



Artificial neural network-based line monitoring sensor with digital twinning

Deliverable No: D6.7

Work package: WP6

Official delivery date: 30.06.2026

Actual delivery date: 26.06.2026

Dissemination level: Public



Co-funded by
the European Union

Project: 101136119 | HORIZON-CL5-2023-D3-01 | www.twineu.net

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Document Authors	
Andreas Strasser (APG)	Marton Markovits (BME)
Balint Nemeth (F4STER)	Levente Racz (BME)
David Szabo (BME)	Tamas Angster (E.ON)
Gabor Szücs (MAVIR)	

Version	Date	Author(s)	Notes
0.1	2025.04.14.	BME, F4STER, E.ON, APG, MAVIR	1 st draft
0.2	2025.04.29.	All partners	2 nd draft
0.3	2026.06.01.	BME, F4STER	3 rd draft
0.4	2026.06.22.	All partners	4 th draft (after internal reviews)
0.5	2026.09.23.	All partners	5 th draft
0.6	2026.09.29.	All partners	6 th draft

Responsible Partner	BME
Checked by WP leader	Daniele Porcu (ENEL), 29.06.2026 (v1), 29.09.2026 (v2)
Verified by the appointed Reviewers	Efstratia Sarafoglou (HEDNO), 10.06.2026 Javier Rodrigo Gutiérrez de la Cámara (OMIE), 10.06.2026
Approved by Project Coordinator	Padraic McKeever (Fraunhofer), 26.06.2026 (v1), 30.09.2026 (v2)

Table of Contents

Executive Summary	9
1 Introduction	10
1.1 Task 6.8	10
1.2 Objectives of the Work Reported in this Deliverable	11
1.3 Outline of the Deliverable	11
2 Monitoring systems for overhead lines	13
2.1 Dynamic Line Rating of Overhead Lines	14
2.2 General challenges of a complex line monitoring system	15
3 Digital twin for power line monitoring systems	18
3.1 Relationship to Previous Projects and Novelty of the Demo	18
3.2 Motivation for the Digital Twin of Line Monitoring	20
3.3 Sensor placement algorithm	21
3.4 ANN-based conductor temperature model	21
3.5 DLR - Ampacity calculation model	23
3.6 Advantages and usability of power line monitoring DT	23
4 Methodology of Task 6.8	25
4.1 Timeframe and Structure of the Demo	26
4.2 The core of the use case scenarios	28
4.3 Preparation for the demonstration	29
4.3.1 Involved grid parts and hardware	30
4.3.2 Collection of the Historical Data	32
4.3.3 IT structure of the project (communication, DT layers, interoperability)	33
4.3.4 Definition of the initial KPIs	35
4.4 Installation of hardware	38
4.5 Demonstration Stage 1	39
4.5.1 Data collection	39
4.5.2 Processing the data	40
4.5.3 Digital Twin implementation	41
4.5.4 Evaluation system	44
5 Results and Discussion	49
5.1 Results for APG grid	51
5.1.1 Results for APG grid with the time-series split	52

5.1.2	Results for APG grid with the seasonal split	53
5.1.3	Results for APG grid with the weekday split	55
5.1.4	Discussion of results for APG grid	56
5.2	Results for MAVIR grid.....	57
5.2.1	Results for MAVIR grid with the time-series split	58
5.2.2	Results for MAVIR grid with the seasonal split	59
5.2.3	Results for MAVIR grid with the weekday split.....	60
5.2.4	Discussion of results for MAVIR grid	62
5.3	Results for E.ON grid.....	63
5.3.1	Results for E.ON grid with the time-series split	64
5.3.2	Results for E.ON grid with the seasonal split	66
5.3.3	Results for E.ON grid with the weekday split.....	67
5.3.4	Discussion of results for the E.ON grid.....	69
5.4	Discussion of the results.....	70
5.5	DLR calculation algorithm.....	74
5.5.1	DLR values calculated for the various grids	77
5.6	Connection to Task 7.3 in Work Package 7	79
5.6.1	Concept of the joint work	79
5.6.2	Data forwarded to Task 7.3.....	80
5.7	Key Performance Indicators	82
6	Lessons learned	84
7	Future works.....	88
7.1	Recommendations for regulation.....	88
7.2	Potential future application of the DT.....	88
7.3	Future Monitoring, Communication and Asset Management Strategy	90
7.4	Potential industrial utilization of the DT	90
8	Conclusions.....	92
	References	94

List of Figures

Figure 1 - Concept of complex line management technology [6].....	13
Figure 2 - Installation of dynamic line rating equipment in the Hungarian distribution and transmission system: a) line monitoring sensor, b) weather station c) weather station [7]	14
Figure 3 - Optimization of DLR sensors' placement on power lines based on clearance values (left) and risk of thermal overload (right) [9][10]	16
Figure 4 Concept of the power line monitoring system DT	18
Figure 5 Schematic structure of an applied neural network [15]	22
Figure 6 Schematic structure of the Power line monitoring DT (HUN Demo).....	25
Figure 7 Gantt chart of the Hungarian Demo	26
Figure 8 Structure of the preparation steps	30
Figure 9 The general IT and communication structure of the Demo	33
Figure 10 Hardware installation on the E.ON grid	38
Figure 11 Structure of the Demonstration Stage 1.....	39
Figure 12 Quantity-check for two different months.....	41
Figure 13 Different ANN configurations and optimization of the momentum and learning rate	44
Figure 14 Plots for model A1_6m with time-series split.....	52
Figure 15 Plots for model A2_12m with time-series split.....	53
Figure 16 Plots for model A1_6m with seasonal split.....	54
Figure 17 Plots for model A2_12m with seasonal split.....	54
Figure 18 Plots for model A1_6m with weekday split	55
Figure 19 Plots for model A2_12m with weekday split	56
Figure 20 Plots for model B1_6m with weekday split	58
Figure 21 Plots for model B2_12m with time-series split.....	59
Figure 22 Plots for model B2_6m with seasonal split.....	60
Figure 23 Plots for model B2_12m with seasonal split.....	60
Figure 24 Plots for model B1_6m with weekday split	61
Figure 25 Plots for model B2_12m with weekday split	62
Figure 26 Plots for model C1_6m with time-series split.....	65
Figure 27 Plots for model C2_12m with time-series split.....	65
Figure 28 Plots for model C1_6m with seasonal split.....	66
Figure 29 Plots for model C2_12m with seasonal split.....	67
Figure 30 Plots for model C1_6m with weekday split	68
Figure 31 Plots for model C2_12m with weekday split	68

Figure 32 Comparison of boxplots between the three grids	71
Figure 33 Scatter plots for the final models.....	72
Figure 34 Signed error distribution of the final models.....	73
Figure 35 Time-series plot for model C1 during one whole day.....	74
Figure 36 Capacity gain can be achieved by the application of DLR technology	75
Figure 37 DLR calculation results with different models	76
Figure 38 Capacity gain for one year on the E.ON line (summer and winter)	77
Figure 39 DLR values calculated from the real and virtual sensors	78
Figure 40 Forecasted DLR values for E.ON grid in summer	81
Figure 41 Forecasted DLR values for a week in winter for the E.ON grid	81

List of Tables

Table 1 Inputs and Outputs of the Preparation for the Demonstration subtask	26
Table 2 Inputs and Outputs of the Installation of Hardware Subtask	26
Table 3 Inputs and Outputs of the Demonstration Stage 1-2 Subtasks.....	27
Table 4 Inputs and Outputs of the Evaluation of the DT models Subtask	27
Table 5 Inputs and Outputs of the Dissemination activities.....	27
Table 6 HUN DT Use Case 1 Partners.....	28
Table 7 HUN DT Cases for the ANN-based DT	29
Table 8 Information exchange in the preparation phase	29
Table 9 Involved power lines in the Hungarian Demo.....	31
Table 10 Involved hardware in the Hungarian Demo	32
Table 11 HUN DT Initial KPIs	36
Table 12 Information exchange in the Demonstration Stage 1.....	39
Table 13 Real-time data collected in the Demonstration Stage 1	40
Table 14 Description of the model names and results achieved during the project (seasonal split)...	50
Table 15 Metrics for APG with time-series split	52
Table 16 Metrics for APG with seasonal split	53
Table 17 Metrics for APG with weekday split.....	55
Table 18 Metrics for MAVIR with time-series split	58
Table 19 Metrics for MAVIR with seasonal split	59
Table 20 Metrics for MAVIR with weekday split.....	61
Table 21 Metrics for E.ON with time-series split	64
Table 22 Metrics for E.ON with seasonal split	66
Table 23 Metrics for E.ON with weekday split.....	67
Table 24 Metrics for the final models.....	71
Table 25 Final KPIs as a result of the project	82
Table 26 Metrics for KPI No. 3 and 4	83

List of Abbreviations and Acronyms

Acronym	Meaning
AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
APG	Austrian Power Grid – Austrian TSO
BME	Budapest University of Technology and Economics
CGMES	Common Grid Model Exchange Standard
CIGRE	The International Council on Large Electric Systems
CIM	Common Information Model
DLR	Dynamic Line Rating
DSO	Distribution System Operator
DT	Digital Twin
E.ON	E.ON EED - E.ON Észak-dunántúli Áramhálózati Zrt. – Hungarian DSO
F4STER	Future 4 Sustainable Transport and Energy Research Institute Zrt. – Technology provider
FARCROSS	Facilitating Regional CROSS-border Electricity Transmission through Innovation – EU H2020
FBCC	Flow-based Capacity Calculation
IEEE	Institute of Electrical and Electronics Engineers
LiDAR	Light Detection and Ranging
MAE	Mean Absolute Error
MAVIR	Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zrt. – Hungarian TSO
MSE	Mean Squared Error
NIS2	Network and Information Security, 2 nd version
OHL	Overhead Line
RES	Renewable Energy Sources
RMSE	Root Mean Square Error
SCADA	Supervisory Control and Data Acquisition
SO	System Operator
TSO	Transmission System Operator
WP	Work Package
WS	Weather Station

Executive Summary

The electricity sector across Europe - and globally - is undergoing unprecedented transformation. Among the most pressing changes are the rapid increase in electricity demand, the high penetration of renewable energy sources (RES), and the aging of transmission and distribution infrastructure, which requires upgrading and making the existing grid infrastructure more flexible. High-voltage overhead lines are crucial components of the electricity grid, on which line monitoring is rapidly becoming a core tool for network operators. Dynamic Line Rating (DLR) technology represents a significant breakthrough in overhead line operations, enabling network operators to unlock 20–40% additional line capacity without costly infrastructure upgrades. DLR provides economic and operational advantages, enhancing cross-border power exchange and market integration. Several challenges in Central Europe have been identified in pilot line monitoring projects, including scalability, cost-effectiveness, and coordination between Transmission System Operators (TSOs) and Distribution System Operators (DSOs). The TwinEU project aims to address these barriers through regional collaboration and innovation, promoting the digitalization of the power line sector and the application of Digital Twin (DT) solutions.

The Hungarian Demonstration focuses on developing a DT model for power line monitoring and DLR systems. This model integrates real-time data from sensors on phase conductors, weather stations, and SCADA systems to estimate conductor temperature using an Artificial Neural Network (ANN). The DT model enables dynamic line rating and enhances asset management strategies, increasing grid resilience and reliability. The objectives include defining the DT concept and scope, preparing for the demonstration, developing ANN models, collecting required data sets, implementing the ANN models, integrating a line rating calculation module, evaluating the ANN models, and disseminating project results. The main emphasis during the project was on developing and testing digital twins to monitor the thermal state of the phase conductors on overhead lines. While physical sensors typically track conductor temperature with an accuracy of ± 0.5 to ± 2 °C (depending on the specific type of physical equipment), the project demonstrated that ± 2 °C of conductor temperature calculation accuracy can also be achieved with ANN-based digital twins. Although the accuracy of input parameters – mainly the weather parameter certainty – can significantly influence the accuracy of the developed digital twin, most cases yield an accuracy around the range of ± 2 °C. Based on the developed digital twin solution, the transmission capacity of the overhead lines can also be calculated with higher precision, including the thermal inertia of the conductors. Moreover, these results can be applied not just for congestion management, but also for co-optimization in market simulations.

The developed digital twin solution is pivotal for the application of DLR expert systems at both the DSO and TSO levels, as the use of virtual sensors (digital twins) enables the optimization of the number of physical devices. This optimization process can result in two different strategies. As more line sections can be observed with the same monetary expenditure, the overall system reliability can be increased. Moreover, with fewer physical devices required, the medium- and long-term (after the training period) CAPEX optimization of the DLR systems can be achieved from the SOs' perspective. Renting is also emerging as a business model among technology providers, opening new opportunities for them as well. To fully unlock the benefits identified in this project, coordinated regulatory support, standardized frameworks, and continued innovation in AI-driven Digital Twin solutions will be essential to enable scalable deployment and realize their significant operational and sustainability potential in future power systems.

1 Introduction

High-voltage overhead lines are key elements of the electricity grid [1]. Historically, their real-time monitoring was considered less critical due to the robust nature of the electricity system. Thermal ratings were (and are) based on conservative static assumptions, and the operational margin was not fully exploited [2]. However, as power systems have become more complex and stressed, dynamically understanding and responding to line loading has become essential. Today, line monitoring is not just common; it is rapidly becoming a core tool for network operators [2].

Implementing extensive line monitoring systems can hardly be generalized. Power line monitoring solutions can be complex, often requiring careful planning, integration of diverse sensor technologies, reliable communication infrastructure, and advanced data processing capabilities [3]. As a result, successful deployment depends on thorough preparation and the alignment of technical, organizational, and regulatory frameworks.

One key output of line monitoring systems is Dynamic Line Rating (DLR). DLR represents a significant breakthrough in transmission and distribution operations. Network operators can unlock 20–40% of extra line capacity without costly infrastructure upgrades [4]. By dynamically assessing the thermal limits of conductors, DLR allows for the more efficient utilization of existing assets, providing economic and operational advantages. Furthermore, DLR enhances cross-border power exchange and market integration by offering a more accurate picture of real-time grid capabilities.

In addition to the benefits of line monitoring and DLR systems, several challenges were identified in the pilot projects of the last few years. The TwinEU project offers a unique opportunity to address these barriers through regional collaboration and innovation [5]. The project promotes the digitalization of the power line sector and the application of Digital Twin solutions for optimal and secure operation.

The key component of the Hungarian Demonstration within the Work Package 6 is providing a Digital Twin solution for power line monitoring and DLR systems. The development and implementation of such DT tools can increase the utilization of power lines, while sharing the lessons learned can improve the cooperation and interaction of the partners across Central Europe.

1.1 Task 6.8

In the Hungarian Demonstration, a revolutionary approach is aimed to be adopted by developing a Digital Twin (DT) model that redefines the monitoring and management of power lines. This DT model integrates real-time data from strategically placed sensors on phase conductors, which measure temperature, with an advanced algorithm that defines the optimal locations for these sensors to mitigate sag-clearance issues and enhance system reliability.

The algorithm considers various factors, including power line span characteristics, conductor specifications, tension at installation, and comprehensive cross-sectional drawings, ensuring sensors are positioned where they are most needed. Surrounding weather stations contribute critical environmental data, capturing wind speed, solar radiation, ambient temperature, and more, while the SCADA system provides invaluable line load parameters. Over a year of data collection lays the groundwork for creating an Artificial Neural Network (ANN), which, with weather conditions and

SCADA load data inputs, is trained to accurately predict conductor temperature. This predictive capability is pivotal, enabling the DT to calculate phase conductor temperatures without direct sensor input, thereby allowing for strategically relocating sensors to other critical areas, optimizing the monitoring network, and reducing costs. Integrating this ANN into the DT framework empowers operators with dynamic line rating abilities and enhances asset management strategies, paving the way for a more resilient and reliable power grid. By leveraging predictive analytics, the DT model prevents potential sag-clearance problems and supports the power grid's adaptation to varying environmental and operational conditions. Ultimately, this DT concept embodies a forward-thinking strategy to increase grid resilience and ensure the reliable operation of power transmission networks, setting a new standard for the industry.

The following chapters present how Task 6.8 of the TwinEU project is structured and implemented.

1.2 Objectives of the Work Reported in this Deliverable

The objectives of the work reported in this deliverable are the following:

- Define the Digital Twin (DT) concept and scope for power line monitoring.
- Prepare for the demonstration, including:
 - Selection of the target power line(s)
 - Selection and validation of appropriate sensors
 - Ensuring sensors are placed at critical points
 - Verification and fine-tuning of the IT and communication systems
- Install additional hardware where necessary to support the monitoring setup.
- Develop Artificial Neural Network (ANN) models and optimize their parameters.
- Collect all required data sets, including technical specifications, historical records, and real-time measurements.
- Define a comprehensive testing framework for evaluating the ANN models.
- Implement the ANN model using the optimal structure determined through testing.
- Determine the minimum dataset size required for reliable and accurate ANN performance.
- Integrate a line rating calculation module based on sensor and model outputs.
- Disseminate the project results to relevant stakeholders and the wider community.

1.3 Outline of the Deliverable

This deliverable is structured into the following chapters:

- **Chapter 2 – Monitoring Systems for Overhead Lines:** This chapter introduces various overhead line monitoring systems, focusing primarily on Dynamic Line Rating (DLR) as a key approach for optimizing line capacity. It also discusses the technical and operational challenges of implementing complex, multi-component monitoring systems in the field.
- **Chapter 3 – Digital Twin for Power Line Monitoring Systems:** This section presents the concept and motivation behind applying a Digital Twin (DT) to power line monitoring, emphasizing the benefits of real-time system representation. It details key DT components,

including sensor placement strategies, an ANN-based temperature model, ampacity calculation methods, and practical advantages of using DTs in operational environments.

- **Chapter 4 – Methodology of Task 6.8:** Chapter 4 outlines the structured methodology used in Task 6.8, from timeline planning to scenario definition and demonstration execution. It covers the physical and digital infrastructure setup, including involved grid parts, communication systems, KPIs, and the stepwise approach to testing and data validation.
- **Chapter 5 – Results and Discussion:** This chapter presents the key findings from the demonstration and evaluates the effectiveness of the implemented monitoring and DT solutions. It includes reflections on data and model accuracy, system performance, and insights gained during the two demonstration stages. This chapter also contains the results of the KPI metrics.
- **Chapter 6 – Lessons learned:** This chapter summarizes the key lessons learned during the project, drawing on the experience gained throughout the Hungarian Demo implementation. It covers insights into data quality and quantity, the underlying IT infrastructure, and the development of the digital twin framework. In addition, it highlights findings on ANN optimization processes and the challenges and best practices associated with Dynamic Line Rating (DLR) calculations.
- **Chapter 7 – Future works:** This chapter outlines the recommended pathways for advancing digital twin-based line monitoring projects. It addresses key needs and gaps at both company and regulatory levels, while also identifying how the developed models can be extended to additional use cases beyond their original scope. Furthermore, it discusses the potential market applications and value creation opportunities associated with the presented results.
- **Chapter 8 – Conclusion:** The final chapter summarizes the main steps and outcomes of the project and highlights the potential of Digital Twin-enhanced monitoring systems in improving grid flexibility and reliability.

2 Monitoring systems for overhead lines

One of the most recent technologies for overhead line monitoring is the dynamic line rating (DLR) technology [3]. Conventionally, the transmission capacity of overhead lines is determined based on the worst-case environmental conditions to maximize the conductor temperature, such as unfavorable cooling conditions, to a safe threshold. This static line rating technology is a significant limiting factor for the power flows most of the time [6]. The recent monitoring systems – including hardware solutions such as weather stations and line monitoring sensors – allow the continuous monitoring of the indirect environmental conditions and the thermal state of the conductors. In this way, the actual current-carrying capacity of the phase conductors can be adjusted to the prevailing heating and cooling effects, which can mean a significant increase in transmission capacity on average. Based on the geographical conditions, design of the line, and the applied conductor, DLR technology can guarantee 20 to 40 % capacity gain on average [4].

