Eidgenössisches Departement für Umwelt,
Verkehr, Energie und Kommunikation UVEK
Bundesamt für Verkehr BAV
Abteilung Politik
Sektion Direktionsgeschäfte
Mühlestrasse 6, 3063 Ittigen,

Lausanne, 19.11.2025

Final Progress Report «Project INITIATE»

1 Introduction

This is the final progress report for Project INITIATE, a research initiative conducted from 2021 to 2025. Project INITIATE was funded by the funding scheme for railway infrastructure "Bahninfrastrukturfonds" BIF of 2014. In this project, EPFL and SBB worked together to improve the detection and prediction of critical conditions in railway power network systems by developing intelligent algorithms for control system data.

The project was initially started at ETH Zurich in 2021 and was transferred to EPFL in March 2022 due to the transfer of the research team from ETHZ to EPFL. In 2024, the project was extended for the duration of 2025.

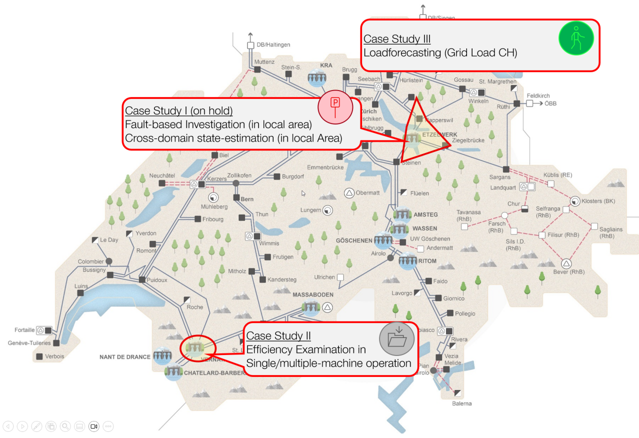

Figure 1: Power grids sites from which the generated data of case study I, II and III originates.

In June 2021, a group of domain experts from different fields collected and rated various case studies as candidates for investigation. The domain experts estimated the highest overall benefits of all case studies as well as immediate benefits if selected as the first case study. Based on these scores, two studies were selected to be the most promising and, thus, were processed first: a case study on state estimation and fault detection from power grid sensors in local areas such as power plants with connected substations (**case 1**) and a case study on hydro-power plant efficiency estimation from hydro and electrical sensors (**case 2**). In 2023, case study 1 was temporarily halted to focus on a new case study, load forecasting in traction power grids (**case 3**) based on findings from case study 1.

While the original proposal directly aimed at predicting component failures or sensor drifts, sufficiently large and high-quality failure datasets were not available during the project to reliably validate detection accuracy. As a result, the project's case studies focused on advancing methodological capabilities for accurate efficiency and state estimation, which form a critical foundation for residual-based anomaly detection. The remaining step of setting the threshold on the residual is straightforward and can be readily implemented even by non-experts.

During the final reporting period, the focus was on finalizing case study 1 and disseminating key findings from Project INITIATE through publications. The model from case study 1 was refined, tested on a novel dataset with higher temporal resolution, and subsequently, the manuscript summarizing this research was submitted to a journal. The new publication puts an emphasis on real-time state estimation using our newly proposed Heterogeneous Graph Attention Network (HGAT). Furthermore, following the peer review feedback, the publication on case study 3 was revised and extended with a building energy data set to demonstrate the broader applicability of the proposed contextual forecasting method.

