschwiler, T., Martins, L. S., Puech, A., Varbella, A., ... & Sobolevsky, S. (2024). Foundation models for the electric power grid. Joule, 8(12), 3245-3258.

Opportunities for Transient Simulation Foundation Model

Accelerating transient simulation with GridFM can unlock new use cases with more comprehensive scenario assessments

Potential use cases:

- Planning and commissioning HVDC/FACT devices
  - Investigating interaction phenomena between HVDC and the rest of the grid
  - Testing controllers and software updates
  - Verification with onsite measures
- Investigating inter-area oscillations
  - Getting common in weak grids
  - Root cause of some of these events still unknown
  - Sub-synchronous oscillations from controller interactions need EMT simulations

Non-conventional use cases:

- How to expand the system that maximises transient stability for k-faults?
- How to operate the system secure against transients?

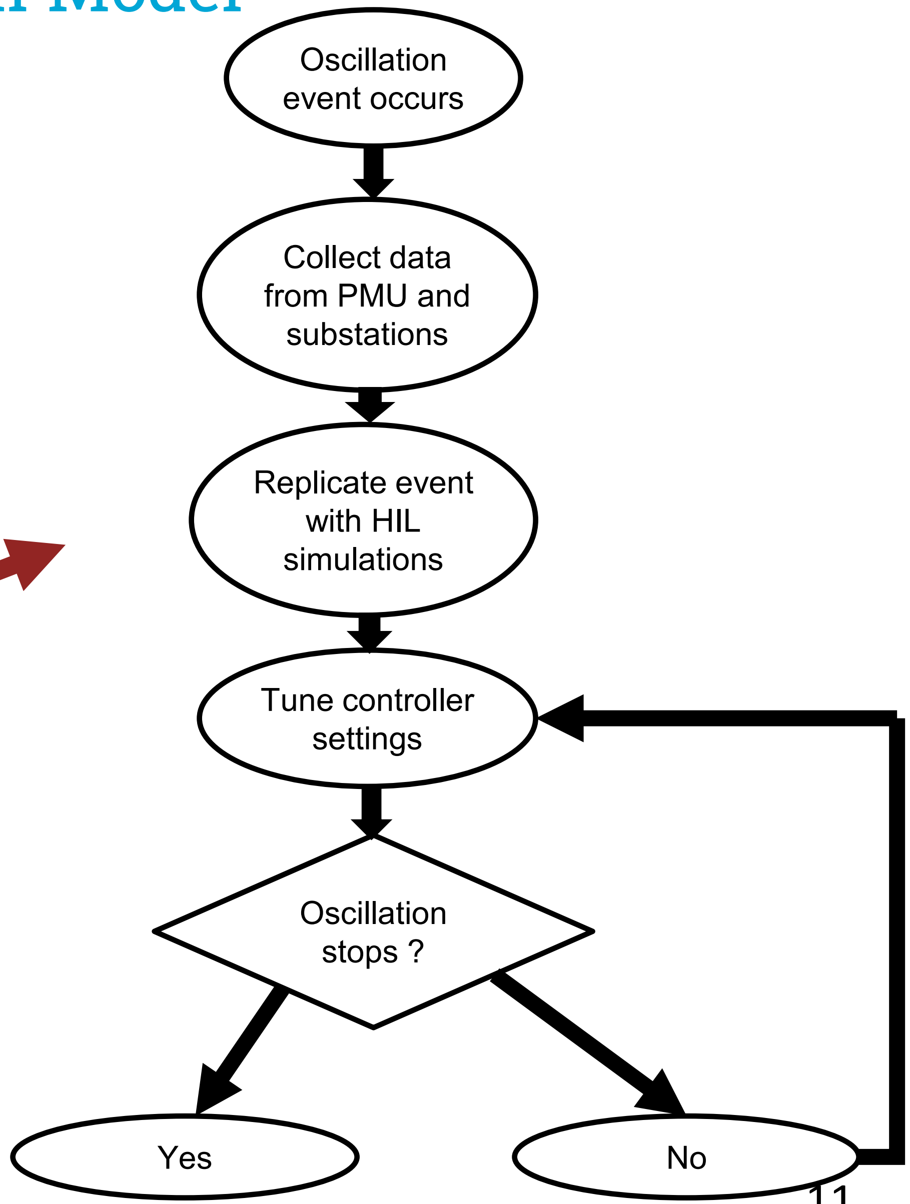
[12] Bastien Giraud, Ali Rajaei, Jochen L. Cremer "Constraint-Driven Deep Learning for N-k Security Constrained Optimal Power Flow", *Electric Power System Research and 2024 IEEE Power System Computation Conference*
[13] Mert Karaçelebi, Jochen L. Cremer, "Power system frequency monitoring and emergency control with neural ordinary differential equations", *12th Bulk Power System Dynamics and Control Symposium and Sustainable Energy, Grids and Networks Journal, 2025*