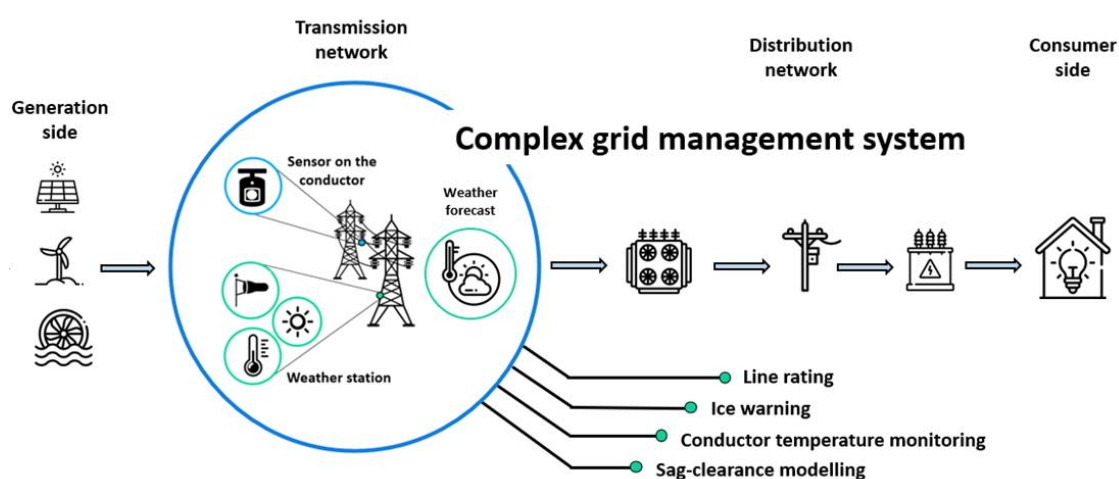


Figure 1 - Concept of complex line management technology [6]

The primary goal of DLR is to maximize the efficiency and reliability of the power grid by utilizing the full potential of transmission lines under favorable weather conditions. This approach not only enhances the capacity of existing infrastructure but also reduces the need for costly upgrades and expansions. By continuously monitoring and adjusting the line ratings, DLR helps to mitigate congestion, integrates renewable energy sources more effectively, and improves overall grid stability. For that purpose, a complex line monitoring system can be implemented. In this way, the DLR-based system can help system operators in the following aspects [6]:

- Calculation of **real-time line** rating: The real-time line rating calculation can achieve a 20-40% average capacity gain, which primarily helps in congestion management, integrating renewable energy sources, and handling N-1 network constraints.
- Dynamic **line rating forecast**: The weather forecast can be used to project the transmission capacity of power lines, which is helpful for optimizing the electricity market, maintenance planning, and network development plans.
- Real-time **sag-clearance tracking**: By knowing the conductors' thermal state, the phase conductors' sag can be tracked in real time. From the sag, the clearance values can also be determined, resulting in safer power line operation.

- **Ice forecasting (not in the scope of this project):** Extreme weather events can also be predicted using the same weather forecast used for the dynamic line rating forecast. Moreover, preventive actions based on Joule-heating can minimize the risk of power line icing.

In Germany, the implementation of DLR systems has shown significant potential in increasing the continuous current-carrying capacity of existing networks by approximately 20%, with higher utilization possible around 98% of the time. This method aligns with the NOVA principle (Network Optimization before Reinforcement before Expansion), emphasizing optimizing existing resources before considering new investments.

Altogether, the different subsystems of a DLR-based complex line management system can contribute to the safe and optimized operation of power lines.

2.1 Dynamic Line Rating of Overhead Lines



Figure 2 - Installation of dynamic line rating equipment in the Hungarian distribution and transmission system: a) line monitoring sensor, b) weather station c) weather station [7]

Dynamic Line Rating (DLR) technology is based on the real-time monitoring of overhead lines, allowing for the determination of the current carrying capacity of the lines in real-time. DLR systems calculate the maximum transmission capacity of the lines by considering the actual prevailing environmental conditions such as ambient temperature, wind speed and direction, and solar radiation. This approach offers significant advantages over static line rating as it dynamically adapts to changing conditions, increasing the efficiency and reliability of the network. The use of such systems

allows for an increase in network capacity. In this way, the investment, such as building new lines or reconductoring, can be postponed in time or even completely redeemed [7].

In Hungary, the implementation and application of DLR systems are still in their early stages but are continuously developing. Several pilot projects are underway in the country to test the technology and examine its practical applicability. For instance, E.ON Hungary has installed DLR systems at 3 locations to monitor the condition of the lines in real time and optimize network operations [7].

MAVIR, the operator of Hungary's transmission system, is also actively involved in developing and applying the DLR technology. As part of the MAVIR DLR project, sensors have been installed at 12 locations to monitor the thermal state of the phase conductors and environmental conditions in real time. Moreover, the extension of the system with a forecast for ampacity forecasting is also implemented. The optimized capacity flows on the critical lines equipped with the DLR system are significant during peak periods and energy transfers. The project aims to increase network capacity and reliability and optimize network operations. Based on the project's results, MAVIR plans further developments and expansions of the so-far-realized system to leverage the benefits of DLR technology more.

In Austria, the application of DLR technology has also started recently [8]. Austrian utilities, such as the transmission system operator, Austrian Power Grid (APG), have implemented an ambient-adjusted system using weather data in the grid since 2013. The current system increases the potential using weather data with max. 15-minute intervals, but to further exploit the potential of capacity increase and the ampacity forecasting on OHLs, APG started testing sensor-based DLR systems from different technology providers in its network in 2021. Therefore, around 18 sensors at 8 locations from three different manufacturers were installed on one OHL. Depending on the manufacturer, the sensors measure conductor temperature, line current, ambient temperature, and sag. The manufacturers differ in weather data; some have provided weather data from external providers, and some come with their weather station. One manufacturer comes with a forecasting system. The data, such as the calculated ampacity, is still analyzed, and based on the results, APG plans further expansion of sensor-based DLR systems.

2.2 General challenges of a complex line monitoring system

The challenges of complex line monitoring systems arise from numerous factors, which have emerged based on the experiences of DLR systems already implemented in the demonstration countries.

General scalability - The scalability of the power line monitoring project presents a significant challenge. While pilot initiatives typically involve only 1–2 sensors (fewer than five), full-scale deployment introduces complexities that must be carefully addressed. One major issue is determining the optimal locations and number of line monitoring sensors and weather stations along the overhead line routes. Due to the variability in topographical and environmental conditions, it is difficult to identify where sensors will provide the most valuable data. In regions with limited prior data or highly variable weather patterns, more sensors may be required to ensure accurate and representative measurements. To ensure optimal placement, thorough preliminary site surveys and planning based on these surveys may be necessary. For example, placing sensor pairs every 4–5 kilometers along a power line can be advisable, especially when spans with critical sag (minimum clearance) need

monitoring. However, proper installation across long transmission routes remains a complicated and labor-intensive process, requiring linemen to install and calibrate devices in often rugged terrain.

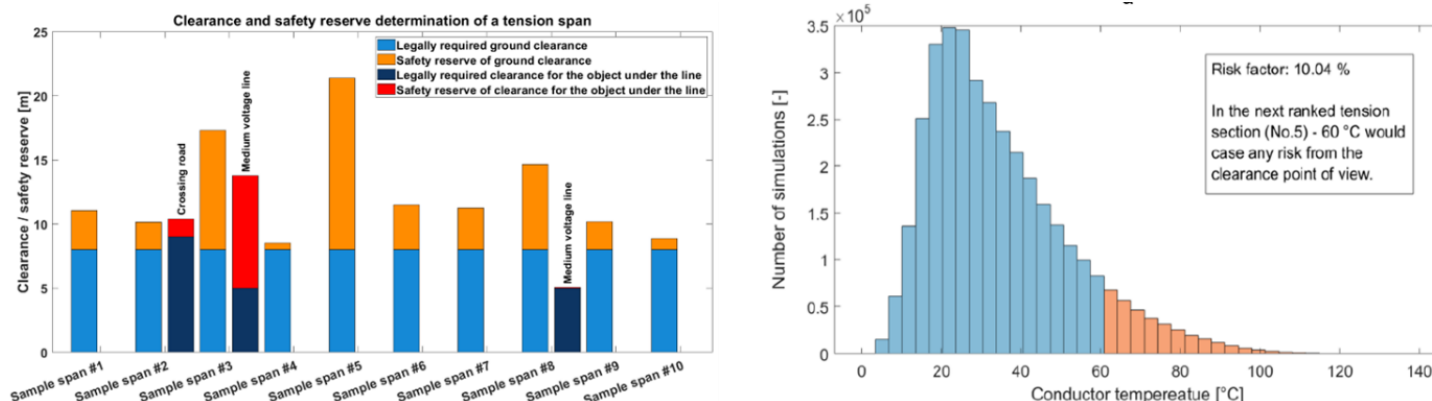


Figure 3 - Optimization of DLR sensors' placement on power lines based on clearance values (left) and risk of thermal overload (right) [9][10]

Project and hardware costs: Another major constraint is cost [10]. Procuring and installing sensors, weather stations, and other necessary infrastructure entails a substantial investment. If the price of the equipment becomes too high, it can hinder the full-scale implementation of Dynamic Line Rating (DLR) systems. This is particularly relevant in the pilot regions of Europe, where cost-effective approaches are actively sought to maximize the benefits of DLR technology. Weather-based virtual sensors have emerged as a possible solution, reducing the number of physical devices while still delivering high-resolution, location-specific information. However, line monitoring systems that rely primarily on weather patterns are less accurate than those systems that integrate physical sensors. A real challenge is investigating hybrid sensor systems where physical devices and reliable virtual sensors support each other in line monitoring systems.

TSO-DSO interactions: Operational readiness and coordination between TSOs and DSOs are essential to successfully deploying DLR systems [7]. Close collaboration supports contingency analysis, helping to assess load distribution and identify potential overloads under various normal and emergency operating conditions. This increases both the reliability and the security of the power network. DLR systems, where implemented, often operate in isolation without integration into broader system operations. In the future, this is expected to change with the extension of interaction across system components. Establishing clear communication protocols and integrated operational procedures is crucial for seamless data sharing and system interoperability. Robust partnerships and shared technical frameworks must support these efforts to overcome the challenges posed by large-scale implementation. At the same time, the increasing exchange of operational and sensor data introduces additional challenges related to data management, cybersecurity, and regulatory compliance. Real-time analytics and data fusion capabilities must be complemented by appropriate cybersecurity measures to protect communication interfaces and operational data and to ensure reliable system operation. As DLR deployment becomes more widespread and the interaction between transmission and distribution networks increases, these requirements highlight the need for coordinated risk-management practices and a common regulatory framework across organizational boundaries. In this context, NIS2 establishes a common baseline for both TSOs and DSOs, while Delegated Regulation (EU) 2024/1366 (NCCS) provides additional sector-specific requirements. Both

transmission and distribution operators may fall within the high- and critical-impact categories, ensuring that cybersecurity requirements apply across the operational interface rather than being limited to a single entity. Furthermore, the NCCS risk-assessment framework, extending from the Union level through the Member State level to individual entities, provides a common methodological basis for assessing risks associated with interconnected and boundary processes.

3 Digital twin for power line monitoring systems

The primary goal of the TwinEU project's Hungarian Demonstration (part of WP6) is to offer industrial partners already involved in line monitoring projects a solution that not only provides a technically and economically appropriate response to the existing challenges (scalability, reliability, economic investments) but also further expands the possibilities that are the cornerstones of a flexible network operation.

This chapter presents the methodology, a key component of the Hungarian Pilot. We detail the main characteristics of the Digital twin to be built, the technical background of the models, and their applicability. It also details how the ampere capacity (ampacity) value provided as the output of the models can be utilized in the Hungarian pilot, which is part of WP7. The main goal of WP7 is to use these ampacity inputs to enhance the day-ahead and intraday capacity calculations, which precede the market co-optimization scenarios.

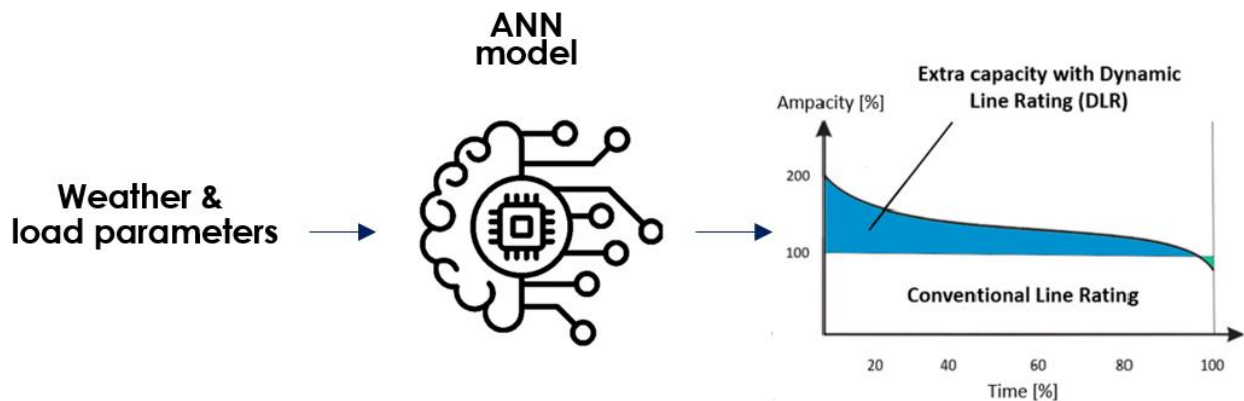


Figure 4 Concept of the power line monitoring system DT

In addition to the models, the project structure, the definition of the partners' roles, and the schedule are presented. Each component examines the key steps crucial for a successful technical implementation. Finally, the project's progress and status are reported in the context of the preliminary schedule.

3.1 Relationship to Previous Projects and Novelty of the Demo

The present study represents the latest stage of a research pathway that has progressively evolved through the FLEXITRANSTORE, FARCROSS, and TwinEU projects. Rather than introducing an entirely new concept, it builds upon methodologies, datasets, and lessons learned from these previous investigations, while addressing several limitations identified in earlier work.

The first investigations into ANN-based virtual sensors were performed within the FLEXITRANSTORE project. The original objective was to assess whether artificial neural networks could support or potentially replace physical monitoring devices in overhead-line applications. This work was motivated by the observation that different thermal rating methodologies, including IEEE and CIGRE approaches, may produce different conductor temperature and ampacity estimations under identical operating conditions. Due to the primary focus of FLEXITRANSTORE on anti-icing and de-icing

applications, the available datasets were relatively limited in both temporal and geographical coverage, and model development relied on measurements from only a small number of monitored locations.

The concept was further developed in the FARCROSS project, where the analysis was expanded to include eight transmission lines across four countries [11] [12]. The main objective was to investigate whether a generalized ANN architecture could reproduce measurements obtained from different sensor technologies and operating environments. The results demonstrated that virtual sensing is feasible not only for directly measured conductor temperatures but also for temperature values derived from sag-monitoring systems. The project also highlighted the importance of dataset length and diversity, leading to the conclusion that approximately one year of training data may be desirable to achieve robust performance. However, the available datasets generally covered only a single annual cycle, requiring the same dataset to be divided between training and validation purposes.

Complementary developments were performed within the TwinEU project, where Digital Twin concepts for power system applications were further refined, and their role in supporting operational decision-making was investigated. These activities helped establish the broader Digital Twin framework within which virtual sensor technologies can be deployed for monitoring and Dynamic Line Rating applications.

Building on these earlier achievements, the present work advances the state of the art in several important aspects. First, substantially longer monitoring datasets are available than in the previous studies. This enables a systematic investigation of how training data duration affects model performance and provides a significantly more robust basis for validation across diverse weather and loading conditions. Second, unlike the earlier studies that relied on predefined ANN architectures, the current work applies a comprehensive hyperparameter optimization process. The investigation includes the optimization of the number of hidden layers, neuron counts, learning rates, and other training parameters, enabling a more rigorous assessment of the achievable prediction accuracy. Third, the study evaluates not only the influence of training data duration but also the impact of weather data quality on the performance of virtual sensors. This aspect is of practical relevance because weather information represents a key input to Digital Twin-based line monitoring systems. Fourth, this project reflects those real scenarios about system operator connection and scalability that reveal the key aspects of line monitoring projects. Harmonizing communication and data sharing between TSO and DSO is a key factor in the future monitoring system and requires a clear framework and a uniform system to enable a higher level of cooperation.

Consequently, the principal contribution of this project is not the introduction of a new framework or the initial proposal of ANN-based virtual sensors, both of which have been presented in earlier publications. Instead, the novelty lies in the systematic assessment of training data requirements and weather data quality, combined with ANN hyperparameter optimization, to improve the reliability and practical deployment readiness of virtual sensors. The results provide stronger evidence that ANN-based virtual sensors can serve as robust Digital Twin components for system operators for Dynamic Line Rating and transmission line monitoring applications.

3.2 Motivation for the Digital Twin of Line Monitoring

System operators in the Central and Central Eastern European region have recently launched several pilot projects to increase flexibility, in which line monitoring has played a significant role [7][8]. In these projects, there is a great emphasis on properly implementing real-time monitoring, as it provides real-time data on the thermal state of power line conductors, which is closely related to sag-clearance conditions. On the other hand, it is of paramount importance that line monitoring systems, supplemented with forecast data, can also provide day-ahead and intraday transfer capacity data to the electricity market.

The first major challenge of such line monitoring systems is that, according to current practice, there is a lack of clear guidance on the number and location of sensors [9][10]. Sensor manufacturers are interested in selling as many devices as possible but often fail to provide suppliers of original equipment with adequate information. Although the SOs have extensive operational experience, they are not experienced in real-time line monitoring, as it was not previously considered an important aspect to know the thermal conditions along the entire transmission line. This largely determines that the detection of classic hot spots is not always trivial, and this affects the location and number of sensors. The other main problem is that if every tension or dead-end section is equipped with a physical sensor, the cost increases significantly, thereby worsening the project's profitability. Although efforts are currently underway to predict ampacity on power lines using virtual sensors without hardware devices, these solutions do not account for the fundamental limitations, conditions of the transmission line, and the degree of aging of the power line components. The third main problem is that the results of the currently available thermal condition calculation models often do not align with the measurement results of actual physical devices. This can be traced back to several roots, which were examined by a complex group of experts within the framework of the previous H2020 FARCROSS project [11]. In addition, a significant challenge lies in calculating and predicting the ampacity value for a transmission line in a manner that accounts for the system's current thermal conditions and the physical, heat-storing nature of the conductors. Finally, a significant challenge is line monitoring, including integration of DLR systems at the national or regional level and integrating the DLR systems of transmission lines operated by different operators and at various voltage levels into a unified, coordinated DLR system.

The Hungarian Demo aims to provide an integrated, digital twin-based solution to the challenges of line monitoring systems. Line Monitoring DT consists of three main parts, which focus on the following areas:

- Sensor placement algorithm
- ANN-based conductor temperature model
- DLR – Ampacity calculation model

These models integrate real-time and predicted data from hardware units, independent organizations, and existing power line infrastructure. They also contain developed software solutions, some of which are supported by artificial intelligence. Combining these parts provides a technical solution that can result in a more economical and faster return on investment for line monitoring systems.

3.3 Sensor placement algorithm

The number of necessary hardware devices must be adequately defined to create a line monitoring DT that is representative from a technical point of view. These hardware devices include two types of sensors. On the one hand, weather stations record environmental changes with a resolution of 10-15 minutes. On the other hand, line monitoring sensors, which are often installed on the phase conductor (in most cases, but there are exceptions as well), monitor the thermal behaviour of the phase conductors or closely related parameters (sag, tension, clearance, etc.) with the exact temporal resolution as the previous one [6].

The number of the hardware unit is also closely related to the power line sections in which the physical devices will be placed. The BME research group initiated an investigation into this issue within the FARCROSS H2020 project [12]. This part of the line monitoring DT must consider which physical limiting factor dominates along the power line under thermal conditions. On the one hand, there are power lines and areas where sag-clearance conditions impact determining thermal limits. However, when designing other transmission lines, such a safety factor is often reserved for the sag-clearance conditions, as the annealing process, typically occurring in the range of 90-93 °C, becomes critical [13].

The sensor placement algorithm is a software solution developed for DT based on analysing the entire route of the transmission line. In all sections of each power line, whether sag-clearance conditions or operation at excessively high temperatures cause problems sooner is investigated. The safety margins taken from the clearance point of view at the maximum conductor temperature are examined in each dead-end section. Those line sections that are critical from this perspective must be equipped with a physical device. The advanced algorithm defines the optimal locations for these sensors to mitigate sag-clearance issues and enhance system reliability. The algorithm considers various factors, including power line span, conductor specifications, tension at installation, and comprehensive cross-sectional drawings, to ensure that sensors are positioned where they are most needed.

The sensor placement algorithm that identifies critical sections is fundamentally essential in DT. It enables the interpretation of the conductor temperature and ampacity calculation model's results locally and along the entire length of the transmission line, which is a primary aspect regarding the further usability of the results.

3.4 ANN-based conductor temperature model

The second central part of the line monitoring DT is an ANN-based conductor temperature model. Since the sensor placement algorithm identifies the critical power line sections, the next task is to equip them with hardware units. Based on previous project experiences, it can be said that installing only a few sensors per power line is technically inadequate, except in cases where the power line length is short, typically up to a maximum of 15 km. In most cases, a distributed power line installation protocol is required to avoid overloads. This normally involves considerable hardware costs. The ANN-based conductor temperature model provides a solution to this [14].

The essence of the concept is that, following the installation of physical devices, which can last 6 months, 1 year, or longer, a virtual sensor is developed in parallel with the operation of the physical sensors. An integral part of the virtual sensor is an artificial neural network (ANN), which can replace

specific physical devices with an appropriate training dataset [15]. The ANN dataset required for training can be divided into two large groups. One is the weather data provided by the weather stations installed on the tower. In the case of the weather parameters, the reliability of the measured data can be handled with error flags to have information about the input data of the ANN. Moreover, at algorithm side outliers can be detected and credibility check can also be run. In addition, it is necessary to record the real-time load, which can be retrieved from the SCADA system. In the case of the SCADA data, a validity flag can also be transferred via the appropriate protocols to use the data only if it is reliable. The training output is the outer conductor temperature, which is measured by the physical sensor on the phase conductor. In the next step, the ampacity of the line can be determined from the conductor temperature. In the case, if the ANN network cannot provide the conductor temperature (e.g.: due to unreliable or missing data), the ampacity can be locked for the static line rating (SLR) value as a fallback procedure, in this way achieving maximum reliability in daily operation.