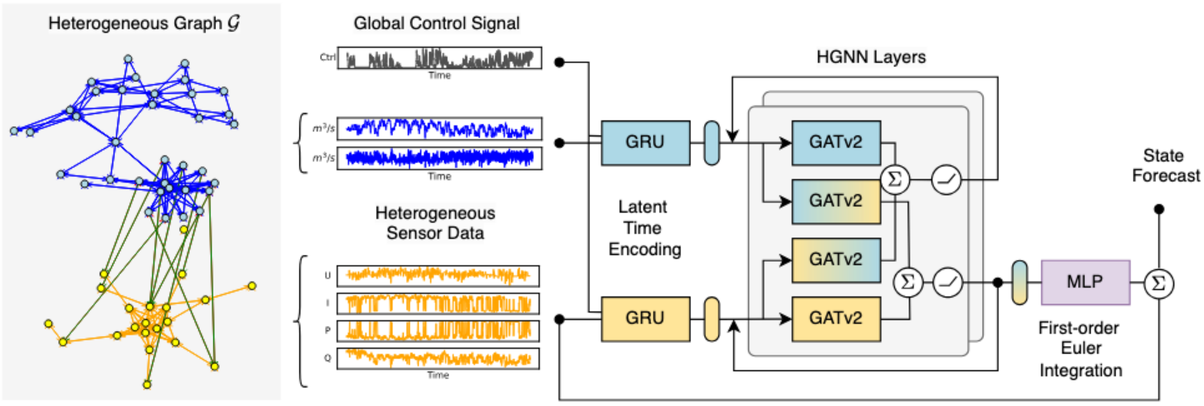

Figure 2: An overview of the processing steps of our proposed Heterogeneous Graph Attention Network (HGAT), applied to the case study of a pumped-storage hydropower plant. By operating on a heterogeneous graph, the HGAT efficiently extracts information from the hydraulic and electrical sensor data to forecast the electrical state variables.

2 State Estimation from Power Grid Sensors (Case 1)

The 2022 INITIATE report called for a systematic study of Graph Neural Networks (GNNs) for sensor data fusion in hydropower plants. Our preliminary work, a proof of concept presented at the PHME 2024 conference [1], demonstrated the efficiency of sensor data fusion for short-term state forecasting in a pumped-storage hydropower (PSH) plant. However, the published Spectral-Temporal Graph Neural Network (STGNN) had two primary limitations: First, it derived the topology solely from univariate sensor data, which failed to account for the actual grouping of installed sensors. Second, validation was restricted to a single dataset with a coarse 1-minute temporal resolution, making it unsuitable for real-time state estimation demanded by many practical applications.

Research Objective: consequently, our subsequent efforts focused on exploring the high-resolution dataset collected during the 2024 reporting period, at the second level, to enable real-time sensor fusion and improved state estimation across power plant subsystems. The main objective was to develop an enhanced model capable of handling all temporal resolutions and to reintroduce the asset and sensor topology.

To account for the grouping of installed sensors and to reintroduce the topology, we explored Heterogeneous Graph Neural Networks (HGNN) in the 2024 reporting period. HGNN are a specialized form of Graph Neural Network capable of handling graphs with multiple node and edge types, each with distinct features. HGNN can simultaneously learn homogeneous relationships (within the same subsystem) and heterogeneous relationships (between different subsystems). However, the HGNN performance was limited on the high-resolution data set. Therefore, during this reporting period, we extended the previous STGNN architecture for power grid state forecasting to develop a model that can operate at a one hertz temporal resolution while integrating the asset and sensor topology.

Heterogeneous Graph Attention for Hydropower Plants: To address this research objective, we introduced Heterogeneous Graph Attention Network (HGAT) to integrate sensor data from the electrical and hydraulic subsystems. Like HGNN, HGAT are a specialized type of Graph Neural Networks designed to handle heterogeneous graphs. The HGAT architecture learns relation-specific message-passing functions across node and edge type pairs, enabling context-aware forecasting of system dynamics by leveraging attention mechanisms over sensor and control data. Furthermore, HGAT uses the attention mechanism to weigh the importance of neighboring nodes and edges, which enables the model to focus on the most informative relationships during message passing. HGAT can simultaneously learn from homogeneous relationships (within the same subsystem) and heterogeneous relationships (between different subsystems). To better handle varying sensor sampling rates, we incorporate a forecasting mechanism based on first-order forward differentiation, which improves the model's ability to generalize across temporal scales, ensuring that the model is compatible with both raw measurements at second resolution and the minute resolution aggregates commonly used by SBB for monitoring. These steps allow the model to capture the rapid transient behaviors of monitored subsystems at the 1-second resolution. As a core contribution, we propose a modular architecture for HGAT comprising three complementary components, designed to work either independently or in combination for enhanced state forecasting performance as illustrated in Figure 2.