Opportunities for Transient Simulation Foundation Model

Integration of offshore wind farms

- ➢ Require detailed models of turbines and controllers
- ➢ Assessment of background harmonic amplification
- ➢ Locating sources of sub-synchronous oscillations

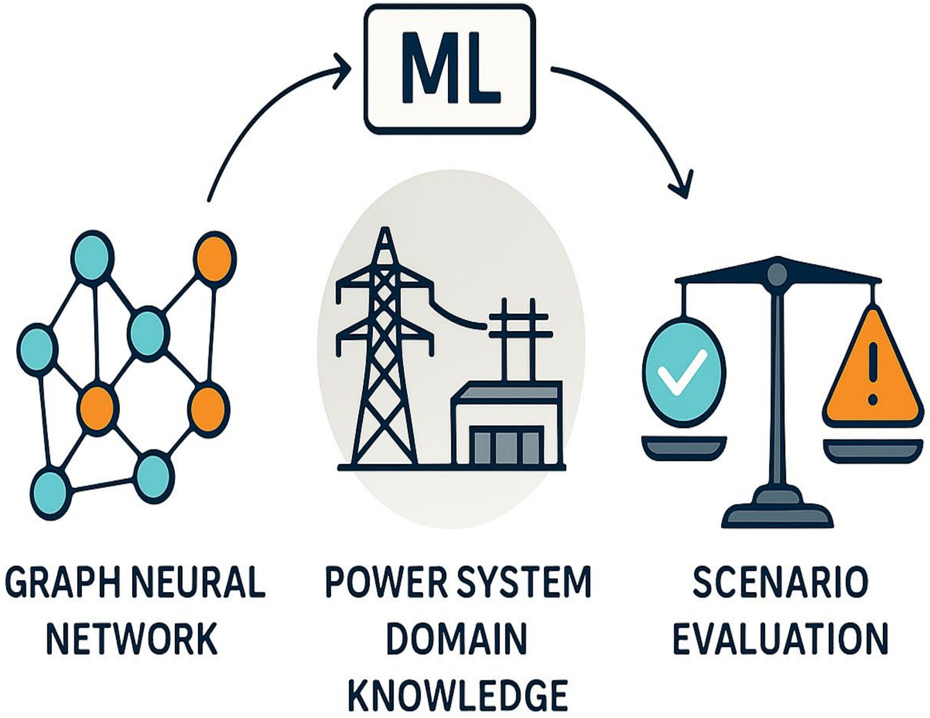
The existing setting is time-consuming



TU Delft | AI Initiative | DAI ENERGY

Approaching a Foundation Model for Transient Simulation

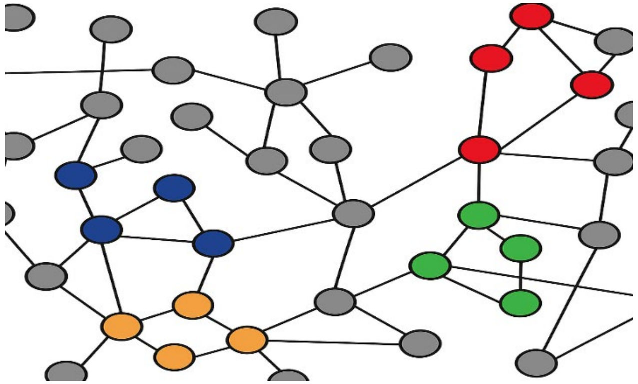
- Potential for graph-based modelling of power system transients
- Graphs can easily handle discrete changes to system topology
- EMT and RMS simulations may be unified by simulating a system of DAEs.
- The best way to formulate an appropriate graph is an open question