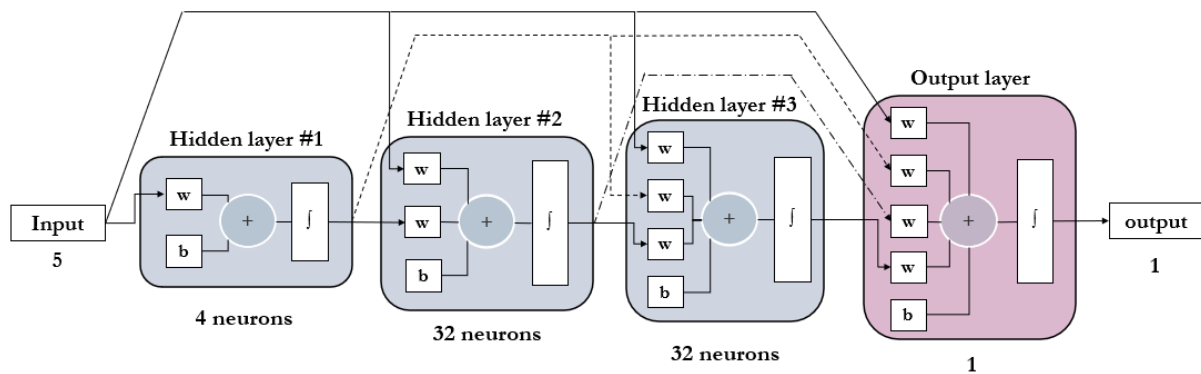


Figure 5 Schematic structure of an applied neural network [15]

Simulations based on the H2020 FARCROSS data provided the concept's basis [12]. However, in previous projects, the devices were relocated to another power line after the one-year operational period. Considering that unexpected circumstances occurred in many cases during that period, it was not possible to perform either the one-year training or the complete evaluation, as the physical devices used for validation were not in their original location.

Within the framework of the TwinEU project, unique power line-specific ANN solutions are developed, which feature hyperparameter optimization tailored to the specific transmission line section, making them optimal for the given line. The two main strengths of the DT concept are as follows: on the one hand, a conductor temperature simulation model is developed based on the ANN model, which more accurately approximates the values measured by the sensors compared to previous models. On the other hand, the developed solution is suitable for replacing specific physical devices after appropriate training, thereby accelerating the adoption of line monitoring projects and reducing their associated costs.

The hardware replacement primarily concerns the line monitoring sensor, which is more expensive and complicated to install than weather stations. It is essential to note that the solution does not intend to replace all physical devices. On the one hand, this approach provides continuous real-time data from every power line. On the other hand, the operation of the ANN model can be further refined with the addition of more data collection.

3.5 DLR - Ampacity calculation model

The ANN-based conductor temperature model provides helpful information to system operators during real-time operation. The relationship between the maximum and the operating temperature and the current degree of sagging and clearance at critical points of the transmission line can be monitored by displaying the conductor temperature in real time. However, this function alone is insufficient to increase network flexibility; this can only be achieved by dynamically determining the power line rating. The third central part of DT is an algorithm that, depending on the input data set, can evaluate the transfer capacity value of the transmission line in real time and predictively.

The existing literature mainly focuses on the IEEE, CIGRE, and IEC models when determining dynamic line rating values [16][17]. However, these models use neglects (precipitation, corona effect around the phase conductor), so their results are of limited application. In addition, calculating the terms affecting convective cooling in the heat equation can show significant differences in the models in some cases. Although temperature tracking can be implemented in these physical models in smaller time sections, determining the line rating takes little account of the heat storage nature of the conductors. This also means that the line rating determination can be calculated in parallel with the conductor temperature for given environmental parameters. So, these models result in instantaneous amplitude changes regarding the line rating, which are difficult to track from an operational perspective.

The peculiarity of the model implemented as part of the DT is that its input parameter is the conductor temperature, which is the output of the ANN-based virtual sensor. This allows us to consider the effect of weather variables on the conductor temperature only during each iteration – via the conductor temperature as an input parameter. At the same time, this temperature effect can be neglected in the time interval between calculation steps. Since then, the Joule heat has determined the change in the thermal state of the conductors. The model's main advantage is that it can provide the additional transmission capacity available in addition to the current line load during the period under consideration - thus providing the dynamic line rating. Of course, during the calculations, several additional parameters must be examined to avoid thermal problems, such as substation limiting elements, the issue of radial thermal distribution, etc.

3.6 Advantages and usability of power line monitoring DT

The implemented power line monitoring DT has numerous advantages and practical applications. The following advantages can be highlighted:

- Critical transmission line sections can be identified
- The number of sensors required can be determined
- The location of the sensors can be determined
- The ANN-based solution allows conductor temperature calculation even when the physical sensor is not working (load is less than it required for the energy harvesting system to operate)
- Scalability - Some parts of the physical equipment can be replaced over time, which can save costs (fewer devices are sufficient in the long run, and maintenance costs are reduced)
- Time can be saved (already purchased sensors can be moved to other locations, no time-consuming new public procurement procedures are required)

- The conductor's thermal behaviour can be monitored in real-time (thermal overloads and sag-clearance conditions can be avoided)
- The transmission capacity of transmission lines can be optimized using the DLR algorithm
- Redispatch costs can be reduced using the DLR algorithm

The implemented power line monitoring DT also has several possibilities, two of which should be highlighted.

One of its main advantages is that the results of models supported by real data can be later integrated into the asset management system of the given SO. On the one hand, monitoring of thermal conditions can reveal trends that can be observed in the change of conductor temperature. On the other hand, operating conditions that accelerate the aging of the mechanical parameters of the conductor can be identified, thereby affecting its service life. Both pieces of information can influence the maintenance schedule of a network section, sometimes indicate necessary investments, or prepare and support the work of the decision-maker.

Another main possibility of power line monitoring DT lies in the fact that the DLR data provided as input can be used. The DLR data can serve as a good input for the Hungarian demonstration included in the WP7 of the TwinEU project, in which the modelling and optimization of market processes take place on an integrated platform. Real-time DLR data can simulate intraday market processes, while forecasted DLR data support day-ahead simulation cases. The power line monitoring DT is suitable for generating real and day-ahead DLR values. This requires that a 24-hour forecast is available for the parameters serving as model inputs, at least with hourly resolution. Thus, transfer capacity constraints can be simulated multiple times on a given day for the following day. In this way, the DT platform developed in Hungarian Demo WP6 and WP7 can be rationally connected.

4 Methodology of Task 6.8

The Hungarian Demo aims to create a power line monitoring DT that uses ANN-based virtual sensors to track thermal processes and provide weather-dependent dynamic line rating data that can contribute to a higher level of utilization of the power grid and provide data to the electricity market platform (Figure 6).

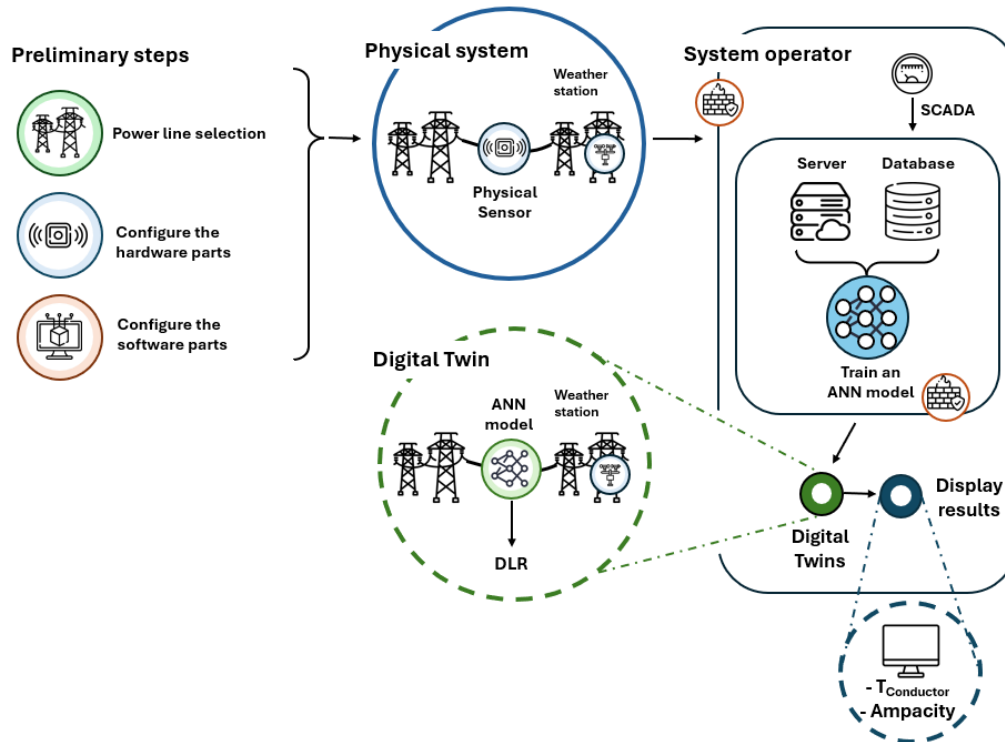


Figure 6 Schematic structure of the Power line monitoring DT (HUN Demo)

The goal is to demonstrate this implementation on three Central European system operator networks. The system operators involved in the project are:

- MAVIR - TSO (Hungary)
- APG - TSO (Austria)
- E.ON - DSO (Hungary)

Each partner participating in the Demo has, from previous pilot DLR projects, the hardware park, and IT system necessary to build a basic DLR system. The TwinEU project builds on this existing infrastructure so that no additional equipment, such as sensors and weather stations, is purchased during the project. Communication systems and IT protocols are a good starting point for an improved version of the software system.

The existing pilot systems provide the basis for the Hungarian Demo methodology. As a result, of the three main parts of the power line monitoring DT, the sensor placement algorithm can be considered given and only verifies that the existing devices are used to monitor critical points.

During the Demo, the main DT development focused on the other two subsystems: optimizing the ANN-based virtual sensors and the DLR calculation algorithm.

4.1 Timeframe and Structure of the Demo

The official project duration of the Hungarian Demo is between M4 and M33. For better planning, the project period is divided into five subtasks, one submission period and one dissemination period. After each subpart, an internal milestone is set, which provides a framework and checkpoint during the project. The project Gantt chart is presented in Figure 7.

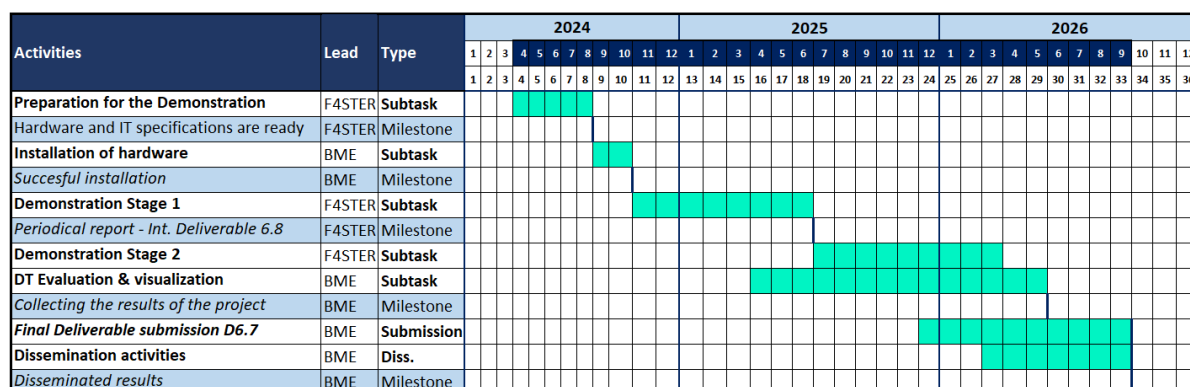


Figure 7 Gantt chart of the Hungarian Demo

Subtask 1: Preparation for the Demonstration

This subtask contains all the preparation tasks needed for the successful hardware installation and demonstration phase (choosing power lines, defining the type and number of the sensors, the critical span analysis, and fine-tuning the IT and communication background).

Table 1 Inputs and Outputs of the Preparation for the Demonstration subtask

Inputs needed	Outcome earned
• Technical details of the power lines. • 2D cross-sectional drawings. • Conductor details. • Sensor details • Communication and IT specifications. • SO cyber security protocols etc).	• The power lines involved in the project. • The sensor types and numbers. • The device installation locations. • The IT and communication system. • The scenarios for the DTs.

Subtask 2: Installation of hardware

The installation of the extra devices is potentially involved in the DT evaluation phase.

Table 2 Inputs and Outputs of the Installation of Hardware Subtask

Inputs needed	Outcome earned
• Installation location. • Field details for the linemen.	• The extra sensors and weather stations are available for data collection.

Subtask 3-4: Demonstration Stage 1&2 (data collection and implementation of the DT)

This subtask is the core of the 6.8 subproject. The physical devices collect the required thermal data from the conductor (conductor temperature and, in some cases, sag) while developing the ANN-based DT models in parallel. In parallel, bi-monthly performance reports are provided. Ampacity calculation from the thermal state of the conductor is also available.

Table 3 Inputs and Outputs of the Demonstration Stage 1-2 Subtasks

Inputs needed	Outcome earned
- Measurements of the physical devices (sensors, weather stations). - Technical details of the conductor. - Historically collected data from the physical devices. - Proper data device operation and data collection.	- Data based on the physical devices' measurements. - Developed ANN-based DT models for the evaluation. - Real-time ampacity based on the thermal behaviour of the phase conductor.

Subtask 5: Evaluation of the DT models

This subtask compares the physical sensors' measurements to those of the ANN-based thermal DT model. The evaluation will focus on average accuracy and outliers. The ampacity calculation for the SOs will also be evaluated and visualized.

Table 4 Inputs and Outputs of the Evaluation of the DT models Subtask

Inputs needed	Outcome earned
- Measurements of the physical devices (sensors, weather stations). - Results from the ANN-based DT. - Results from the ampacity calculation. - Proper data device operation and data collection.	- Final results about the performance of the ANN-based DT model and the average ampacity increase gained with the application of the model. - Ampacity based on the thermal behaviour of the phase conductor.

Dissemination activities

This activity focuses on publishing the achieved results at scientific forums (conferences, journal papers, etc.)

Table 5 Inputs and Outputs of the Dissemination activities

Inputs needed	Outcome earned
Final results from the demonstration and evaluation period.	Published results at scientific forums.

4.2 The core of the use case scenarios

In connection with the construction of the power line monitoring DT system, the primary focus during the project period is on the ANN-based virtual sensor DT part and the ampacity calculation model. For the Hungarian Demo, 2 national use cases related to the work of WP6 were defined:

- HUN DT Use Case 1 - Digital twin for power line monitoring
- HUN DT Use Case 2 - Enhanced flow-based capacity calculation for market co-optimization using DLR data

In HUN DT Use Case 1, the physical sensor-based line monitoring systems (necessary for DLR calculations) will be partially substituted by ANN-based digital twins (DTs). These kinds of DTs can provide accurate temperature results (inputs for the ampacity calculations) and decrease the number of sensors in the grid.

In the use case, ANN-based DTs acting as virtual sensors in the high-voltage network are developed. Table 6 lists the involved partners and their roles.

Table 6 HUN DT Use Case 1 Partners

Company	Country	Type	Role
APG	Austria	Transmission system operator (TSO)	• Grid Access Provider • Data Provider • System operator
BME	Hungary	Academic partner	• Independent evaluator • Technical supporter
E.ON	Hungary	Distribution system operator (DSO)	• Grid Access Provider • Data Provider • System operator
F4STER	Hungary	Technology provider	• Project manager • DT developer
MAVIR	Hungary	Transmission system operator (TSO)	• Grid Access Provider • Data Provider • System operator

The DTs are trained with different training datasets, transmission lines, and sensor data, and their technical accuracy is tested on a defined time series. The various cases are listed in Table 7.

Table 7 HUN DT Cases for the ANN-based DT

Case ID	Grid operator	Description	Training dataset	Output
1	APG	Short-period ANN training	ca. 6 months of data	- Optimized ANN structure - Conductor temperature - Dynamic Line Rating (DLR)
2		Long-period ANN training	ca. 12 months of data	
3	E.ON	Short-period ANN training	ca. 6 months of data	
4		Long-period ANN training	ca. 12 months of data	
5	MAVIR	Short-period ANN training	ca. 6 months of data	
6		Long-period ANN training	ca. 12 months of data	

During the DT's operation, it is investigated whether some physical sensors could be replaced with ANN-based virtual sensors in the Hungarian pilot.

HUN DT Use Cases 1 and 2 connect via the output of the line monitoring system DT. HUN DT Use Case 2 applies the calculated line rating values, demonstrating how these DLR results improve flow-based capacity calculations for cross-border electricity markets. By integrating line monitoring data, the DLR-enhanced approach enables more accurate and adaptive capacity calculations, which optimizes transmission network utilization, reduces congestion, and supports greater integration of renewable energy sources. A deeper analysis of this use case and methodology is presented during the work of WP7.

4.3 Preparation for the demonstration

The preparation for the demonstration is an essential part of the HUN Demo, as it provides a baseline for the whole project. The preparation phase includes several subparts as follows:

- Definition of the Demo plans
- Choosing power lines
- Defining the number of sensors and WSs
- Defining the initial KPIs for the demonstration
- Collecting all the historical info from the SOs
- Configuration of the software part of the Monitoring System

Figure 8 shows a schematic of the preparation phase, and Table 8 shows the information exchange during the preparation phase.

Table 8 Information exchange in the preparation phase

ID	Name of information	Description of information exchanged
I-01	Name of involved power lines	Grid parameters (Limiting elements in substation, line length, historical load level).
I-02	Spans for sensors	Cross-section drawings of the power line and technical datasheet of the line elements.

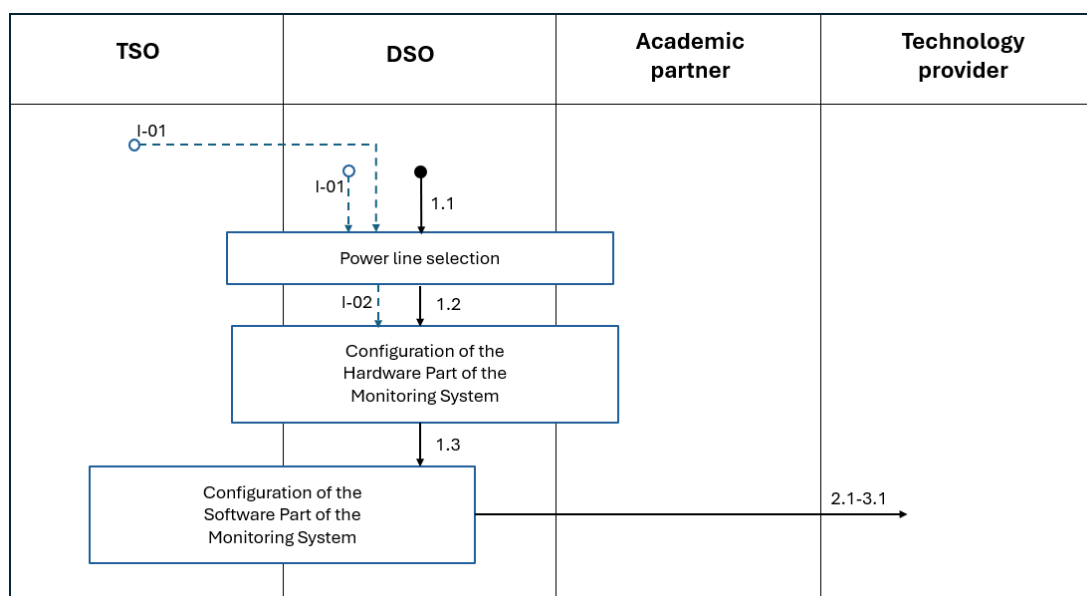


Figure 8 Structure of the preparation steps

The preparation phase lasted between M4 and M8. The provisions of the TwinEU project Grant Agreement defined the Demo plan. The agreements for each sub-part were finalized during monthly meetings involving all partners. As a result, the concept presented in Chapter 3 was defined.

4.3.1 Involved grid parts and hardware

After the concept was finalized, each SO involved in the project stated that they would like to participate in the DT implementation as a demonstrator, with measurements and simulations on their grid parts. In connection with this, it was defined which transmission lines and how many sensors would be involved in the project.

4.3.1.1 APG

After the FARCROSS project, APG started a DLR pilot on one specific line to investigate the potential of sensor-based DLR systems in the case of increasing load. Therefore, a ca. 100 km long, 220-kV double bundle line in the western part of Austria was chosen. The line routes through valleys and mountainous terrain are a good example of how the conditions can differ along the spans.