- The proposed method begins by augmenting the system state with first-order forward differences, explicitly capturing short-term dynamics. These enriched representations are then passed through a gated recurrent unit (GRU)-based encoder, which models temporal dependencies and generates latent states for each sensor node.

- To incorporate spatial and relational context, the latent system state representations are refined in the subsequent module through attention-based message passing over a heterogeneous graph, enabling the model to account for both intra- and inter-modality interactions.
- Finally, the model predicts finite differences from the refined node embeddings, which are then integrated using Euler's method to estimate the next system state - resulting in a forecasting approach that is both dynamically informed and guided by structural priors.

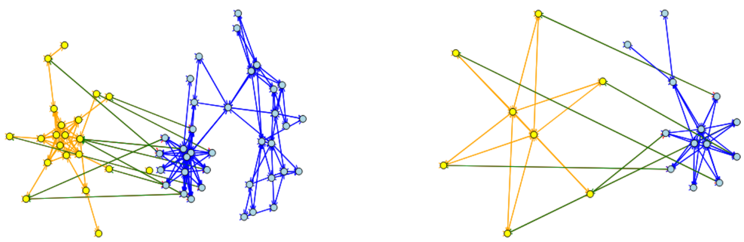

Figure 3: The heterogeneous graph used for message passing derived from schematical drawings of the hydropower plant.

As of November 2025, we have successfully finalized and submitted a publication for this work, covering this initial case study (Case 1). The publication confirms that integrating electrical and hydraulic data significantly enhances short-term state forecasting accuracy.

Heterogeneous Graph Construction: The HGAT model uses the same rule-based heterogeneous sensor graph construction as the previously proposed HGNN, which enables a direct path for comparing model performance. The graph is constructed by associating nodes with groups of sensors. We define an edge wherever a direct physical interaction exists between two sensor nodes. Therefore, in the pumped-storage hydropower plant considered in this case study, the hydrological network captures interactions such as the basin level of the dam influencing the water flow in the downstream pipe. For the electrical network, we consider the placement of phasor measurement units (PMUs) within the asset's single-line diagram. For heterogeneous connections, we establish connections between sensor nodes located near the generator that electromechanically integrate the two systems. The resulting graph as illustrated in Figure 3, serves as the computational graph for the results.

Results and Discussion: In contrast to the previous report, which was limited to current signals, we evaluate the full state PQUI - representing Active Power (P), Reactive Power (Q), Voltage Magnitude (U), and Current Magnitude (I) - in predicting the highly dynamic currents within the electric subsystem. The results, shown in Figure 4, demonstrate the superior performance of the Heterogeneous Graph Attention Network (HGAT) for short-term state forecasting. Specifically, in terms of Normalized Root Mean Square Error (NRMSE), HGAT significantly outperforms all evaluated baselines, including sequential models (LSTM and 1D-CNN), node-wise baselines (LSTM* and 1D-CNN*), the previous best-performing model (StemGNN, a Spatio-Temporal Graph Neural Network), and the homogeneous approach (GAT). This establishes HGAT as the new state-of-the-art model for both, the 1-Second and the 1-Minute Dataset. Comparing StemGNN to HGAT reveals that constructing a heterogeneous graph from data schematics also appears to be effective. However, the proposed approach has revealed certain limitations. First, the graph was constructed based on physical connectivity and plausibility and is static. For example, position information on switches and circuit breakers could be used to refine the graph. Achieving an optimal message passing graph may also require additional experimentation with subgraph representations of physical assets and additional expert input. To enhance the model's performance further, incorporating additional heterogeneous node types could be beneficial. For instance, metrological nodes could provide insights into both hydrological (e.g., rainfall impacting dam levels) and electrical systems (e.g., temperature affecting equipment efficiency). Similarly, railway traffic data, which influences electricity demand could also improve forecasting capabilities. For detailed results we would like to direct the reader to the publication [3].