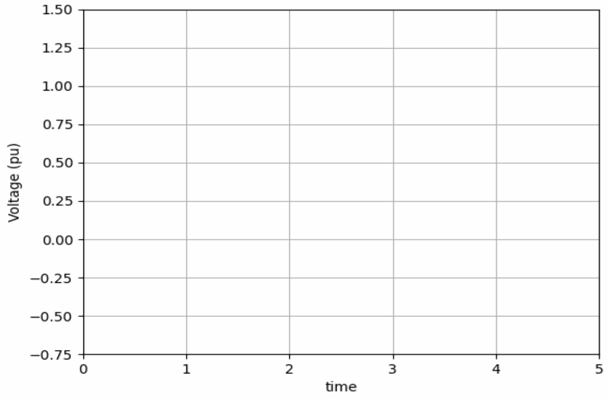
Challenges

- Large-scale simulation training dataset
  - ➢ Data variety (synthetic and Real)
  - ➢ Data validity
  - ➢ Data privacy
- Modelling challenges
  - ➢ Architectural definition
  - ➢ Computational cost of pretraining
  - ➢ Consistency with physical laws
- Application / Validation
  - ➢ Experimental or Physical validation

Graph input



Simulation output

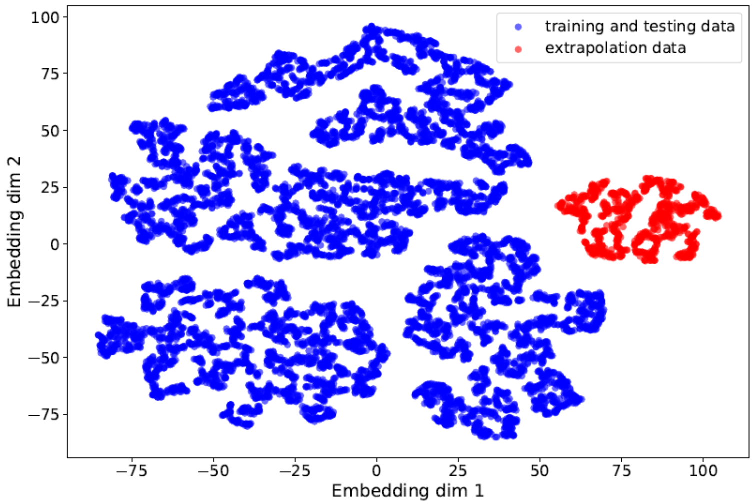


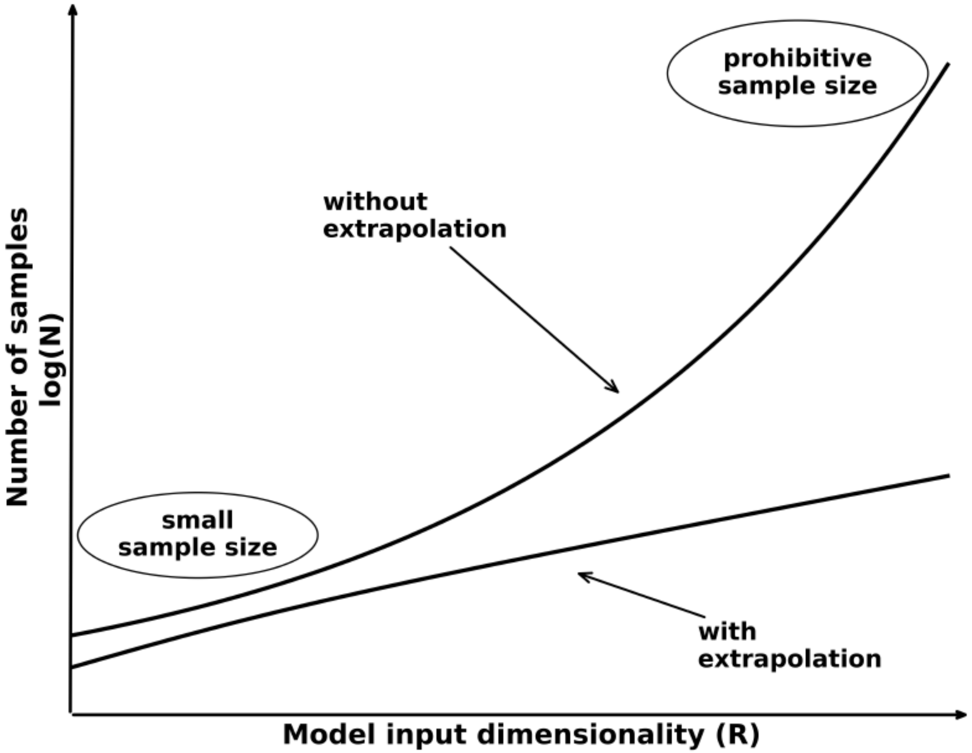
[14] Al-Amin Bugaje, Jochen L. Cremer, Goran Strbac ”Generating Quality Datasets for Real-Time Security Assessment: Balancing Historically Relevant and Rare Feasible Operating Condition” *International Journal of Electrical Power & Energy Systems, 2023*

[15] F. Bellizio, J. L. Cremer, G. Strbac, *"Machine-learned security assessment for changing system topologies." International Journal of Electrical Power & Energy Systems 134. 2022: 107380.*