Three different manufacturers were chosen for the project, and around 18 sensors were installed at 8 different locations. Five locations were selected together with Graz University of Technology to cover exposed spans, such as the highest point of the line. Three further locations were chosen, with one manufacturer of the DLR system and the manufacturer's algorithm.

The three systems differ in the provided data; they use the provided weather data or come with their own weather station. The calculated ampacity based on the sensor data is compared to APG's current ambient-adjusted system to analyse the potential of the sensor-based systems. At the beginning of the project, there were some challenges regarding continuous measurements, and some sensors had to be exchanged. Therefore, the investigation is still running.

4.3.1.2 E.ON

Recognizing the advantages of Dynamic Line Rating (DLR) systems, ELMŰ, one of Hungary's E.ON DSOs, initiated a pilot project in 2021. During that project, a 132 kV power line was involved, for which a high risk of overload was predicted in the coming years. The reason behind the heavy utilization was the increased solar power generation in the region.

To support this initiative, ELMŰ evaluated sensors from multiple suppliers, considering accuracy, reliability, communication capabilities, and suitability for the chosen line. After identifying critical spans and assessing the proportional length of the monitored section, they installed two conductor monitoring sensors and two weather stations.

This line, equipped with sensors, has been chosen to be integrated into the TwinEU project. In the first year of the TwinEU project, the sensors will be relocated to another 132 kV line to collect more data from DSO power lines as a backup system.

4.3.1.3 MAVIR

MAVIR, the Hungarian TSO, has been actively piloting DLR systems for several years. As part of the H2020 FARCROSS project, line monitoring sensors and weather stations were installed on an internal and cross-border 400 kV power line. On both lines, 1-1 sensor with different measuring principles (contact measurement vs. LiDAR measurement) were installed and tested. This system was later extended with additional transmission lines and sensors, covering much of the Hungarian 220 and 400 kV network.

A 220 kV transmission line, which was included in the previously mentioned successful DLR pilot project, was also included in the TwinEU project. Since 2021, 2 pieces of contact temperature measurement sensors and 2 pieces of weather stations have been located along that transmission line.

In the case of MAVIR, several transmission lines can be included in the demonstration phase of the TwinEU project in case of a data shortage. The selected backup transmission line is a single-circuit 400 kV power line equipped with monitoring units similar to the 220 kV transmission line.

The power lines involved in the Hungarian Demo are listed in Table 9.

Table 9 Involved power lines in the Hungarian Demo

SO	Power line ID	Voltage Level	Number of circuits	Status
APG	A1	220 kV	Double	Involved
	A2	220 kV	Double	Backup
E.ON	E1	132 kV	Single	Involved
	E2	132 kV	Single	Backup
MAVIR	M1	220 kV	Double	Involved
	M2	400 kV	Single	Backup

The power line sensors and weather data involved in the Hungarian Demo are listed in Table 10.

Table 10 Involved hardware in the Hungarian Demo

SO	Power line	Line monitoring sensor			Weather data	
		Type	Measuring principle	Quantity	Type	Quantity
APG	A1	“A”	Contact Temperature	2 pcs	Weather agency	9 pcs
	A2	“A”	Contact Temperature	2 pcs	Weather agency	3 pcs
E.ON	E2	“ANN”	Contact Temperature	2 pcs	Weather station	2 pcs
	E2	“B”	Contact Temperature	1 pc	Weather station	1 pc
	E1	“B”	Contact Temperature	2 pcs	Weather station	2 pc
MAVIR	M1	“B”	Contact Temperature	2 pcs	Weather station	2 pcs
	M2	“B”	Contact Temperature	2 pcs	Weather station	2 pcs

The training datasets and the developed ANN models will be made accessible to the scientific community to the greatest extent permitted by data ownership and confidentiality constraints. However, the underlying datasets originate from TSOs and DSOs and are therefore subject to strict confidentiality requirements. Consequently, restrictions apply to the open-access publication of both the datasets, even in anonymized form, and the associated algorithms.

To ensure compliance with these confidentiality obligations, direct public dissemination of the raw data and model implementations will not be possible. Nevertheless, the scientific outcomes of the research, including model performance evaluations, methodological findings, and derived conclusions, will be disseminated through scientific publications, conference presentations, and other publicly accessible channels, thereby contributing to the advancement of knowledge within the research community.

4.3.2 Collection of the Historical Data

After selecting the transmission lines and sensors for the project, the next step was to collect historical data. This included all the design, technical, and environmental documents that may affect the operation and define the operational constraints.

- General power line data
- Power line route
- Power line design parameters
- 2D Cross-sectional drawings
- Conductor datasheet
- Substation limiting equipment
- Average load of the line
- Extreme operational cases from the past
- Extreme weather events in the past

4.3.3 IT structure of the project (communication, DT layers, interoperability)

A general IT and communication structure of the Hungarian Demo is shown in Figure 9.

The Hungarian Demo DT data architecture was defined in the preparation phase. This architecture consists of 5 layers as follows:

- Physical layer
- Data ingestion layer
- Data processing and storage layer
- Artificial intelligence and analytics layer
- Digital twin interface layer.

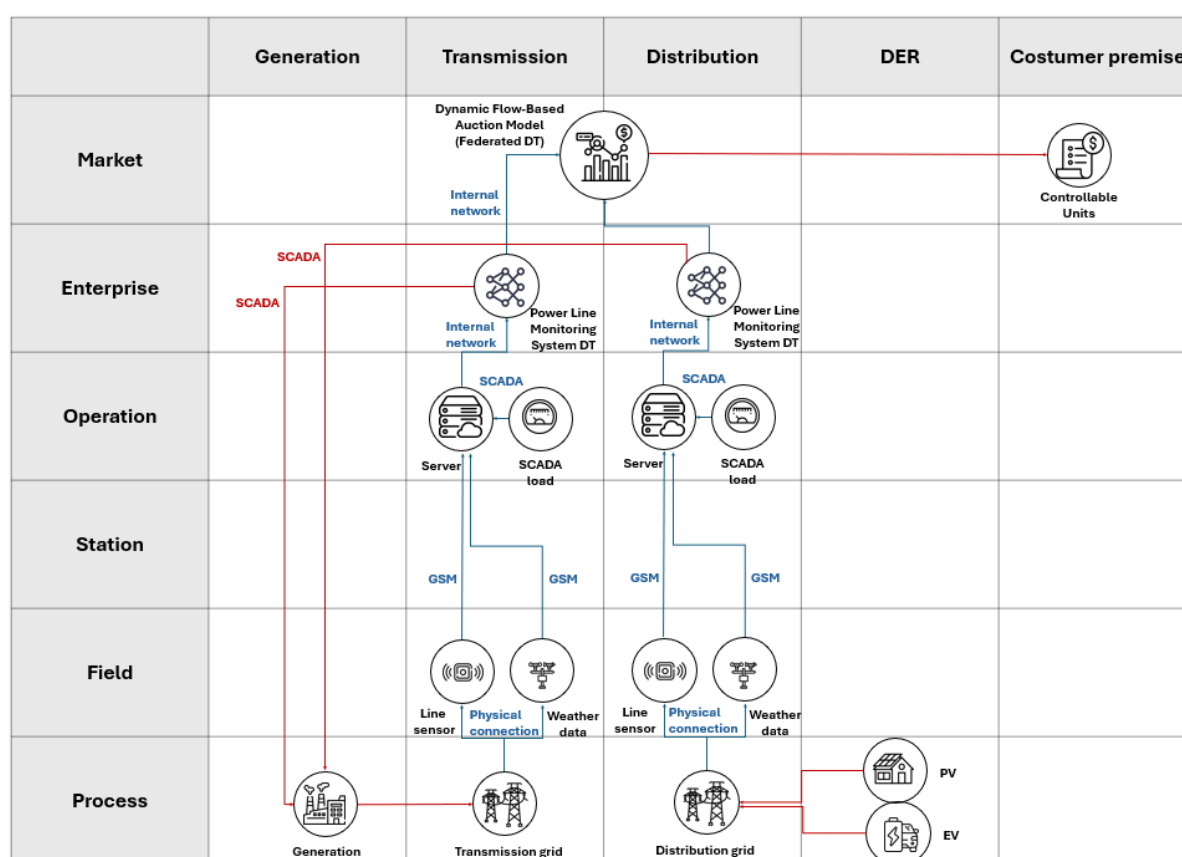


Figure 9 The general IT and communication structure of the Demo

Regarding the software and communication systems, the project aimed to follow the IT structure of the existing DLR pilot systems. An exception to this is E.ON, where, according to the company specification, an update must be made by the end of the project. The goal is to implement a more general environment integrating data from multiple platforms.

4.3.3.1 Physical layer

The physical layer forms the foundation of the Digital Twin (DT) system by incorporating three real-time data sources. The line monitoring sensor captures the temperature at specific conductor location, offering direct insight into the thermal state of the power line. Complementing this, the weather station gathers essential environmental parameters such as wind speed, solar radiation, and ambient temperature, which are crucial for accurately modelling thermal behaviour and optimizing line performance. Lastly, the SCADA system supplies detailed operational data, including line load information, which reflects the dynamic electrical conditions and their effect on conductor temperatures.

4.3.3.2 Data ingestion layer

The data ingestion layer interfaces the physical components and the DT system, efficiently collecting and transmitting diverse data streams. The core of this layer is the data aggregation module, which combines inputs from sensors, weather stations, and SCADA systems into a cohesive format suitable for analysis or storage. Robust communication protocols such as GSM enable smooth and secure data flow, providing reliable connectivity and real-time transmission capabilities necessary for maintaining the DT's operational accuracy.

4.3.3.3 Data processing and storage layer

The data processing and storage layer ensures that all incoming data is preserved and prepared for analysis. Its storage sublayer archives both historical and real-time data related to environmental and load conditions, creating a valuable resource for long-term modelling and predictive analytics. Simultaneously, the data preprocessing module enhances data quality by organizing, cleaning, and filtering the incoming information, correcting errors, and removing anomalies to produce clean datasets that support accurate Digital Twin simulations and insights.

4.3.3.4 Artificial intelligence and analytics layer

The artificial intelligence and analytics layer equips the Digital Twin (DT) with advanced computational intelligence to interpret and act on incoming data. At its core, the Artificial Neural Network (ANN) leverages historical and real-time data to predict conductor temperatures under various environmental and loading conditions, enabling proactive and efficient asset management. The predictive analytics engine builds on these capabilities by suggesting optimal sensor placements based on data trends, enhancing the flexibility and cost-efficiency of the monitoring strategy. Complementing these tools, the Dynamic Line Rating (DLR) module evaluates real-time line capacity, empowering grid operators to make informed decisions about line utilization and safety under changing operational conditions.

4.3.3.5 Digital Twin Interface Layer (implemented later)

The Digital Twin interface layer is the system's visual and interactive component, presenting vital insights to users through a graphical user interface (GUI). The visualization dashboard offers a clear and intuitive display of real-time conditions, including temperature trends, sag-clearance margins, and risk indicators. Operators can engage directly with these insights through the user control panel. In addition, the alerts and notifications system provides automatic warnings when critical thresholds are

breached, such as excessive conductor temperature or sag, facilitating swift, data-driven responses to potential issues.

4.3.3.6 Integration and Interoperability Layer (implemented later)

Designed for future deployment, the integration and interoperability layer envisions a scalable and secure ecosystem for Digital Twin applications. This layer will include APIs and communication protocols that allow smooth data exchange between the DT and broader grid management platforms, ensuring the DT can be embedded within the existing operational landscape without friction. Additionally, a robust data security module will be implemented to safeguard sensitive information, using encryption and controlled access mechanisms to uphold data integrity and prevent unauthorized access—an essential step as DT systems evolve into critical components of power grid infrastructure.

Data models: The project plans to adopt standard models for data exchange where possible. The Common Information Model (CIM) is considered at a basic level. At the same time, the more focused Common Grid Model Exchange Standard (CGMES) is preferred for more advanced and detailed grid data sharing. However, in cases where neither CIM nor CGMES can be implemented, individual solutions tailored to the project's needs may be used to ensure data interoperability [18].

Data exchange: Regarding real-time data exchanges, the system manages multiple data streams to support thermal calculations. These include sensor measurements from the power lines, weather station (WS) measurements, external weather agency data, and inputs from the SCADA system. A distinction is made between hard and soft real-time data; soft real-time data is sufficient for this project. This means that a delay of up to 20 minutes in receiving the data is acceptable and still valuable for monitoring and analysis.

The project operates with two leading types of data exchange to support planning and real-time decision-making. The first is data for operational planning, which is typically exchanged daily and used for day-ahead planning processes. The second type is real-time data exchange, which operates on a soft real-time basis and is primarily used for intraday calculations and system adjustments. This real-time exchange relies on edge computing to process data locally and efficiently, reducing latency and ensuring timely insights. Additionally, the data is often aggregated to optimize transmission and make the system more scalable and manageable.

4.3.4 Definition of the initial KPIs

During the preparation phase, primary emphasis was placed on evaluating the DT results. Thus, preset key performance indicators were defined, which are presented in Table 11.

Table 11 HUN DT Initial KPIs

ID	Name	Description	Target value
1	No. of physical sensors	How many physical sensors were used for the data collection of the ANN-based DT training?	Over 3
2	No. of DT models	How many ANN-based DT models have been implemented?	Over 3
3	No. of accurate DT models	How many ANN-based DT models have been made where the average achieved accuracy is within plus or minus 2 degrees Celsius?	Over 1
4	Virtual sensor accuracy (max & mean)	What is the maximum and average DT accuracy achievable (one-month exam period)?	Avg: ± 1 °C Max: ± 2 °C
5	Virtual sensor accuracy (outliers)	What is the outlier percentage of the DT model (one-month exam period)?	Below 5%
6	No. of power lines covered with virtual sensor	How many transmission lines are covered with virtual sensors (DT)?	Over 1
7	No. of potential substitutions	How many physical sensors can be potentially substituted by ANN-based DT models?	Over 1

The detailed explanation of the KPIs listed in Table 11 is provided below.

KPI No. 1: Number of physical sensors

The first KPI concerns the number of physical sensors used for data collection. The target value was set to more than three sensors. Therefore, the project required the utilization of at least three physical sensors to provide the data necessary for the training and validation of the ANN-based digital twin models.

KPI No. 2: Number of DT models

The second KPI evaluates the number of implemented digital twin models. The target value was defined as more than three models. Consequently, at least three different ANN-based digital twin models were required to be developed and implemented during the project.

KPI No. 3: Number of accurate DT models

The third KPI focuses on model accuracy. Accuracy is evaluated using the whisker values of the prediction error boxplots according to MATLAB's boxplot methodology. In MATLAB, the whiskers extend to the most extreme observations located within 1.5 times the interquartile range (IQR) from the first and third quartiles, while observations outside this range are classified as outliers. The target accuracy was defined as ± 2 °C. This KPI measures the number of models whose whisker range remains within the specified ± 2 °C limit. The initial target value was set to be more than one model.

KPI No. 4: Virtual sensor accuracy (max & mean)

The fourth KPI complements KPI No. 3 by evaluating the achievable accuracy of individual models. The accuracy is defined as the deviation of the errors from the ideal, zero error line. Therefore, the maximum accuracy defined in the KPI refers to the whisker range of the error boxplot, reflecting the spread of the majority of prediction errors. The average accuracy is represented by the width

of the box itself, corresponding to the interquartile range and therefore approximately 50% of all predictions. The target values were defined as below ± 2 °C for the maximum accuracy and below ± 1 °C for the average accuracy.

KPI No. 5: Virtual sensor accuracy (outliers)

The fifth KPI evaluates the proportion of outliers produced by the models. It should be noted that the definition of an outlier used for this KPI differs from the statistical outlier definition applied in MATLAB boxplots and discussed in KPI Nos. 3 and 4. While MATLAB classifies outliers based on observations lying outside 1.5 times the interquartile range (IQR), KPI No. 5 uses a project-specific definition directly linked to the required prediction accuracy.

For the purpose of this KPI, an outlier is defined as any prediction error exceeding the predefined ± 2 °C acceptance range. The outlier percentage is therefore calculated as the ratio of predictions with an absolute error greater than 2 °C to the total number of predictions. This definition was chosen because it directly reflects the practical accuracy requirements of virtual conductor temperature sensors and provides a straightforward measure of model reliability. The target value for this KPI was set below 5%.

KPI No. 6: Number of power lines covered with a virtual sensor

The sixth KPI measures the number of power lines that can be covered by virtual sensor technology. A power line is considered covered if at least one virtual sensor model achieves sufficient accuracy for practical application on that line. The initial target value was set to more than one power line.

KPI No. 7: Number of potential substitutions

The seventh KPI evaluates the number of physical conductor temperature sensors that can be substituted by virtual sensors. A physical sensor is considered replaceable if its corresponding virtual sensor meets the required accuracy and reliability. The initial target value was set to more than one substituted sensor.

4.4 Installation of hardware

In the Hungarian Demo, the selected transmission lines, with one exception (E2), had already been integrated into pilot DLR projects. This also means that power line sensors and weather stations (weather agency data in the case of APG) were available on all transmission lines except the E2 power line.

The hardware installation on the E2 transmission line aimed to have a backup power line in case of any data loss, from which load, environmental, and thermal data are available. The hardware installation (Figure 10) took place in the M9-M10 period.



Figure 10 Hardware installation on the E.ON grid

Before the installation, E.ON experts determined the critical transmission line sections on which the sensors and weather stations should be installed. These devices previously collected data on the E1 power line. The installation took place in a de-energized state after prior agreement. This is because no live work is carried out using any working method on the Hungarian 132 kV network. The line crew placed the power line sensor on the lower phase conductor, while the weather station was on the tower steel structure. Technicians tested the installation and ensured that communication and data flow from the sensors were functioning correctly after the power line energization. As the data stream began, incoming data from the sensors, the SCADA system, and weather forecasts were checked in the DLR central monitoring application. The MATLAB-based DLR pilot system displays data and forecasts regarding transmission capacity both graphically and numerically, based on the incoming data.

4.5 Demonstration Stage 1

After the preparation phase and the installation of additional physical devices, the next step is Demonstration Stage 1. The most important part of this phase is ensuring the continuity of data collection. The 8 months of data collected between M11 and M18 are essential in constructing the ANN-based power line monitoring DT.

Figure 11 shows the general structure of the Demonstration Stage 1, and Table 12 shows the information exchange in the project phase.

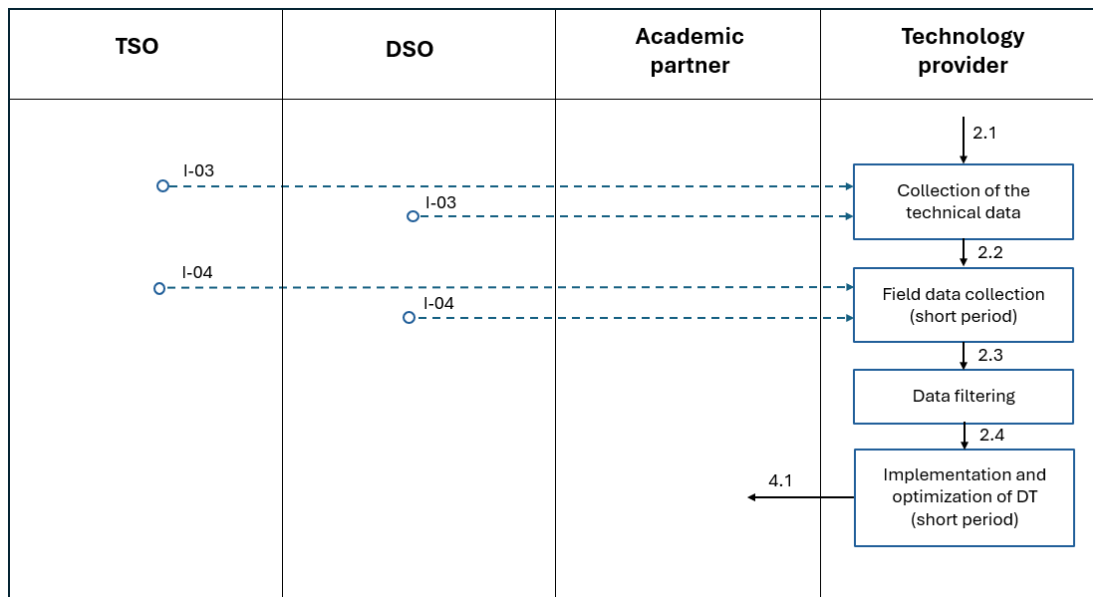


Figure 11 Structure of the Demonstration Stage 1

Table 12 Information exchange in the Demonstration Stage 1

ID	Name of information	Description of information exchanged
I-03	Phase conductor data	Technical datasheet of the phase conductor.
I-04	Field data	Collected weather parameters (temperature, wind, solar radiation), line load (SCADA), and conductor temperature.

4.5.1 Data collection

This phase includes technical data, data from real physical devices, and weather models. These provide the opportunity to train and test the ANN part of the DT. In this project phase, only the short training period is investigated among the use case implementations. Thus, the ANN-based models are trained with approximately half a year of data (detailed in Table 13).