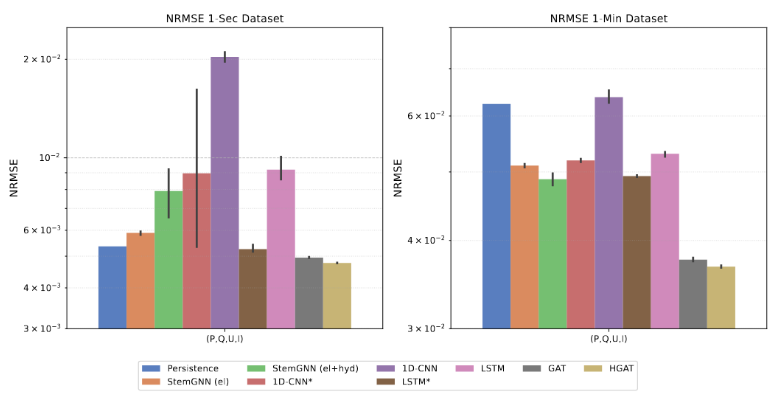

Figure 4: Normalized root mean square error (NRMSE) for short-term state forecasting averaged over all electrical current sensor nodes.

3 Load Forecasting in Traction Power Grids (Case 3)

Timetable-based Energy Forecasting: During the 2022 and 2023 report period, we gathered evidence suggesting that spatial-temporal transformer models – a deep learning architecture that uses attention mechanisms to weigh the importance of different input elements dynamically - are well-suited for forecasting tasks within traction power grid environments. This observation is highly significant for the task of day-ahead forecasting, which is an important source of information for energy trading, grid dispatching and load management applications. During the past report period, we finalized the work on the data-driven model that learns to forecast the grid load of the entire traction power grid of SBB from timetable-based information in a publication. In this review period, we have received feedback on the publication from the peer review process. The feedback suggested including a second case study to show the generality of the approach. Therefore, during the current reporting period, we have added an additional contextual load forecasting study on building energy, where context information is given in the form of detailed building occupancy.

Load forecasting for building energy systems is critical, as they account for approximately 40% of total energy consumption globally [4], with about one-third of global energy consumption attributed to building operation. The increasing electrification of heating and mobility services further elevates the importance of building power demand for their smooth integration into local distribution networks. Accurate load forecasts support energy cost management, improve operational efficiency by optimizing Heating, Ventilation, and Air Conditioning (HVAC) systems, and enable better integration with local renewable energy sources, particularly for larger buildings such as offices.

Buildings are also well-suited for incorporating future contextual information, as occupant behavior and usage patterns - key drivers of energy demand - can provide valuable insights to enhance forecasting accuracy and align energy consumption with expected usage.

Experimental Setup and Training Data: To demonstrate the generalizability of our forecasting framework beyond load forecasting for the railway traction power grid, the auxiliary case study on Building Energy forecasting focuses on an individual medium-sized office building in the United States, a subset of the dataset AlphaBuilding registered with the U.S. Department of Energy's Open Energy Data Initiative (OEDI). The dataset represents 70% of U.S. commercial buildings and is widely used within the building energy community. The dataset is generated using the building simulation software EnergyPlus which models the office building by incorporating building characteristics, HVAC system specifications, and weather data.

Baseline Methods: We evaluated all contextually enhanced transformer models. Namely, Spacetimeformer (CE-STF), Crossformer (CE-CF) and Timeseries Transformer (CE-TST). In addition, we analyzed various models to demonstrate the competitive performance of our contextually enhanced transformers developed in this study case. Specifically, we benchmark our model against the current state-of-the-art in long-range time-series forecasting. From recent advancements in multi-step linear models, we include DLinear and TiDE, a dense residual model known for its effectiveness in long-term forecasting. We also included PatchTST (PTST), a model that applies a patch-based approach to the standard transformer architecture for efficiency and performance. To provide a thorough comparison, we also evaluate the distinctive (inverted) embedding strategies employed by iTransformer traditional time-series forecasting methods, such as bidirectional LSTM, which effectively incorporates future covariates. Furthermore, we extend our analysis to a popular gradient boosting framework, CatBoost, recognized for its robustness and efficiency in diverse predictive modeling tasks. These results are summarized in the publication [2].