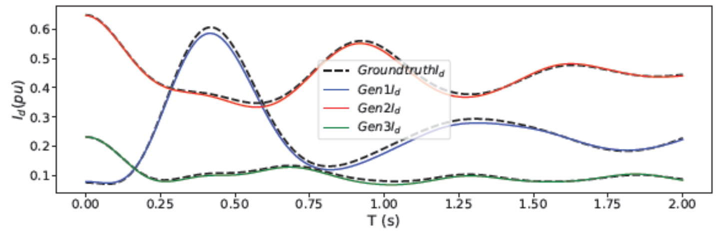
Generalisation to changes in s or m

The model performs well not just on training data, but on ***unseen scenarios*** *— new grid states, topologies, contingencies, or time horizons.*



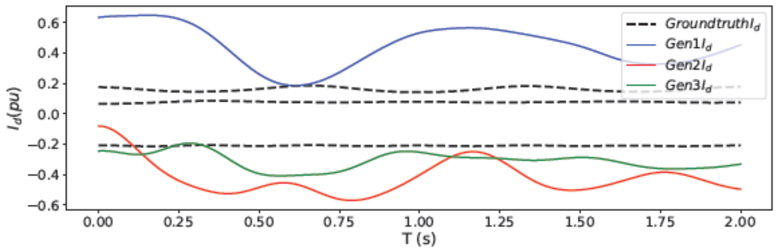


Extrapolation in continuous domain



(a) I_d current trajectory

Extrapolation in nonlinear domain (discrete)



(a) I_d current trajectory

TU Delft | AI Initiative | DAIE ENERGY

AI·EFFECT DRIVING ENERGY TRANSITION THROUGH AI EXPERIMENTATION

[3] Olayiwola Arowolo, Jochen Stiasny, Jochen Cremer, "Exploring the Extrapolation Performance of Machine Learning Models for Power System Time Domain Simulations", *12th Bulk Power System Dynamics and Control Symposium and Sustainable Energy, Grids and Networks Journal, 2025*

Power Flow is at the heart of many power system tasks
How should we formulate a foundational PF?