Table 13 shows that the basic concept in data collection was to test the different data sources in terms of usability. This is reflected in both the weather and the data collected by the sensors. Regarding weather data, local weather stations installed on power line towers on E.ON and MAVIR transmission lines collect data. In contrast, in the case of APG, data provided by an independent weather agency rather than local units served as training data.

Table 13 Real-time data collected in the Demonstration Stage 1

Data source	Data [Unit]	Data Frequency		
		APG	E.ON	MAVIR
Line sensor	Conductor temperature [°C]	15 min	10 min	10 min
SCADA	Current load [A]	15 min	10 min	10 min
Weather Agency (APG) & Local weather stations (E.ON & MAVIR)	Wind speed [m/s]	15 min	10 min	10 min
	Wind direction [°]	15 min	10 min	10 min
	Ambient temperature [°C]	15 min	10 min	10 min
	Solar radiation [W/m ²]	15 min	10 min	10 min
	Relative humidity [%]	15 min	10 min	10 min

In the case of sensors, both hardware measured contact temperature, but the exact type of measuring unit and its recording method were unique to each manufacturer. Examining these differences in advance can help better evaluate the results.

Weather forecasts are also available near the transmission lines involved in the demonstration to ensure that DLR calculations are available day-ahead. These forecasts are for a period of 24 hours ahead, with an hourly resolution. The estimates for a given location are updated at specified intervals, so that the accuracy of the estimates can be monitored. Forecasts are available for the following weather data:

- Ambient temperature
- Wind speed and wind direction
- Solar radiation
- Relative humidity

4.5.2 Processing the data

The work performed on the collected raw data during the Demonstration Stage 1 contained two main parts. First, preliminary data processing was done, where the data needed for the DT implementation was unified and quantity checked.

Uniformization: While the type of collected data is the same for the different network operators, the applied sensors and databases vary, such as the format of the provided information. Therefore, the first step after the data collection period was a processing phase. Then, a suitable data structure was defined, and algorithms were developed to unify the data format. The data structure was created to be optimal for neural network training and testing.

Quantity-check: A preliminary data analysis followed the pre-processing stage. The efficiency of the neural network training mainly depends on the quantity and quality of the training data. Lower quality or quantity leads to an insufficiently trained neural network, which provides less accurate outputs. For the reasons stated above, the quality and quantity of the input data were analyzed for

every grid and every measurement point. The amount of the data was first determined based on the timely manner of the records. In general, it can be calculated for a specific period how many data points should be available based on the predetermined time interval of the records. Comparing this to the available data, the quantity can be calculated. Figure 12 shows an example of a data check – the different colors represent different data availability in each hour of a month.

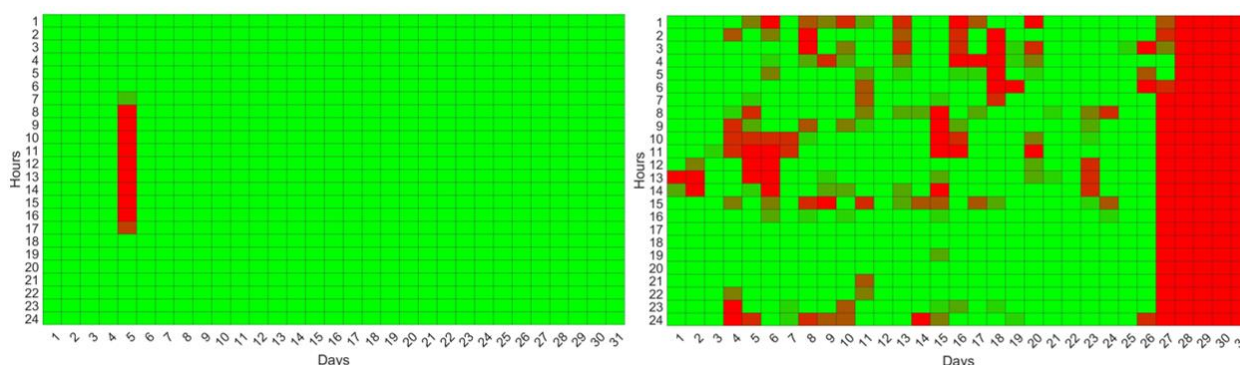


Figure 12 Quantity-check for two different months

Quality-check: Determining the quality of the records is a more challenging task. The quality can be viewed from different aspects. A data point can be incorrect when there is a value, but it's outside the plausible range of that variable. An example of this case is a wind speed record with a negative speed value. Another aspect of data quality can be evaluated using neural network training. The ANNs are trained based on similarities and patterns. As a result, data without correlation can be difficult to train on successfully. However, this task can only be checked after a few iterations of the training period. The overall quality and quantity were characterized by a ratio called availability. It can be calculated for the entire dataset or calculated and visualized hourly.

4.5.3 Digital Twin implementation

The unified, pre-processed data are appropriate for DT development. As highlighted before, WP6's work within the Hungarian Demo focuses on implementing the ANN-based virtual sensors of the Digital Twin. Thus, ANN models were developed in MATLAB for all the physical sensors involved in the project. The thermal state values produced by the ANN models serve as inputs to the ampacity calculation in the DLR part of the DT.

The ANN training utilized all available data, comprising approximately 17 months of records (accounting for missing or invalid entries). According to the operational principles of neural networks, the dataset must be partitioned into training and evaluation subsets. A neural network learns the relationships and patterns between input and output variables by adjusting its internal parameters during training. Therefore, to ensure a fair evaluation of performance, it is essential that the model does not "see" the data used for evaluation. This principle also applies to any modification of the network, as such changes would violate the independence between the training and test sets. This consideration is particularly important in the present case, where the optimization process involves modifying the network based on performance. Consequently, performance must be assessed using data that are not involved in the training or optimization process.

4.5.3.1 Data splitting strategies

To ensure the independence of the training and test datasets, a train-test split is typically performed. In the Hungarian Demo, the total data span was 17 months. The training–test split was performed with two training set sizes, corresponding to approximately 6 and 12 months of data. Consequently, for each grid and each measurement point, two models were trained: one using a smaller training dataset and another using a larger one. Within these predefined training set sizes, a specific splitting strategy was applied to obtain better training results, while meeting the required 6- or 12-month training data subsets. The splitting process refers to partitioning the available data into training, validation and testing sets according to a defined methodology. Numerous splitting strategies exist, and additional approaches can be developed for specific applications. In the Hungarian Demo, the network's task is to predict the conductor's temperature based on ambient parameters and current. In this case, the seasonal effects have a strong influence on the records; therefore, the splitting method needs to be representative of all 4 seasons. In our work, we have used 3 methods, described below.

- **Weekly split:** This split method was implemented in the early stages of the project. In this method, the split is done weekly. The training set consists of data from 5 days of the week, while the validation and testing set consists of 1-1 day of data. This method ensures that the data's seasonality is preserved while keeping the 3 datasets independent of each other, which is essential.
- **Seasonal split:** In this method, the data is split by percentage values inside a season. This means that, for example, 80% of the data from winter, spring, summer, and autumn form the training set, while the remaining 10-10% from each season form the validation and test sets. This results in a semi-random split, similar to the widely used random, percentage-based split method, while maintaining seasonal effects.
- **Time-series split:** The time-series-based split works by dividing the data for training, validation, and test sets based on the date. For example, the first 12 months of data are used for training, the next 2 for validation, and the last 3 for testing. While this method captures the digital twin's behaviour under operating conditions, it does not account for seasonal effects.

The exact ratios in the seasonal and time-series splits were selected to achieve the target training durations of 6 and 12 months, while accounting for missing data. Consequently, the gross training data corresponded to approximately 7 and 13 months, respectively, resulting in usable net data of 6 and 12 months, respectively.

The training process, described in Chapter 4.5.3.3, was performed for all grids using all three split methods described here. This results in multiple models for a single measurement point. In Chapter 5, where the results are shown, the best-performing split with good generalizability will be chosen for the final model.

4.5.3.2 Normalization methods

After splitting the data, an important step is normalizing the input data for training. Neural networks process features differently depending on their range and magnitude. In our case, the input features exhibit significantly different ranges, since the networks use weather and SCADA data. For example, the ambient temperature typically ranges from -10 to 40 °C, while the current flowing through the power lines has large, strictly positive values, often in the range of several hundred to a

few thousand. Without normalization, this disparity would lead to asymmetric training and potentially inaccurate results. Therefore, normalization is essential. In this demonstration, several normalization methods were applied, as described below. For reference, a configuration without normalization was also included.

Z-score normalization: Z-score normalization, also known as standardization, is one of the most widely used normalization techniques in neural network regression and machine learning tasks. It transforms the data so that the resulting distribution has a mean of 0 and a standard deviation of 1, allowing features with different scales to contribute equally during model training.

$$X_{norm,i} = \frac{X_i - \mu}{\sigma} \quad (1)$$

Min-Max normalization: Min-max normalization rescales the values of a feature to a fixed interval, typically [0,1]. This technique preserves the original distribution of the data but changes its scale, making it especially suitable for neural networks using bounded activation functions such as sigmoid or tanh.

$$X_{norm,i} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (2)$$

Robust normalization: Robust normalization scales features using the median and interquartile range (IQR) instead of the mean and standard deviation. This method is designed to be less sensitive to outliers, making it useful for datasets containing extreme values or skewed distributions.

$$X_{norm,i} = \frac{X_i - Q_2(X)}{Q_3(X) - Q_1(X)} \quad (3)$$

The appropriate normalization method depends on both the network and the dataset. Therefore, in the Hungarian Demo, the normalization method was selected during the optimization process for each model.

4.5.3.3 Training process optimization

The performance of an ANN model is largely determined by its architecture and training process. Therefore, in WP6 of the Hungarian Demo, both the specific architecture and the training process were optimized to achieve high accuracy while maintaining generalization. For the computational effectiveness of the optimization, a 3-phase Bayesian optimization method was developed, the 3 phases are listed, then described below:

1. Selection of the normalization method and the network type based on a reference model
2. Bayesian optimization for the network architecture
3. Bayesian optimization for the effectiveness of the training

In the first step, an appropriate normalization method and network type were selected for each model. The normalization method and network type were selected based on their performance on the validation set. For this purpose, a reference model architecture was trained on the training set. The reference model was a two-layer feed-forward or cascade-forward network (depending on the network type), with 20 and 10 neurons in the hidden layers, respectively. The use of a fixed architecture enables straightforward comparison between the different normalization methods and

network types. For this evaluation, only the validation set was used, as modifying the network based on test set performance is strictly prohibited. Such a practice would compromise the independence of the test set and make an objective evaluation of the model's performance impossible.

After initialization, the normalization and model type were selected appropriately. In the next step, an optimization was performed to refine the models' hyperparameters for improved accuracy and an efficient training process.

In the second stage, Bayesian optimization was applied to determine the optimal network architecture, specifically the number of hidden layers and the number of neurons per layer. The best-performing architecture was then selected for the third stage of optimization. This subsequent, third stage also used Bayesian optimization to tune critical hyperparameters, such as the learning rate and momentum. The first Bayesian optimization stage primarily influences overall predictive accuracy, as the network's structural configuration fundamentally determines its capacity to model complex patterns. The second Bayesian optimization stage improves the efficiency and effectiveness of the training process. The performance of the created and trained models was evaluated on the test datasets. The complex evaluation system is described in the next chapter, while the results are described in a separate chapter.

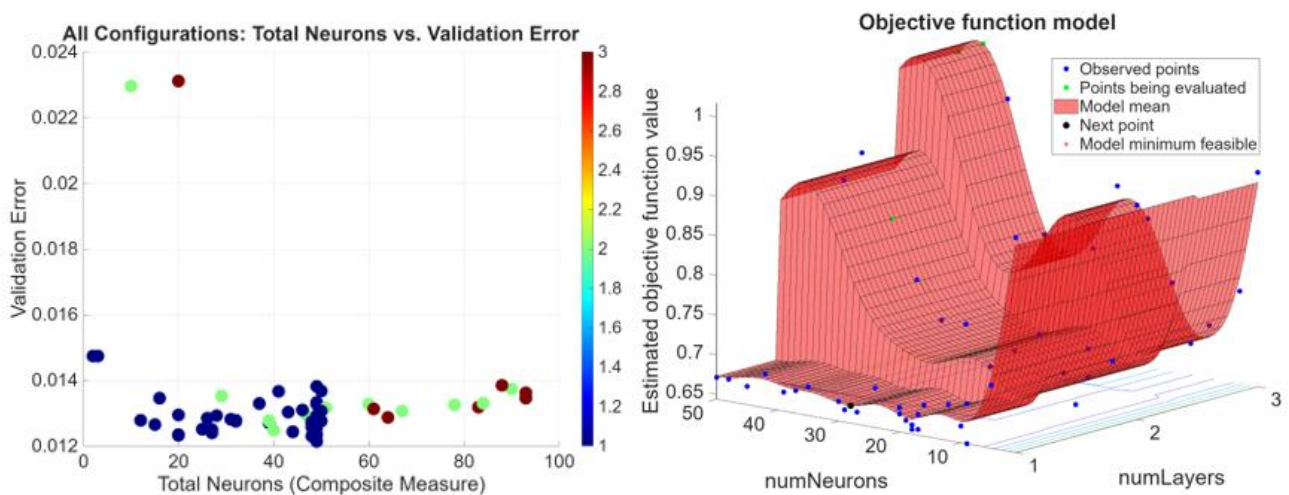


Figure 13 Different ANN configurations and optimization of the momentum and learning rate

As described in Chapter 4.5.3.1, three different splitting methods were evaluated. For all grids, each split method was assessed through the previously described optimization process. From a metric-based perspective, the best-performing approach was the weekday split; however, when generalizability is also considered, the superior split method is the seasonal split. For reference and comparison, all results are presented, with the seasonal split chosen as the final method.

4.5.4 Evaluation system

Evaluating machine learning and deep learning models is a complex task. A widely used approach involves the application of performance metrics and figures for intuitive interpretation; however, models developed for different tasks may exhibit varying behaviour depending on the selected metrics. For instance, in certain scenarios, the sign of the error may have a limited impact, whereas in

others it is of critical importance. Therefore, developing an objective and comprehensive evaluation framework for the specific artificial neural network (ANN) model is essential. This chapter provides an overview of the most commonly used evaluation metrics for assessing ANN model performance and presents the evaluation framework developed for the Hungarian Demo. In this Demo, the model output corresponds to the real-time prediction of the conductor temperature based on the input current and weather variables. Consequently, the problem addressed in this study is a regression task; therefore, the discussion of metrics is restricted to those applicable to regression models.

4.5.4.1 Performance metrics

The accuracy of regression models is commonly characterized by using error-based metrics. In regression tasks, the model's objective is to predict one or more numerical outputs from the input features; therefore, errors can be computed as the differences between the predicted and actual values. The element-wise difference between these values is typically called the residual error vector, which represents the error for each data point. In many cases, performance metrics are derived from the residual vector, in order to summarize the model's accuracy with a single scalar value.

Mean Absolute Error

One of the most commonly used metrics is the Mean Absolute Error (MAE), which is the average absolute difference between the model's predictions and the true values. An advantage of this metric is that it provides an overall measure of accuracy while remaining in the same unit as the target variable. However, a notable limitation is that it does not sufficiently penalize large errors when compared to other metrics.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_{predicted} - y_{true}| \quad (4)$$

Root-Mean Squared Error

Root Mean Squared Error (RMSE) is defined as the square root of the mean squared difference between the model's predictions and the true values, yielding a metric that penalizes large errors while retaining the same units as the target variable. Owing to these properties, RMSE is one of the most commonly used performance metrics.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_{predicted} - y_{true})^2} \quad (5)$$

Coefficient of determination

In neural network regression tasks, the coefficient of determination (R^2) is a statistical metric that measures how well the model's predictions approximate the true target values. It represents the proportion of variance in the dependent variable that the model explains. Within the context of neural networks trained for regression, (R^2) quantifies how effectively the network has learned the relationship between the input features and the output variable. The metric compares the neural network's prediction errors to the variability in the observed data. A value close to 1 indicates that the model closely matches the observed values, while a value near 0 suggests that the model performs no better than predicting the mean of the target variable. Negative values indicate that the model performs worse than this baseline. Therefore, the coefficient of determination provides a useful

measure of model performance by quantifying the proportion of total variation in the observed outcomes that is captured by the model's predictions.

$$SS_{res} = \sum_{i=1}^n (y_{true} - \hat{y}_{predicted})^2 \quad (6)$$

$$SS_{tot} = \sum_{i=1}^n (y_{true} - \bar{y}_{true})^2 \quad (7)$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (8)$$

4.5.4.2 Reliability metrics

In the context of neural networks, a good model is not only accurate but also reliable. This is particularly important in electrical power applications, where the electric grid is complex and critical infrastructure, and even small errors can cause significant system damage. Therefore, the applied models must be evaluated not only for accuracy but also for reliability and robustness.

The previously discussed performance metrics do not fully capture model reliability. While a model with substantially lower error metrics is generally preferable, in cases where two models exhibit similar error levels, the more reliable model is better suited for safety-critical applications. In the context of the demo, the models are applied to power lines, a safety-critical use case. As the models are intended for use during network operation, system operators bear the consequences of potential prediction errors. Consequently, model reliability provides essential information regarding the level of trust that can be placed in the predicted values, thereby supporting informed operational decisions.

Assessing model reliability, however, is a challenging task. In the case of conductor temperature prediction, the model output is a temperature value that can be compared to the corresponding measured value. These measured values serve as a reference during training and evaluation, yet they are also subject to inaccuracies arising from measurement and communication errors. Therefore, it must be considered that the reference values themselves contain uncertainties, which are not explicitly accounted for in the evaluation process.

In summary, the evaluation framework must also address model reliability. In this work, reliability is assessed through outlier analysis. For conductor temperature prediction, large prediction errors may pose safety risks; therefore, in addition to achieving high overall accuracy, it is essential to maintain a low outlier ratio. An error is considered an outlier if it is significantly larger than the majority of errors. The applied metrics for this purpose are described below.

Percentiles

Percentiles provide a measure of model accuracy while accounting for extreme errors. The selected percentile depends on the application; in safety-critical contexts, commonly used values include the 95th, 98th, and 99th percentiles (denoted P95, P98, and P99). The percentile metric defines an interval within which a given percentage of the error values falls. For example, a 95th percentile value of 2.5 °C indicates that 95% of the errors lie within the range of 0 to 2.5 °C. This metric reflects model reliability by indicating that, in 95% of cases, the prediction error remains below 2.5 °C. From an end-user perspective, this can be interpreted as achieving an error of at most 2.5 °C with 95% confidence. In our case, the results will contain the P95 value for every model.

Outlier percentage

The outlier percentage is a metric that quantifies the proportion of outliers relative to the entire dataset. In this context, outliers are defined based on the errors between the predicted and actual values. Although several methods exist for outlier detection, their common objective is to identify values that deviate significantly from the majority of the data. In this demo, a prespecified error limit is applied, corresponding to the initial KPI accuracy. This approach identifies outliers based on the pre-determined ± 2 °C limit set for accuracy. Therefore, an error is considered as an outlier if it resides outside of the ± 2 °C range. The proportion of outliers in the entire dataset, expressed as a percentage, is called the outlier percentage. From a reliability perspective, a lower outlier ratio is desirable; however, due to data imperfections, measurement errors, and the inherent characteristics of artificial neural networks, a small number of larger errors is unavoidable. Therefore, in addition to improving overall accuracy, a key objective of neural network training optimization is the reduction of the outlier percentage.

4.5.4.3 Graphical representation of metrics

While the various metrics provide useful, easily interpretable results, visualizations offer several additional advantages in the evaluation process. Performance metrics summarize model accuracy into single numerical values, thereby limiting the amount of information conveyed. In contrast, plots can capture multiple aspects of model behaviour simultaneously and facilitate more intuitive comparisons.

Therefore, a comprehensive evaluation framework should include relevant, information-rich, and easily interpretable visualizations. In the context of neural networks, several types of plots are commonly employed; the following section describes the visualizations used in this demo.

Boxplot

Boxplots are widely used to describe data distributions. In this demo, they are applied to represent the error distribution of the evaluated models. Typically, and in our case as well, boxplots are constructed based on percentile values. The box captures the central 50% of the data, spanning from the 25th to the 75th percentile, while the whiskers extend to the minimum and maximum values within the non-outlier range. Data points lying outside the whiskers are classified as boxplot outliers.

Boxplots are particularly useful as they convey multiple characteristics of the distribution in a compact form and enable straightforward comparison between models. The spread of the data can be assessed from the width of the box and the length of the whiskers, while potential bias can be inferred from the position of the median within the box. For regression models, the objective is to minimize the spread of the errors while avoiding systematic bias. In terms of the boxplot representation, this corresponds to a narrow box and whiskers, with the median near zero.

In addition to their descriptive value, boxplots play an important role in the evaluation of the project KPIs. KPI No. 3 is directly based on the whisker range of the error distribution, as the target was defined as achieving a whisker range within ± 2 °C. KPI No. 4 further utilizes the boxplot representation by evaluating both the whisker range, referred to as the maximum accuracy, and the width of the box itself, referred to as the average accuracy achieved by approximately 50% of the predictions. Consequently, models exhibiting narrow boxes and short whiskers are considered more accurate and reliable.