Analyzing Forecasting Outliers: In the building energy case study, we observe a comparable performance improvement among the contextually enhanced transformer models in terms of reducing outlier counts through the integration of future contextual information. CE-CF with contextual information consistently records the lowest outlier counts across all MAPE threshold values. Particularly, CE-CF displays an average of 2.97 % significant outliers exceeding a 30% MAPE which is lowered to 0.2 % outliers when FCI is included. CF also reaches a zero count of outliers already at a MAPE threshold of 1.25, whereas the exclusion of FCI raises the threshold to 1.92 MAPE. This pattern reinforces the advantage of incorporating future context in reducing forecasting outliers and establishes contextually enhanced Crossformer as the most robust model against outliers. Moreover, the inclusion of future context universally helps to reduce the outlier counts across models. The outlier analysis is summarized across both the Railway and the Building Energy case studies in Figure 6.

Summary of the Results: The performance analysis in terms of normalized mean absolute error (NMAE), across both the Railway and the Building Energy case studies, is summarized in Figure 5. Our study underscores the critical importance of integrating timetable-based information to improve the accuracy and robustness of load forecasting models. We observed that while modern linear multi-step models (such as Dlinear or TiDE) perform well on datasets with clear trends and periodicity, they struggle in scenarios requiring timetable-based information. In contrast, contextually enhanced transformers demonstrate superior performance in handling multiple data streams, effectively leveraging timetable-based information. The addition of the Building Energy case study demonstrated

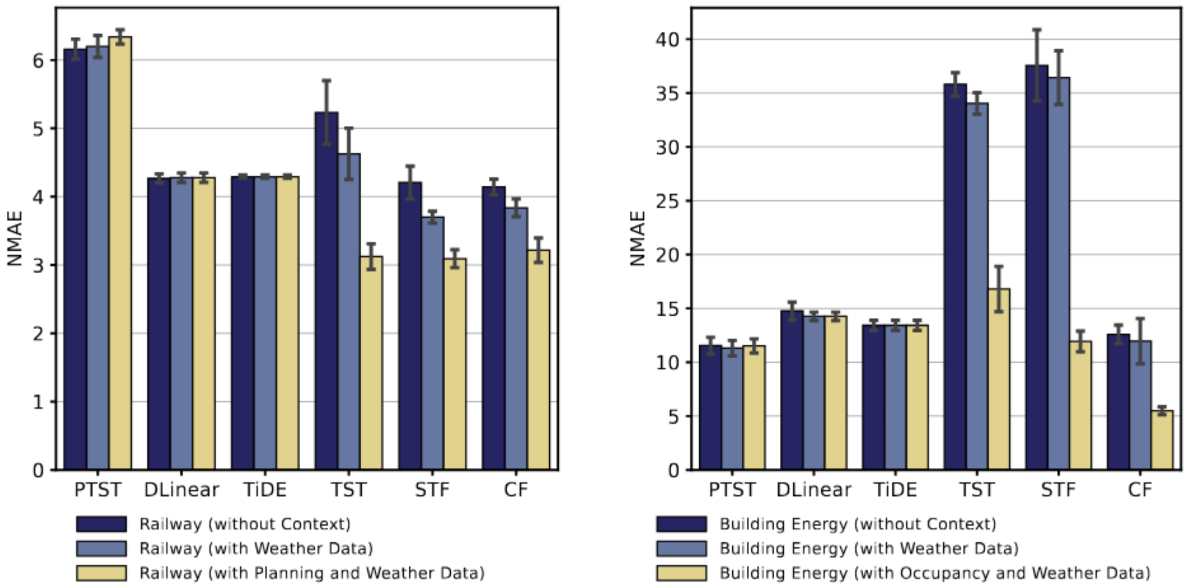

Figure 5: Normalized MAE in watts with and without the addition of timetable-based information. We list all transformer models (PTST, TST, STF, CF) and multi-step linear models (DLinear, TiDE) included in our evaluations for the for the railway traction load forecasting case study (left) and for the building energy case study (right).