Jochen Stiasny

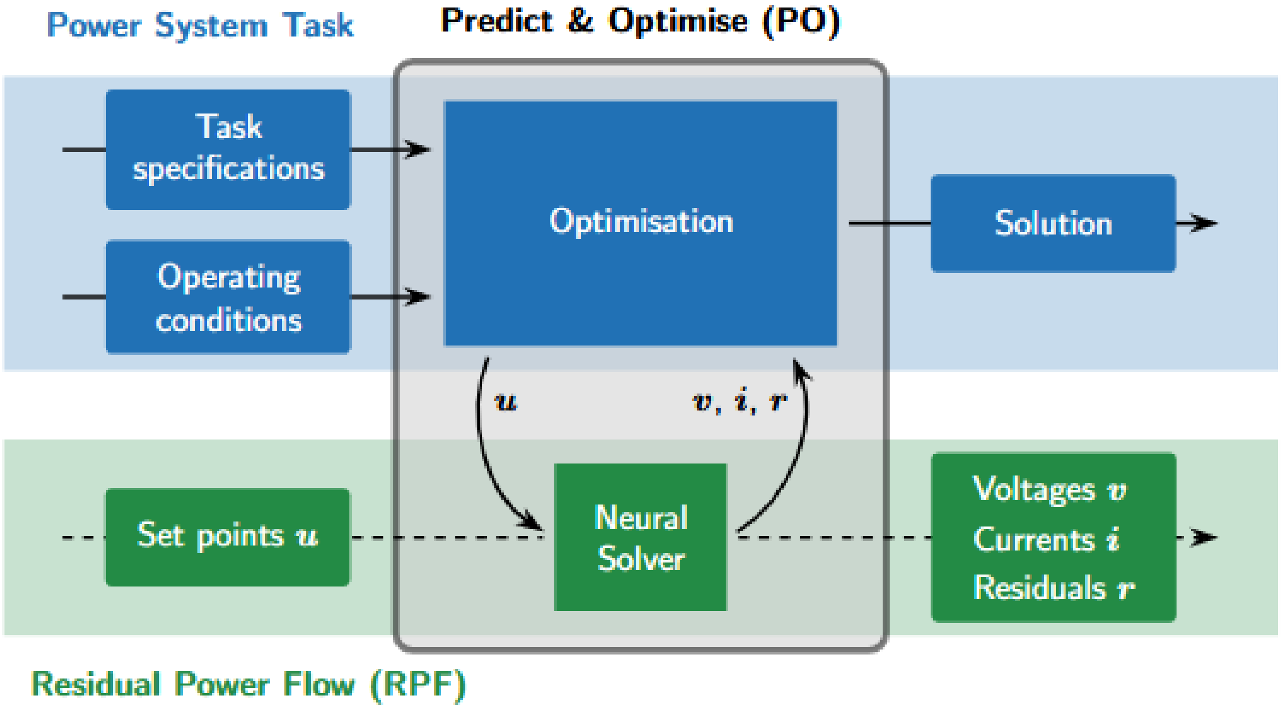
Residual power flow (RPF)

- RPF quantifies infeasibility
- Simpler formulation for neural solvers

Predict-and-Optimise (PO)

- Flexible handling of constraints and objectives
- While minimising infeasibility