Time-series plot

While boxplots provide substantial information about the distribution of errors, they do not capture the model's behaviour under real-time conditions. In contrast, time-series plots display both actual and predicted values over time, enabling visualization of temporal trends and model dynamics.

Residuals plot

The residuals plot is similar to the time-series plot in that it represents data as a function of time. However, instead of displaying the actual and predicted values, it visualizes their differences, referred to as residuals. This representation facilitates the identification of systematic errors, bias, and potential drift in the model's predictions.

Scatter plot

Scatter plots can be used for various purposes; however, in the context of neural networks, a specific type is commonly applied for model evaluation. In this plot, the predicted values are plotted as a function of the corresponding true values, effectively visualizing the regression model's performance. The ideal case corresponds to a 45° line, indicating that the predicted values are equal to the actual values, and thus the predictions are perfect. A good model fit is characterized by the majority of data points lying in close proximity to this line. Conversely, a more dispersed distribution of points indicates lower model accuracy.

5 Results and Discussion

The following chapters present the project's results, summarizing the key findings and the extent to which the initially defined KPIs have been achieved. As previously stated, the training process was carried out via an optimization procedure using three splitting methods. Since the data for each of the three grids consists of two distinct measurement points, and each point was trained using both 6 and 12 months of data, a total of twelve distinct models were developed for each grid. From these, four models (corresponding to the two measurement points trained with 6 and 12 months of data using the seasonal split method), were selected as the final models for each grid. A summary of the created models, their final parameters and the achieved values are summarized in Table 14.

The large number of evaluated models makes interpreting the results more complex. Therefore, the results are presented separately for each grid. Although the results obtained using all applied split methods are included, only a limited number of plots are presented for the worse-performing weekday and time-series splits to improve clarity and readability. The presentation of the results for each grid is followed by a comparative discussion and an overall evaluation.

Table 14 Description of the model names and results achieved during the project (seasonal split)

Grid	Selected power line	Measurement point on the selected power line	Name of models created for the measurement point	Length of training	Model codename	Final Model Parameters and Achieved Values			
						Architecture	Normalization	RMSE [°C]	P95 [°C]
APG	A1	APG1	A1	6 months	A1_6m	CFN	Z-Score	3.015	5.853
				12 months	A1_12m	FFN	Robust	2.927	5.579
		APG2	A2	6 months	A2_6m	FFN	None	1.930	3.984
				12 months	A2_12m	FFN	Min-Max	1.899	3.828
MAVIR	M1	MAVIR 2	B1	6 months	B1_6m	FFN	Min-Max	0.982	1.965
				12 months	B1_12m	FFN	Min-Max	0.962	1.946
		MAVIR 3	B2	6 months	B2_6m	CFN	Min-Max	1.024	2.113
				12 months	B2_12m	CFN	Robust	1.087	2.161
E.ON	E1	E.ON 1	C1	6 months	C1_6m	CFN	Min-Max	0.792	1.647
				12 months	C1_12m	CFN	Min-Max	0.826	1.650
		E.ON 2	C2	6 months	C2_6m	CFN	Min-Max	0.932	1.943
				12 months	C2_12m	CFN	Robust	0.791	1.538

5.1 Results for APG grid

As expressed in Chapter 4, the APG grid line monitoring system differs from the others regarding sensor setup, data sources, and data characteristics. Consequently, the preliminary data processing for APG had to follow a different approach from the other grids. The Type B sensor directly measures the conductor's temperature. The real-time load is obtained from the SCADA system. Since there is no local weather station, weather parameters had to be sourced separately, requiring a customized data integration process.

- **Data Selection for Modeling:**
 - After preprocessing, cleaning, and aligning the database, **multiple measurement points** across several lines were available.
 - Two monitoring points — **APG 1 and APG 2** — were selected for **model training** on the A1 line.
- **Data Characteristics:**
 - APG's dataset is recorded at **15-minute intervals** (4 records/hour), **less frequently** than in other grids.
 - This results in a **smaller overall data volume**, especially noticeable over the same period (e.g., 6 months).
- **Data Availability:**
 - Despite the lower frequency, the **availability rate** is remarkably high, **91.8%** for APG 1 and **91.3%** for APG 2.
 - This **high availability** helps mitigate the effects of the smaller dataset size.
- **Training process:**
 - Four different models were trained with each split method, corresponding to the two distinct grid points APG 1 and APG 2 — denoted as **A1 and A2**, respectively— with 6 and 12 months of training — denoted as **_6m and _12m**, respectively.

Preprocessing was followed by ANN model training. As explained in Chapter 4.5, the data split was performed using the three different methods. Each split was followed by training the models using the described optimization method. The results are shown in the following chapters.

5.1.1 Results for APG grid with the time-series split

The results are shown in Table 15 and Figure 14 -Figure 15.

Table 15 Metrics for APG with time-series split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
APG 1	A1_6m	Z-score	3.032	2.255	0.835	6.178	44.765
	A1_12m	Min-Max	3.700	2.877	0.755	7.216	55.973
APG 2	A2_6m	Min-Max	2.229	1.562	0.916	4.616	26.134
	A2_12m	Robust	2.486	1.712	0.895	5.117	29.496

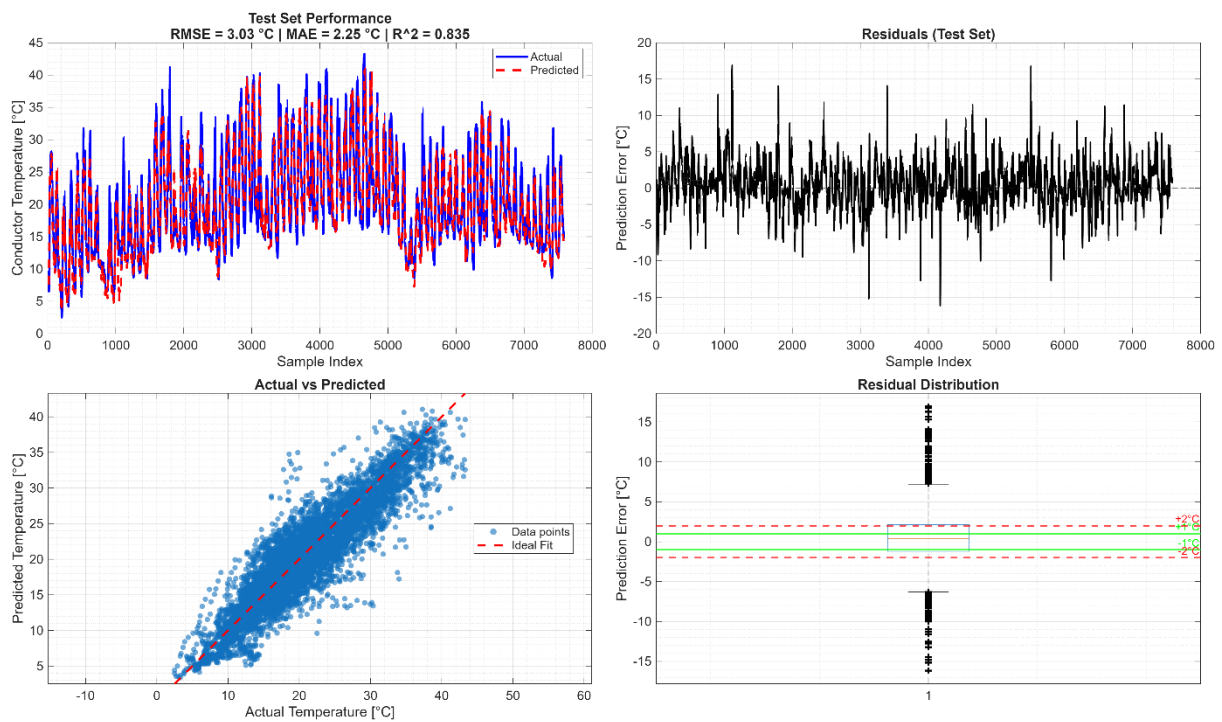


Figure 14 Plots for model A1_6m with time-series split

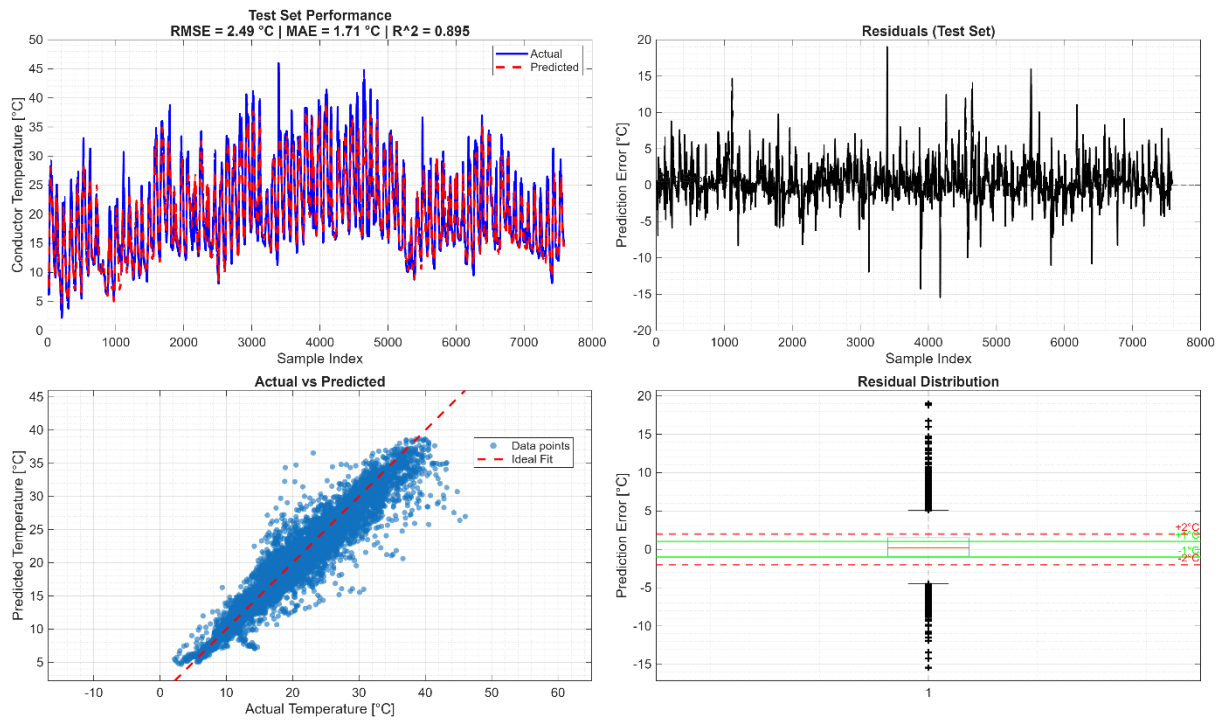


Figure 15 Plots for model A2_12m with time-series split

5.1.2 Results for APG grid with the seasonal split

The results are shown in Table 16 and Figure 16 -Figure 17.

Table 16 Metrics for APG with seasonal split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
APG 1	A1_6m	Z-score	3.015	2.192	0.911	5.853	42.924
	A1_12m	Robust	2.927	2.190	0.907	5.579	43.421
APG 2	A2_6m	None	1.930	1.403	0.963	3.984	22.813
	A2_12m	Min-Max	1.899	1.385	0.955	3.828	23.755

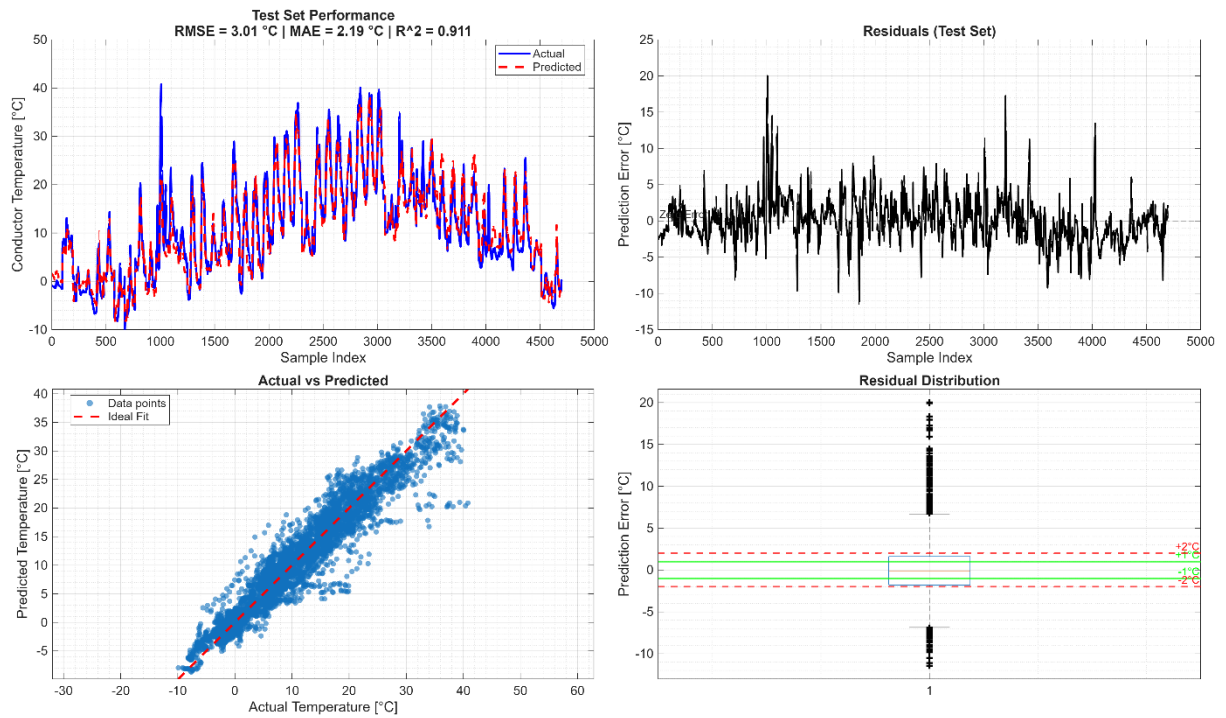


Figure 16 Plots for model A1_6m with seasonal split

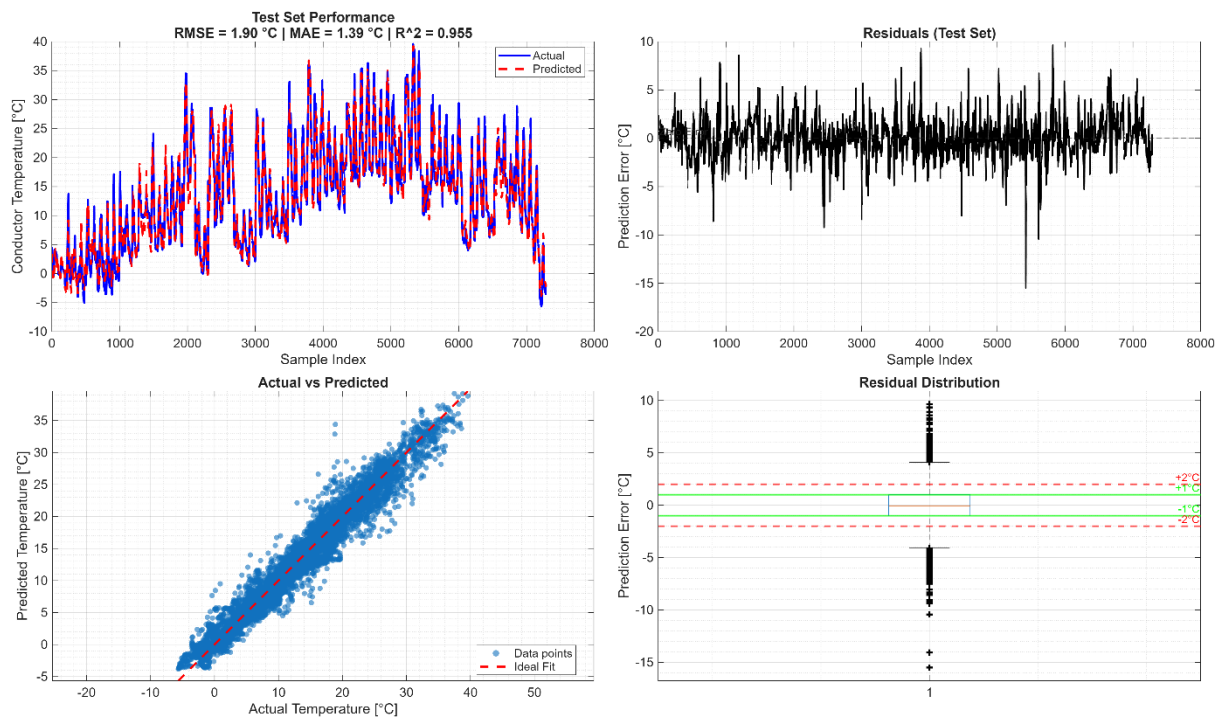


Figure 17 Plots for model A2_12m with seasonal split

5.1.3 Results for APG grid with the weekday split

The results are shown in Table 17 and Figure 18 - Figure 19.

Table 17 Metrics for APG with weekday split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
APG 1	A1_6m	Min-Max	2.865	2.153	0.906	5.764	42.430
	A1_12m	None	2.981	2.269	0.883	5.989	44.881
APG 2	A2_6m	None	2.238	1.612	0.946	4.644	28.079
	A2_12m	Z-score	2.141	1.517	0.945	4.459	26.658

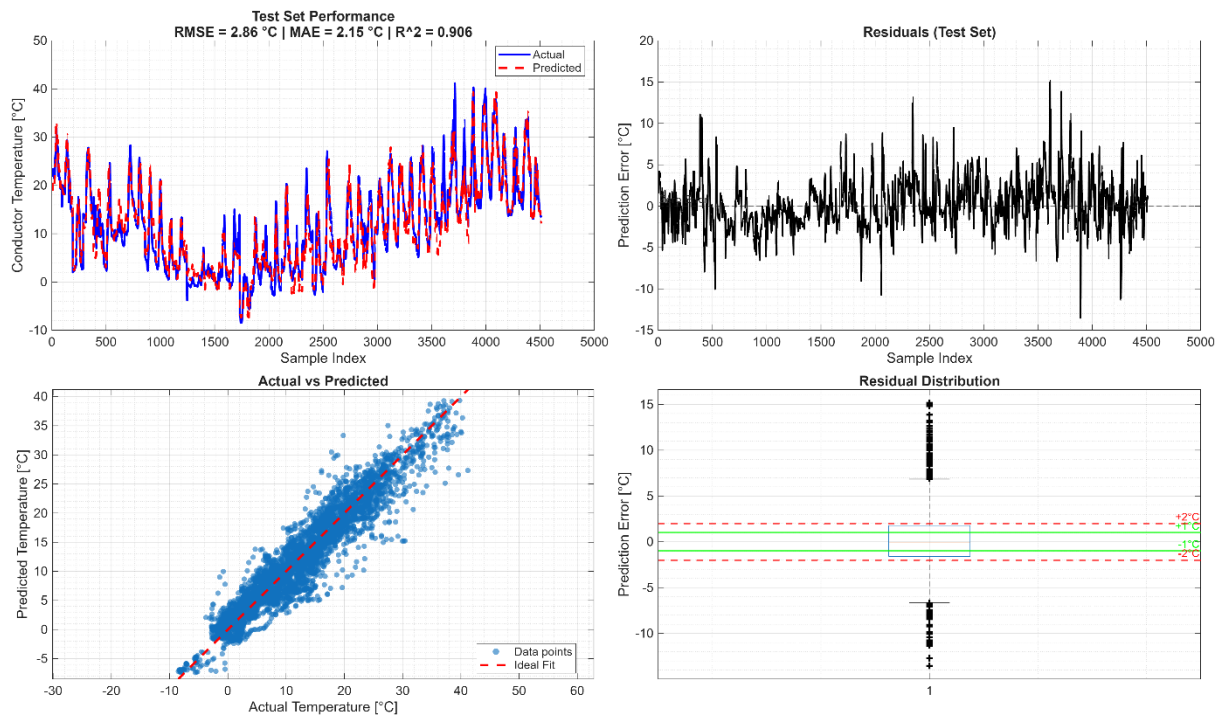


Figure 18 Plots for model A1_6m with weekday split

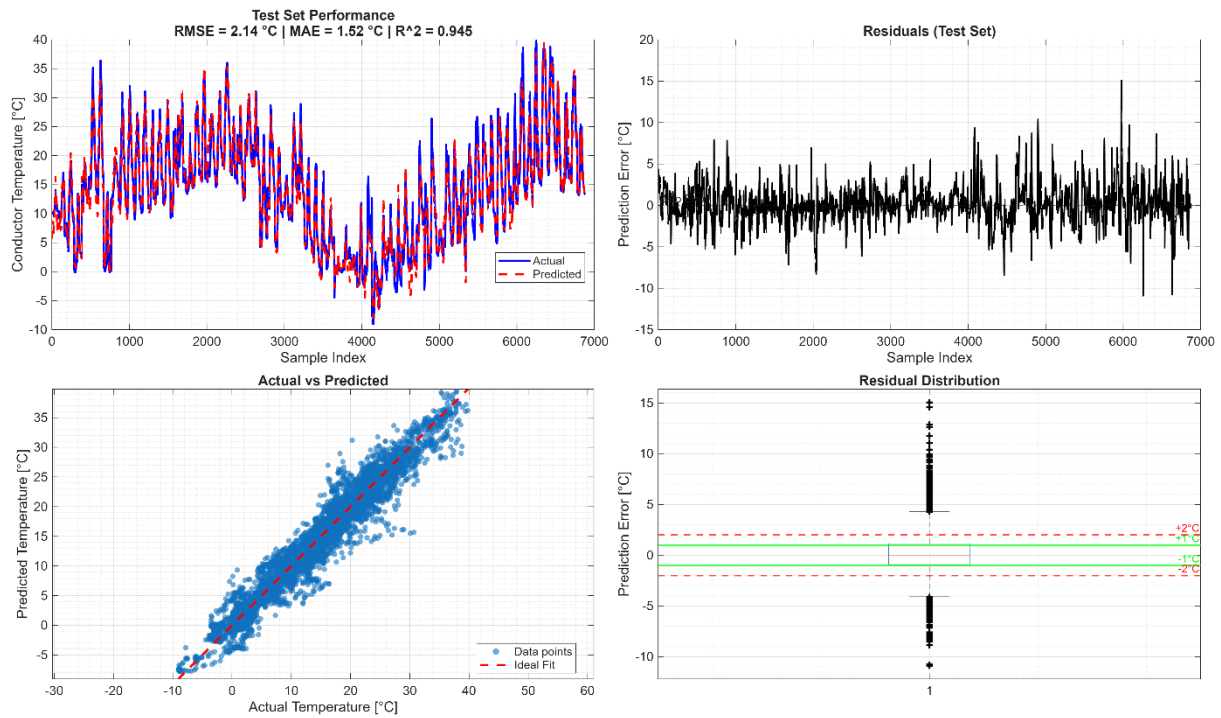


Figure 19 Plots for model A2_12m with weekday split

5.1.4 Discussion of results for APG grid

The results of the neural network optimization in case of APG does not achieve the desired accuracy. This was anticipated, as of the lack of local measurements and the preliminary results. During the early stages of the project, when the first models were created and evaluated, it was determined that the lack of local measurements significantly affected the accuracy of the conductor's temperature prediction. However, the use of more complex models and optimization processes enabled more accurate and reliable results in the later stages of the project.