that, contrary to the baseline approaches, incorporating rich contextual information is a necessary condition for transformers to perform efficiently on the data. Summarizing the work of this case study, contextually enriched encoder-decoder transformers excel at leveraging timetable-based contextual information, achieving superior performance confirmed by their successful application to both multi-year load forecasting case studies in the railway traction power grid and the new building energy case study introduced in this reporting period. For detailed results we would like to direct the reader to the publication [2].

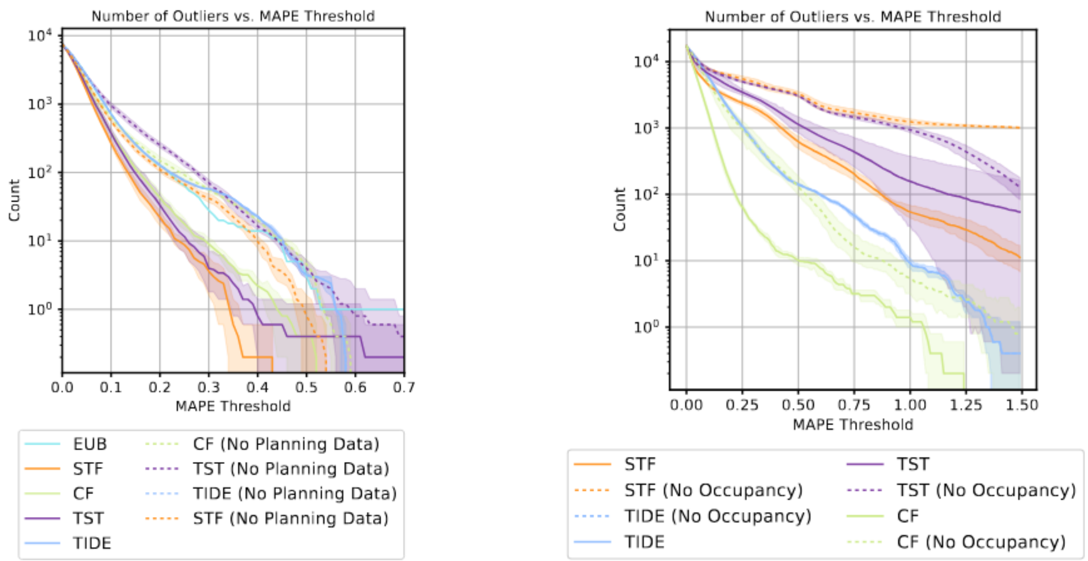

Figure 6: The count of outliers by forecasting model as a function of the MAPE threshold, comparing models with future contextual information and those without (dashed line). We display outlier counts by forecasting model plotted against MAPE threshold for the railway traction load forecasting case study (left) and the same result for the building energy case study (right).

4 Conclusions

Modern industrial systems generate extensive data that are routinely collected in energy management systems and other operational monitoring and control systems. Leveraging this data, project INITIATE has demonstrated that data driven methods are of substantial practical interest to power grid operators:

- Case study 1 addressed the challenge of state estimation for anomaly and sensor fault detection in a pumped storage hydropower plant with heterogeneous sensing. The proposed Heterogeneous Graph Attention Network (HGAT) fuses electrical and hydraulic sensor data over an explicitly constructed heterogeneous graph to significantly improve short term forecasting accuracy. In contrast to existing approaches, the method operates robustly across multiple temporal resolutions and supports real time deployment as demonstrated on the one second resolution case study.
- Case study 2 addressed the problem of accurately estimating hydropower turbine efficiency under variable operating conditions from data. The developed physics-informed convolutional neural network (PICNN) for efficiency estimation learns turbine performance characteristics directly from historical hydro and electrical measurements. The proposed novel framework provides more reliable efficiency maps for operation and planning.
- Case study 3 addressed the task of day ahead load forecasting in the SBB traction power grid given complex timetable dependent demand patterns. The proposed contextually enhanced transformer framework presents a novel method that conceptualizes forecasting as a combined forecasting-regression task, formulated as a sequence-to-sequence prediction problem. This method efficiently integrates timetable information and other future contextual variables with historical load data to produce more accurate and robust day ahead load forecasts for both railway traction power grids and building energy systems.

The case studies demonstrate that data-driven methods systematically benefit from the inclusion of additional physical, contextual and topological information. In particular, the enhanced HGAT architecture for sensor state estimation exploits heterogeneous electrical and hydraulic information and spatial sensor positioning. The prediction and reconstruction performance of the PICNN for hydropower turbine efficiency indicates that incorporating nozzle information significantly improves the estimation accuracy. In parallel, the contextual transformer framework for traction power grids achieved strong performance in processing future contextual information for load forecasting. This information includes temperature predictions, tonnage, kilometers traveled, gross tonne-kilometers, and train counts derived from the timetable for regional, long-distance or intercity, and cargo trains.

Across all investigations, the project has clearly demonstrated that the explicit integration of physical, structural and contextual knowledge is key to achieving robust and reliable data-driven model performance. The resulting methodological advances provide a strong and versatile foundation for the development of next-generation predictive tools for railway infrastructure and related application domains. In particular, the work has significantly advanced the accuracy and reliability of hydropower plant state forecasting, load forecasting, and turbine efficiency estimation under real-world conditions enabling models are increasingly sensitive to distinguishing healthy from anomalous system behavior.

Further validation under anomalous conditions requires records of representative fault data, which is currently limited. Importantly, the remaining step of defining thresholds on residuals for anomaly detection is straightforward and can be readily implemented in practice. To fully exploit the potential of these methods, it is recommended to systematically collect representative datasets that cover the relevant operating regimes and environmental conditions. Notably, no failure data is required for training the models; such data is only needed for validation purposes.

Given the large number of relevant operating conditions, the case studies showed that the performance of current data-driven models remains strongly dependent on the availability of data. Increasing the amount of representative data even by a small factor of two or three, which in practice is often not associated with significant additional costs, can substantially improve model performance. Moreover, the continued explicit integration of physical, structural, and contextual knowledge has been shown to be highly valuable to improve model performance, robustness, and generalization in data-limited conditions. This implies that a promising research direction for future work is to further reduce the dependence of data-driven approaches on operating conditions to increase data efficiency.

The work has resulted in three conference and journal publications covering two of the case studies, thereby making the research widely accessible to the public. Furthermore, we have disseminated the results at various international conferences and created a new dataset that will be released to the public.

5. References

[1] : Theiler, Raffael, and Olga Fink. "Graph Neural Networks for Electric and Hydraulic Data Fusion to Enhance Short Term Forecasting of Pumped Storage Hydroelectricity". PHM Society European Conference, vol. 8, no. 1, Prague, Czech Republic, 2024. https://doi.org/10.36001/phme.2024.v8i1.4129

[2] : Theiler, Raffael, and Olga Fink. "Integrating the Expected Future: Schedule Based Energy Forecasting." arXiv preprint arXiv:2409.05884 (2024).

[3]: Theiler, Raffael, and Olga Fink. "Heterogeneous Graph Neural Networks for Short-term State Forecasting in Power Systems across Domains and Time Scales: A Hydroelectric Power Plant Case Study" arXiv preprint arXiv:2507.06694, (2025)

[4]: OECD. Renewables 2022: Analysis and Forecast to 2027. Organisation for Economic Cooperation and Development, Paris, (2022)