→ Preprint soon available



Model inaccuracy $s \neq m$ *("data quality")*

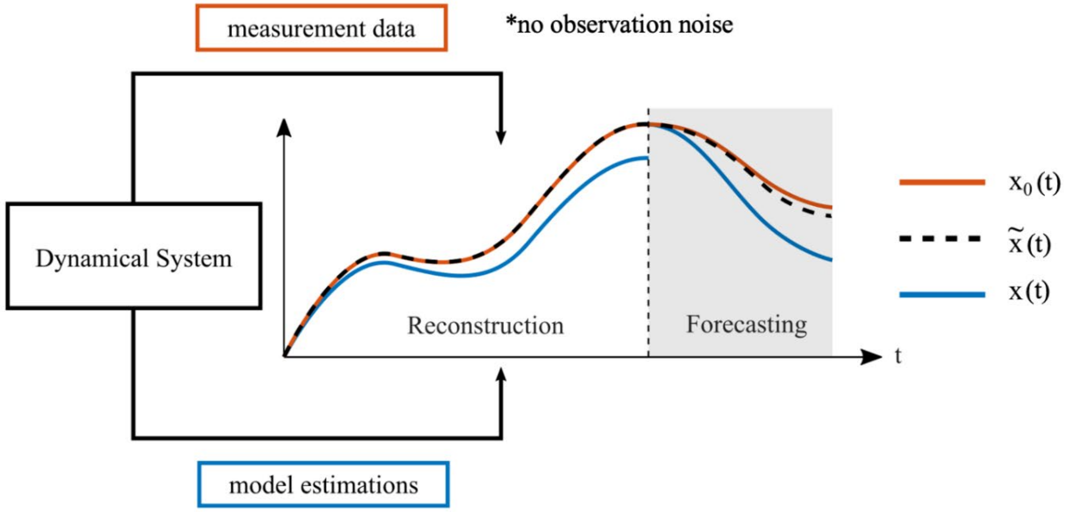
"All models are wrong, but some are useful", George E. P. Box

Example challenges

- Distribution: Inaccurate transformer-tap positions
- Transmission: Converter-based control models are unkown

Sim-to-Real Domain Adaptation

Possible techniques: Parameter estimation to develop probabilistic and deterministic models, discrepancy learning



[16] Ebers, M. R., Steele, K. M., & Kutz, J. N. (2024). Discrepancy modeling framework: Learning missing physics, modeling systematic residuals, and disambiguating between deterministic and random effects. *SIAM Journal on Applied Dynamical Systems*, *23*(1), 440-469.

Conclusion: We have work to do

Data Generation and Synthesis

- Synthetic data generation
- Data integration from real power systems
- Data preprocessing

Model Development

- Defining pretraining task
- Representation learning
- Multi-timescale modelling

Model Validation

- Fine-tuning the pretrained model on different simulation tasks
- Physical validation with hardware-in-the-loop
- Uncertainty quantification

Thank you for your attention

Dr. Jochen L. Cremer

Associate Professor

https://www.Jochen-Cremer.com

TUDelft DAIE ENERGY

https://www.tudelft.nl/ai/delft-ai-energy-lab

More references

- Jatskevich, J. (2023, May 25). Challenges in Transient Simulation of Fast-Evolving Energy Systems, Imperial College London.
- Olayiwola Arowolo, Jochen Stiasny, Jochen Cremer, Exploring extrapolation of machine learning models for power system time domain simulation, Sustainable Energy, Grids and Networks, Volume 43, 2025.
- Jochen Stiasny, Spyros Chatzivasileiadis, Physics-informed neural networks for time-domain simulations: Accuracy, computational cost, and flexibility, Electric Power Systems Research, Volume 224, 2023
- T. Cheng, R. Chen, N. Lin, T. Liang and V. Dinavahi, "Machine-Learning-Reinforced Massively Parallel Transient Simulation for Large-Scale Renewable-Energy-Integrated Power Systems," in *IEEE Transactions on Power Systems*, vol. 40, no. 1, pp. 970-981, Jan. 2025
- Zhao, J., Zhu, Z., Galkin, M., Mostafa, H., Bronstein, M., & Tang, J. (2025). Fully-inductive Node Classification on Arbitrary Graphs. *arXiv [Cs.LG]*. Retrieved from http://arxiv.org/abs/2405.20445
- Ye, Z., Huang, X., Chen, L., Liu, H., Wang, Z., & Dong, B. (2025). PDEformer: Towards a Foundation Model for One-Dimensional Partial Differential Equations. *arXiv [Math.NA]*. Retrieved from http://arxiv.org/abs/2402.12652
- Bodnar, C., Bruinsma, W.P., Lucic, A. *et al.* A foundation model for the Earth system. *Nature* **641**, 1180–1187 (2025)
- Das, A., Kong, W., Sen, R., & Zhou, Y. (2024). A decoder-only foundation model for time-series forecasting. *Proceedings of the 41st International Conference on Machine Learning*. Vienna, Austria: JMLR.org.