The results indicate that one of the APG models, specifically the A2 model, produces slightly more accurate predictions than the A1 model. This observation is consistent across all applied split methods and for both training durations. It should be emphasized that the optimization and training procedures were identical across all models, using the same time span of data within each grid. Consequently, better performance at a specific grid point primarily results from improved data quality.

With the higher-quality data, model A2_12m achieved the best performance when using the seasonal split method. In this configuration, the RMSE was 1.899 °C, while the MAE reached 1.385 °C, indicating an average prediction error of approximately 1.5 °C. Although these accuracy metrics may appear promising at first, the reliability metrics reveal considerable uncertainty in the predictions. Despite the relatively low average errors, the P95 value was 3.828 °C and the outlier percentage reached 23.755%. This indicates that a substantial proportion of predictions exhibit significant deviations, with nearly one-quarter falling outside the ± 2 °C target range. It is important to highlight, however, that this does not mean the APG lacks the appropriate input data for line monitoring systems. A defined objective of the project was to examine the significance of local weather data during training; this is why these inputs were utilized.

For measurement point A1, the weekday split method performed best, whereas the seasonal split was optimal for A2. In this case, model A1_6m trained using the weekday split achieved the best results for this measurement point. The obtained accuracies were 2.865 °C (RMSE) and 2.153 °C (MAE). The reliability metrics indicate lower reliability than those for the models developed for measurement point A2, consistent with the lower overall accuracy. For measurement point A1, the outlier percentage ranged between approximately 40% and 45%, depending on the specific configuration. This implies that nearly half of the predictions fall outside the desired ± 2 °C interval.

Regarding the influence of training duration, the results indicate that, for measurement point A2, a longer training period improved performance for both the weekday and seasonal split methods. For measurement point A1, the longer training period improved accuracy only for the seasonal split; however, it significantly reduced prediction accuracy when the time-series method was applied.

5.2 Results for MAVIR grid

For the MAVIR grid, measurement data from the A-type line monitoring sensor and a weather station are available. The real-time load was retrieved from the SCADA system. The first step was data preprocessing, aligning the datasets, and removing incorrect records.

- **Data Selection for Modeling:**
 - After preprocessing, cleaning, and aligning the database, **two measurement points** across several lines were available.
 - These monitoring points — **MAVIR 2 and MAVIR 3** — were selected for **model training** on the M1 line.
- **Data Characteristics:**
 - MAVIR's dataset is recorded at **10-minute intervals** (6 records/hour), **more frequently** than in APG grids.
 - This results in a **greater overall data volume**, especially noticeable over the same period (e.g., 6 months).
- **Data Availability:**
 - The **availability rate** is not as high as for APG, **86.0%** for MAVIR 2, and **82.7%** for MAVIR 3.
 - This **lower availability** indicates the importance of carefully choosing the training period to make enough quality and quantity data.
- **Training process:**
 - Four different models were trained with each split method, corresponding to the two distinct grid points **MAVIR 2 and MAVIR 3** — denoted as **B1 and B2**, respectively— with 6 and 12 months of training — denoted as **_6m and _12m**, respectively.

Preprocessing was followed by ANN model training. As explained in Chapter 4.5, the data split was performed using the three different methods. Each of the splits was followed by the training of the models according to the optimization method described. The results are shown in the following chapters.

5.2.1 Results for MAVIR grid with the time-series split

The results are shown in Table 18 and Figure 20 - Figure 21.

Table 18 Metrics for MAVIR with time-series split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
MAVIR 2	B1_6m	Min-Max	1.366	0.896	0.956	2.956	10.453
	B1_12m	Z-score	1.359	0.906	0.957	2.892	10.546
MAVIR 3	B2_6m	Z-score	1.354	0.878	0.955	2.716	9.752
	B2_12m	Min-Max	1.297	0.859	0.959	2.586	9.083

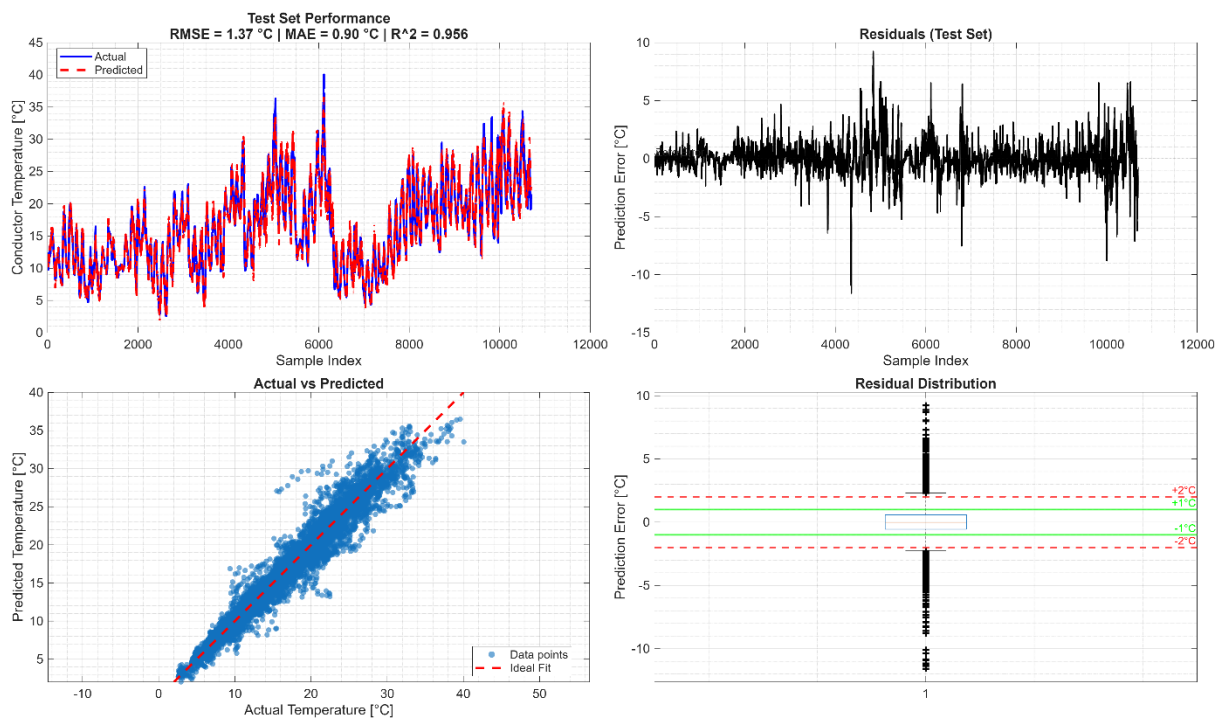


Figure 20 Plots for model B1_6m with weekday split

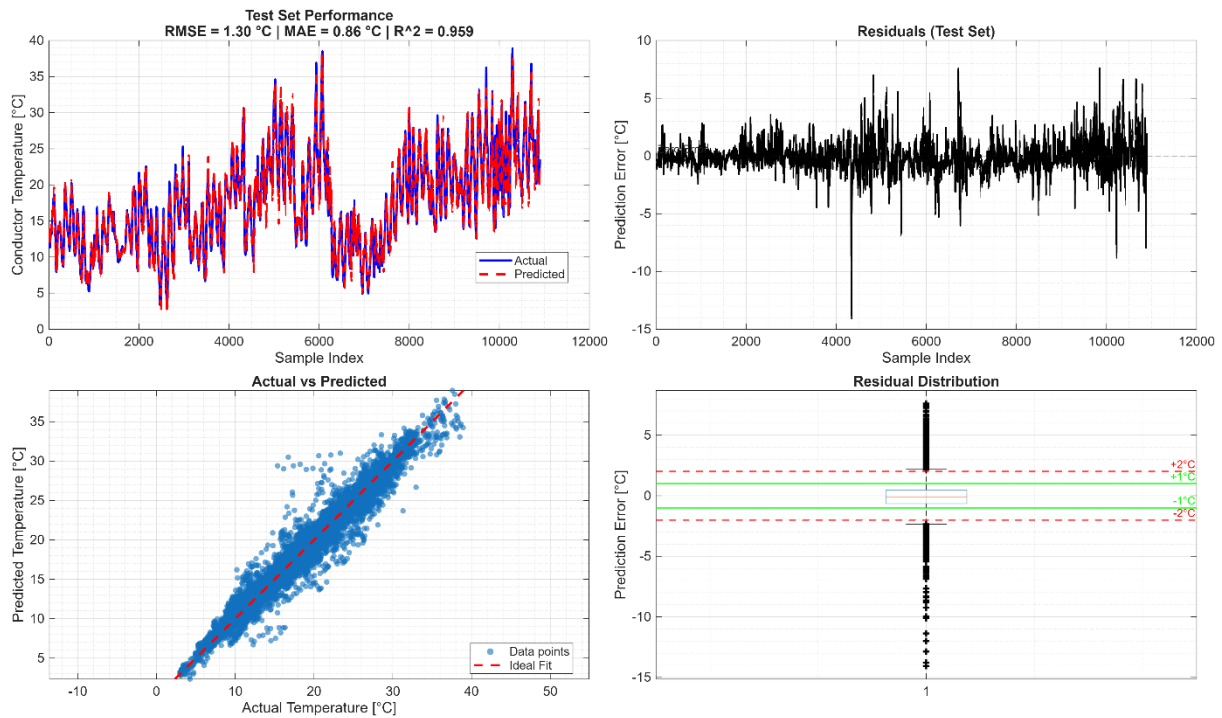


Figure 21 Plots for model B2_12m with time-series split

5.2.2 Results for MAVIR grid with the seasonal split

The results are shown in Table 19 and Figure 22 -Figure 23.

Table 19 Metrics for MAVIR with seasonal split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
MAVIR 2	B1_6m	Min-Max	0.982	0.664	0.989	1.965	4.802
	B1_12m	Min-Max	0.962	0.662	0.989	1.946	4.622
MAVIR 3	B2_6m	Min-Max	1.024	0.725	0.985	2.113	5.930
	B2_12m	Robust	1.087	0.755	0.984	2.161	6.200

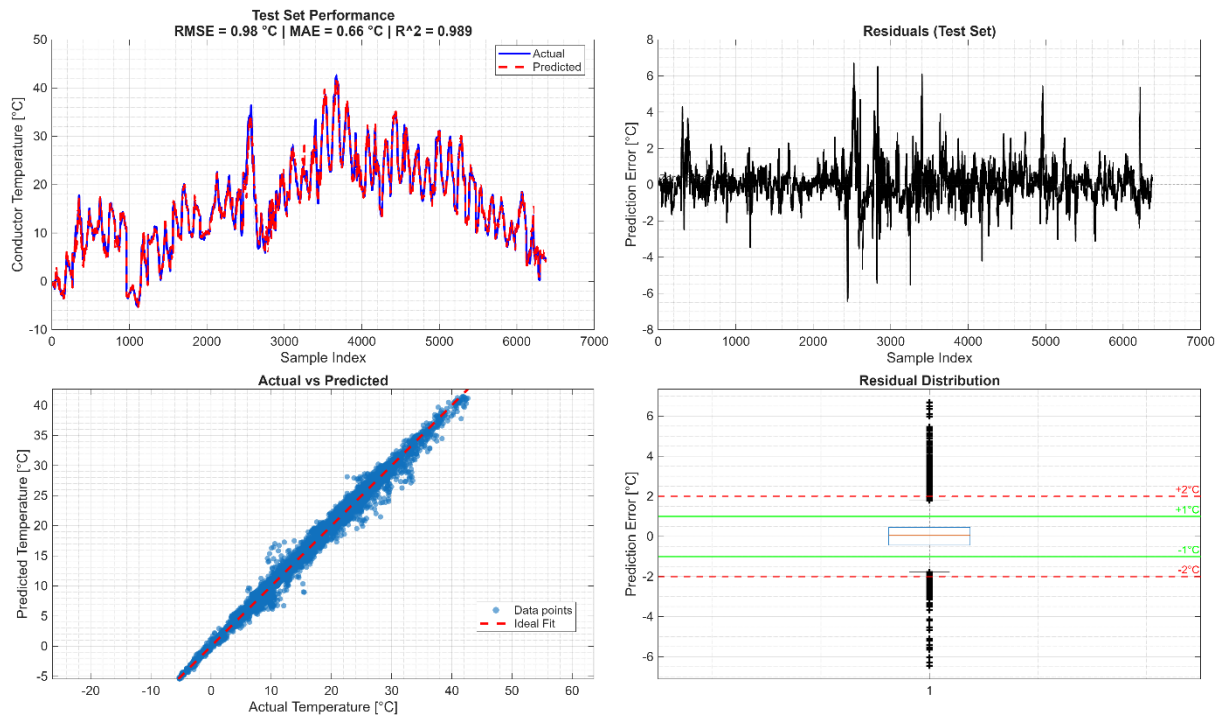


Figure 22 Plots for model B2_6m with seasonal split

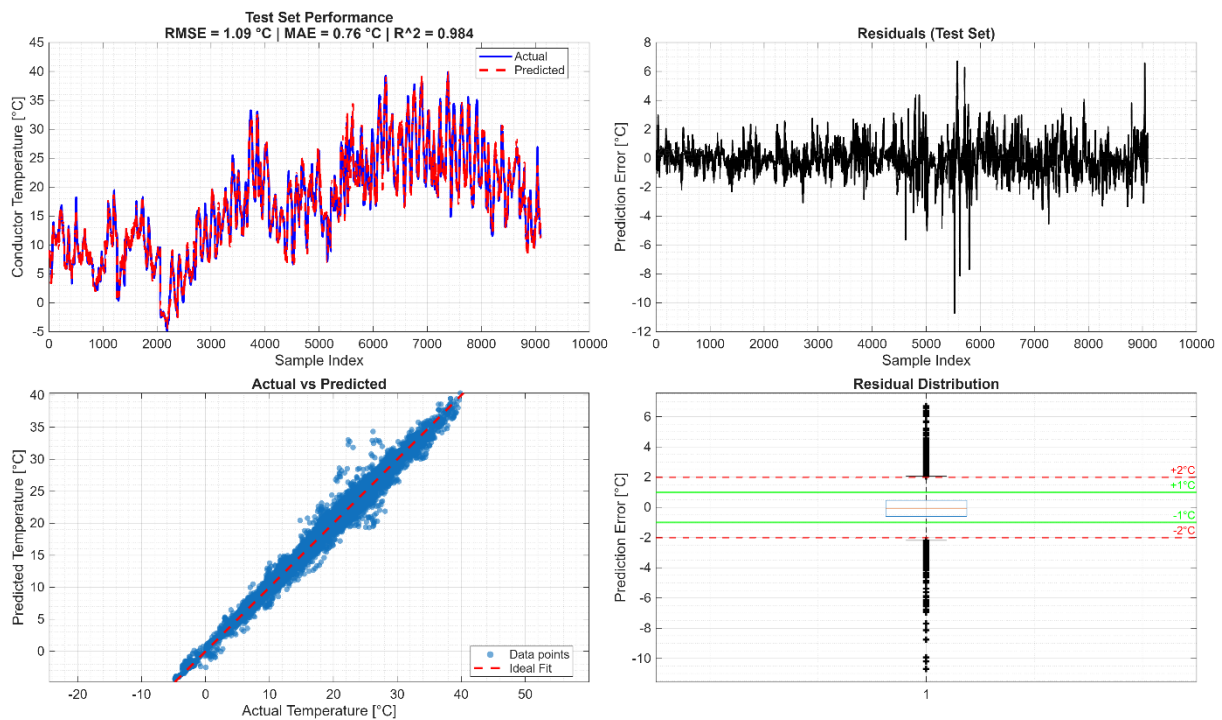


Figure 23 Plots for model B2_12m with seasonal split

5.2.3 Results for MAVIR grid with the weekday split

The results are shown in Table 20 and Figure 24 - Figure 25.

Table 20 Metrics for MAVIR with weekday split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
MAVIR 2	B1_6m	Z-score	1.027	0.691	0.990	2.203	6.370
	B1_12m	Z-score	0.917	0.608	0.990	1.901	4.444
MAVIR 3	B2_6m	Robust	0.986	0.678	0.988	2.007	5.056
	B2_12m	Min-Max	0.897	0.638	0.989	1.795	3.834

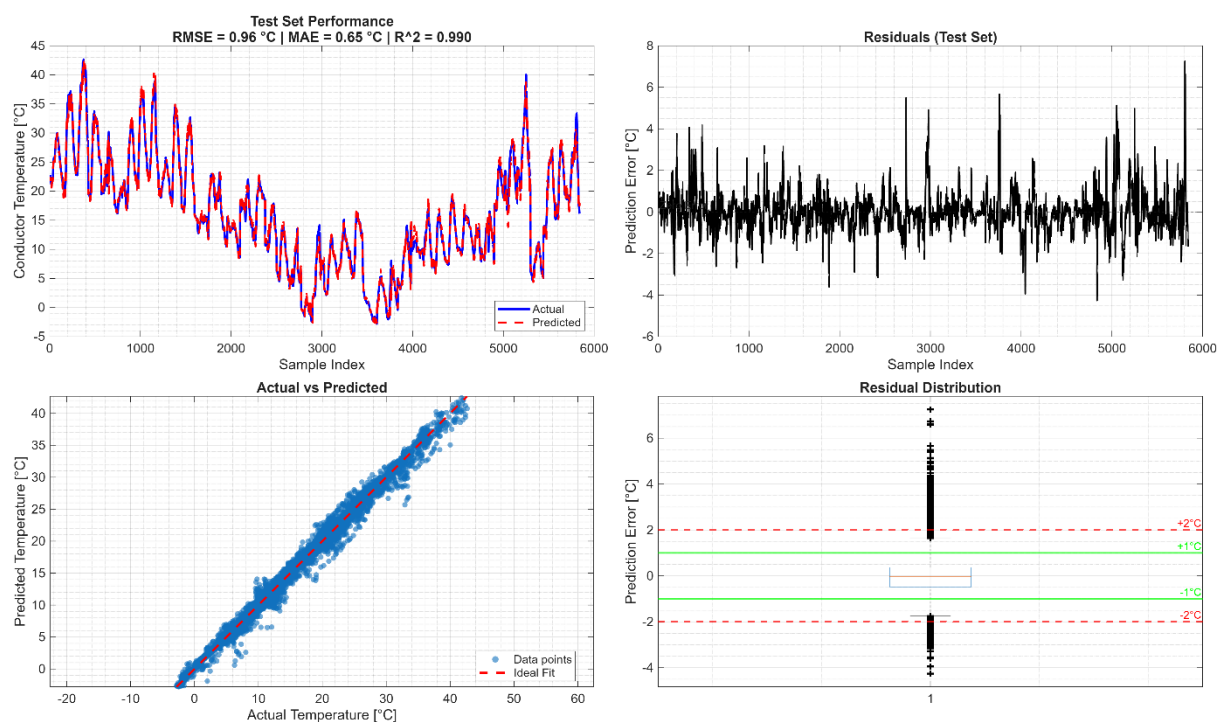


Figure 24 Plots for model B1_6m with weekday split

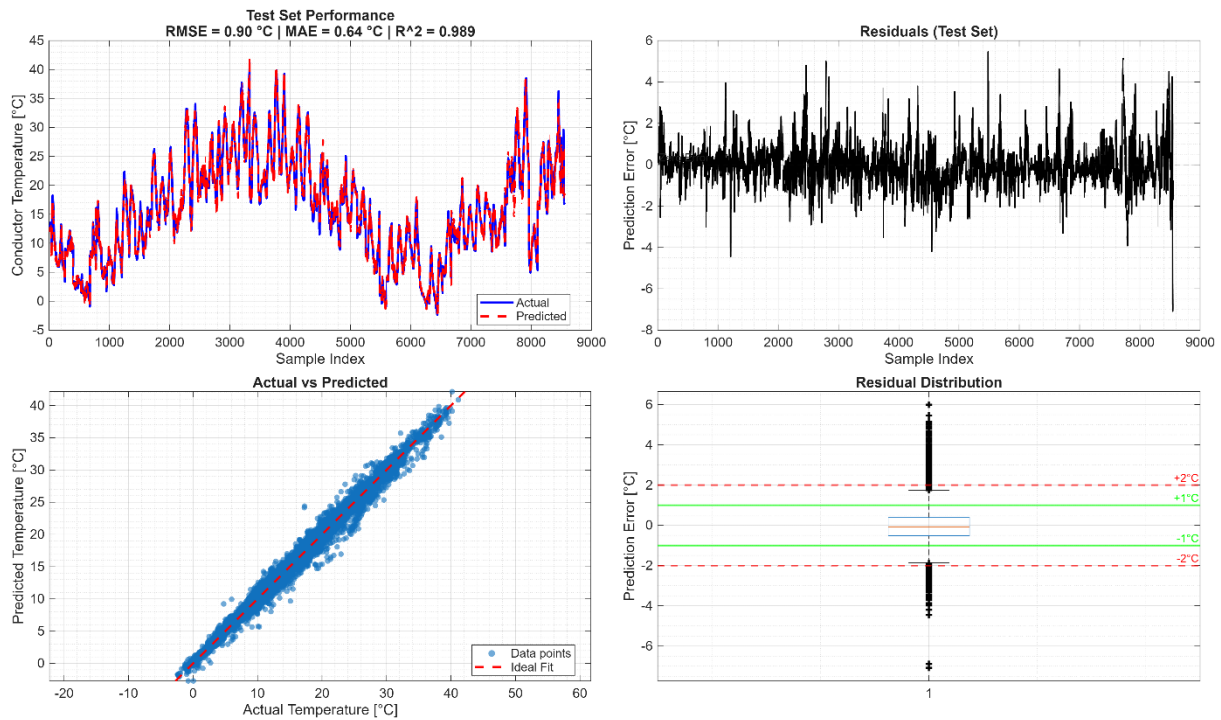


Figure 25 Plots for model B2_12m with weekday split

5.2.4 Discussion of results for MAVIR grid

In the case of MAVIR, the results are significantly better than those obtained for APG. As discussed previously, the absence of local weather measurements appears to have a greater impact on model performance than the APG dataset's lower temporal resolution. In the MAVIR case, weather measurements are collected locally and recorded at 10-minute resolution. However, MAVIR's data showed lower availability compared to the APG dataset. The observed gaps were primarily related to the characteristics of the measurement infrastructure and the operating conditions of the monitored lines. The dataset combines measurements from line-mounted conductor sensors and weather stations installed along the overhead lines. Consequently, data availability depends not only on the continuous operation of the measurement devices but also on the reliability of their communication and data acquisition infrastructure. Temporary communication failures, sensor or device related issues, and interruptions within the data transmission chain can result in missing measurements. In addition, the monitored overhead lines are not continuously in service, as they may be temporarily de-energized for planned maintenance, network operations, power flow management, or protection-related events. During such periods, conductor temperature and loading measurements are either unavailable or have limited analytical relevance. Weather stations with an independent power supply, however, may continue recording during these intervals, resulting in differences in data availability between the two measurement types. Where temporary communication failures occurred, locally buffered measurements could in some cases be recovered and subsequently incorporated into the dataset. These characteristics explain why lower nominal data availability does not necessarily indicate poorer data quality or reduced suitability for ANN model development.

This is also reflected in the consistency of the prediction results obtained at the two MAVIR measurement points. In contrast to APG, the two distinct measurement points in the MAVIR dataset do not exhibit substantial differences in prediction accuracy, as both B1 and B2 show similar performance. The best-performing split method in both cases is the weekday split. The seasonal split achieves comparable results, while the time-series split yields lower accuracy. The RMSE values for the two points are approximately 0.9 °C, 1.0 °C, and 1.3 °C for the weekday, seasonal, and time-series splits, respectively.

Considering the influence of training duration, the results clearly indicate that a longer training period improves prediction accuracy in the MAVIR case. In all configurations, except for measurement point B1 with the seasonal split, the longer training period resulted in improved model performance.

For MAVIR, the weekday split is mathematically the best-performing approach, achieving the highest overall accuracy among the evaluated methods. The results obtained are substantially better than in the APG case, consistent with the preliminary observations. The RMSE values are 0.917 °C and 0.897 °C, while the corresponding MAE values are 0.608 °C and 0.638 °C for models B1_12m and B2_12m, respectively.

However, despite the superior mathematical performance of the weekday split, the seasonal split is considered particularly important for generalizability. Since conductor temperature prediction is strongly influenced by seasonal meteorological factors, the seasonal split provides a more realistic evaluation of the models' ability to generalize to unseen operating conditions across different seasons. In this context, the seasonal split also achieves very strong performance. The RMSE values for the two models with the longer training periods are 0.962 °C and 1.087 °C, while the corresponding MAE values are 0.662 °C and 0.755 °C, respectively. These results indicate that the models maintain high predictive accuracy even under more challenging, representative seasonal conditions.

The reliability metrics also demonstrate considerable improvement. While the APG models exhibited outlier percentages in the range of approximately 25–44%, the best-performing MAVIR model using the seasonal split produces outliers in less than 5% of cases. Specifically, the outlier percentage is 4.622% for model B1_12m and 6.200% for model B2_12m. The P95 metric further confirms the models' reliability, with values of 1.946 °C for measurement point B1 and 2.161 °C for measurement point B2.

5.3 Results for E.ON grid

The line monitoring system applied at the E.ON grid is similar to the system at the MAVIR side. Data from the A-type line monitoring sensor and a weather station are available. The real-time load was retrieved from the SCADA system. The first step was data preprocessing, aligning the datasets, and removing incorrect records.

- **Data Selection for Modeling:**
 - After preprocessing, cleaning, and aligning the database, **two measurement points** across several lines were available.
 - These monitoring points — **E.ON 1 and E.ON 2** — were selected for **model training** on the E1 line.

- **Data Characteristics:**
 - E.ON's dataset is recorded at **10-minute intervals** (6 records/hour), **more frequently** than in APG grids.
 - This results in a **greater overall data volume**, especially noticeable over the same period (e.g., 6 months).
- **Data Availability:**
 - The **availability rate** is not as high as for APG and MAVIR. It is **70.8%** for E.ON 1 and **62.7%** for E.ON 2.
 - This **lower availability** indicates the importance of carefully choosing the training period to collect enough quality and quantity data
- **Training process:**
 - Four different models were trained with each split method, corresponding to the two distinct grid points **E.ON 1** and **E.ON 2**— denoted as **C1** and **C2**, respectively— with 6 and 12 months of training — denoted as **_6m** and **_12m**, respectively.

Preprocessing was followed by ANN model training. As explained in Chapter 4.5, the data split was performed using the three different methods. Each split was followed by training the models using the optimization method described. The results are shown in the following chapters.

5.3.1 Results for E.ON grid with the time-series split

The results are shown in Table 21 and Figure 26 - Figure 27.

Table 21 Metrics for E.ON with time-series split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
E.ON 1	C1_6m	Robust	0.651	0.473	0.989	1.357	1.182
	C1_12m	Min-Max	0.683	0.500	0.988	1.410	1.355
E.ON 2	C2_6m	Min-Max	0.813	0.615	0.982	1.627	2.429
	C2_12m	Min-Max	0.863	0.663	0.980	2.526	3.005

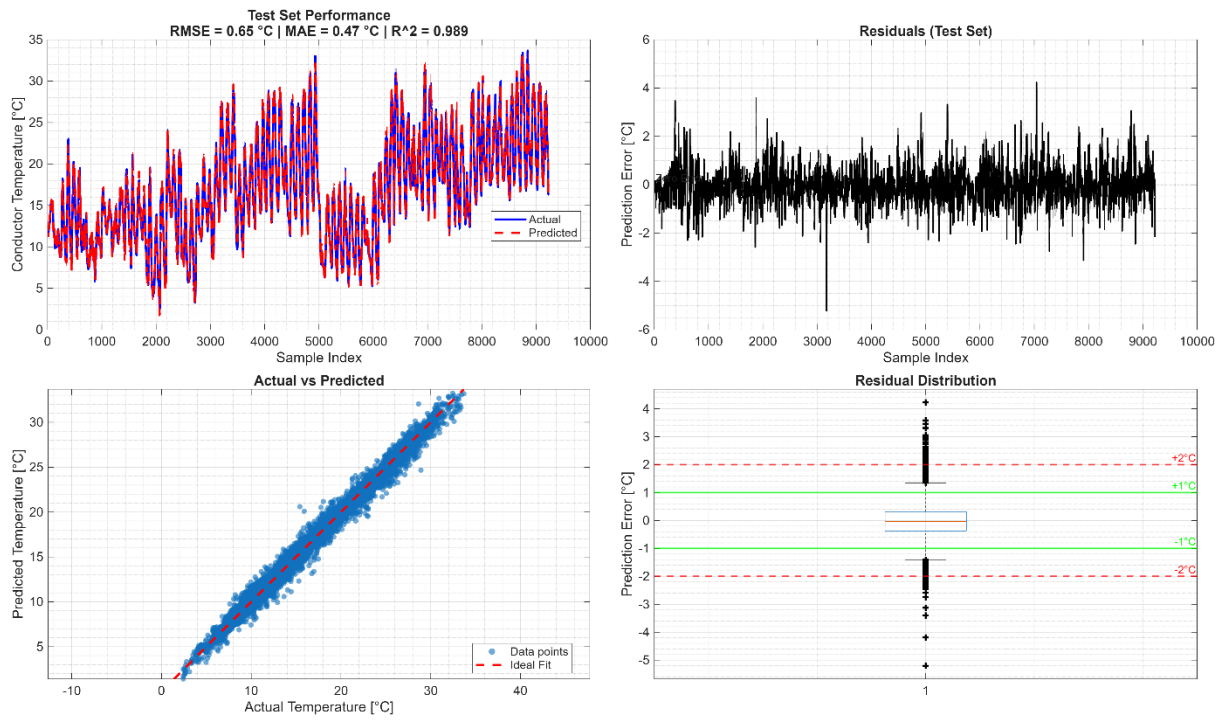


Figure 26 Plots for model C1_6m with time-series split

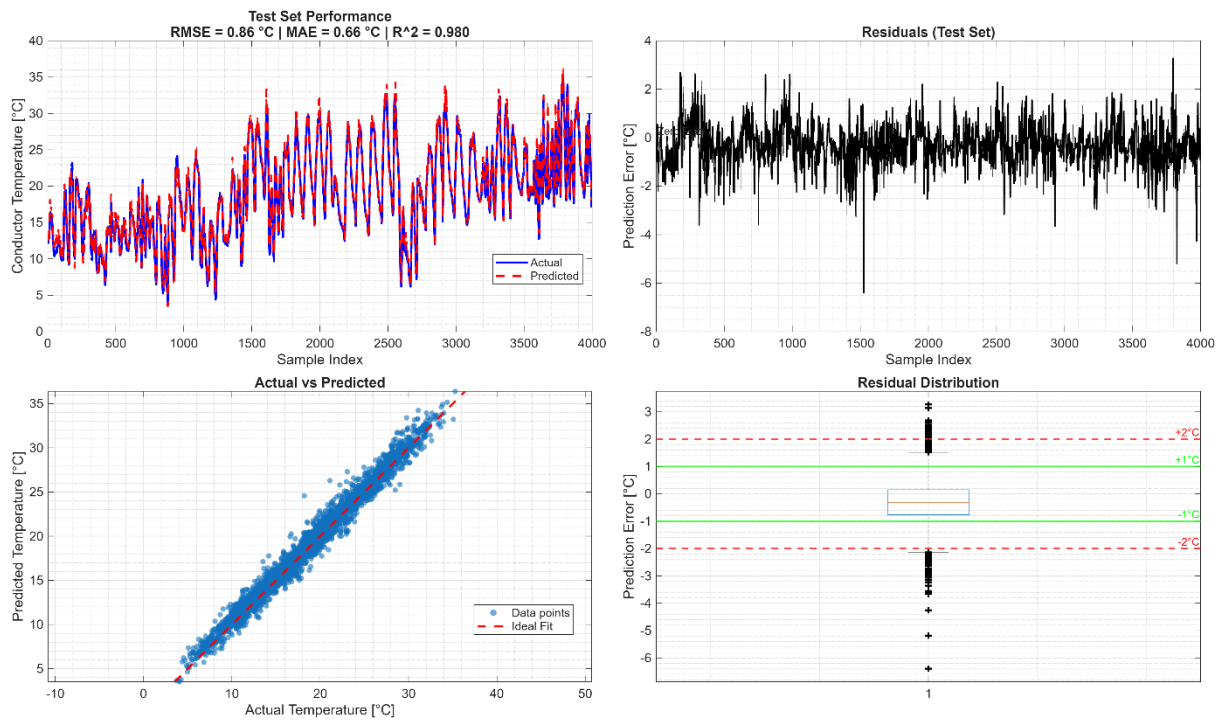


Figure 27 Plots for model C2_12m with time-series split

5.3.2 Results for E.ON grid with the seasonal split

The results are shown in Table 22 and Figure 28 - Figure 29.

Table 22 Metrics for E.ON with seasonal split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
E.ON 1	C1_6m	Min-Max	0.792	0.549	0.992	1.647	2.676
	C1_12m	Min-Max	0.826	0.551	0.991	1.650	3.076
E.ON 2	C2_6m	Min-Max	0.932	0.655	0.990	1.943	4.532
	C2_12m	Robust	0.791	0.556	0.992	1.538	2.524

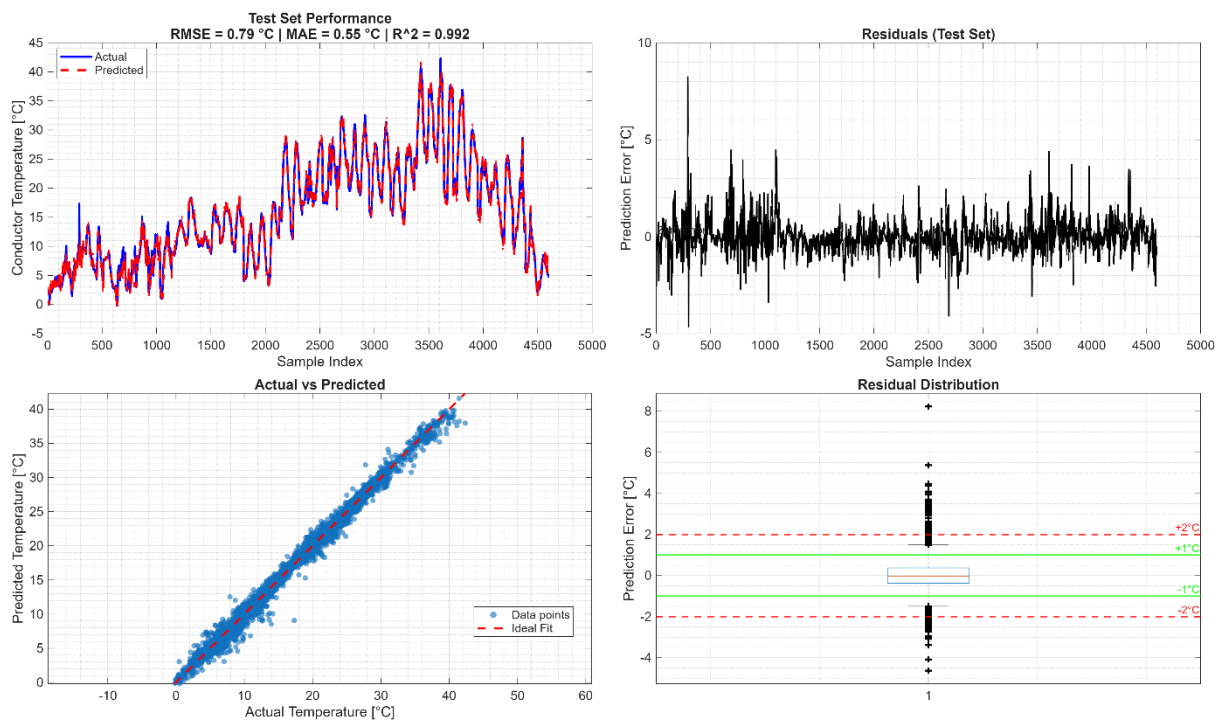


Figure 28 Plots for model C1_6m with seasonal split

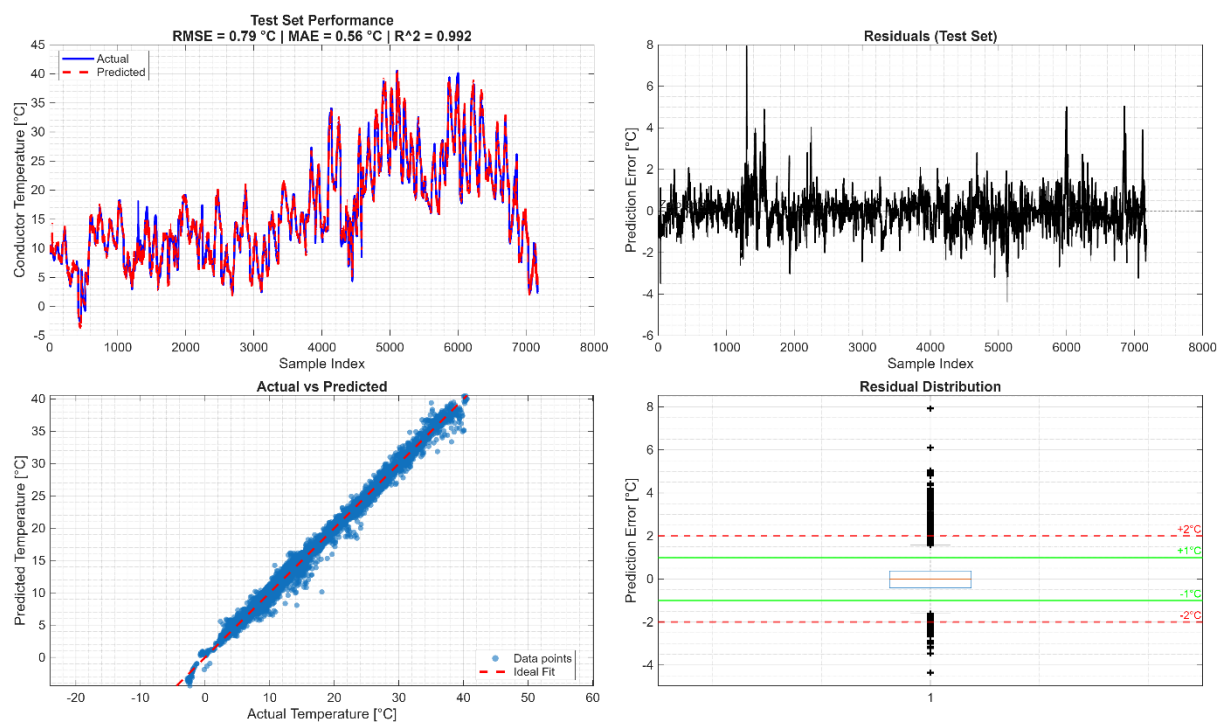


Figure 29 Plots for model C2_12m with seasonal split

5.3.3 Results for E.ON grid with the weekday split

The results are shown in Table 23 and Figure 30 - Figure 31.

Table 23 Metrics for E.ON with weekday split

Data point	Model	Normalization	RMSE [°C]	MAE [°C]	R ² [-]	P95 [°C]	Outlier percentage [%]
E.ON 1	C1_6m	Min-Max	0.643	0.469	0.995	1.321	0.954
	C1_12m	Min-Max	0.691	0.505	0.994	1.435	1.439
E.ON 2	C2_6m	Robust	0.720	0.524	0.994	1.518	1.763
	C2_12m	Min-Max	0.724	0.530	0.994	1.461	1.665

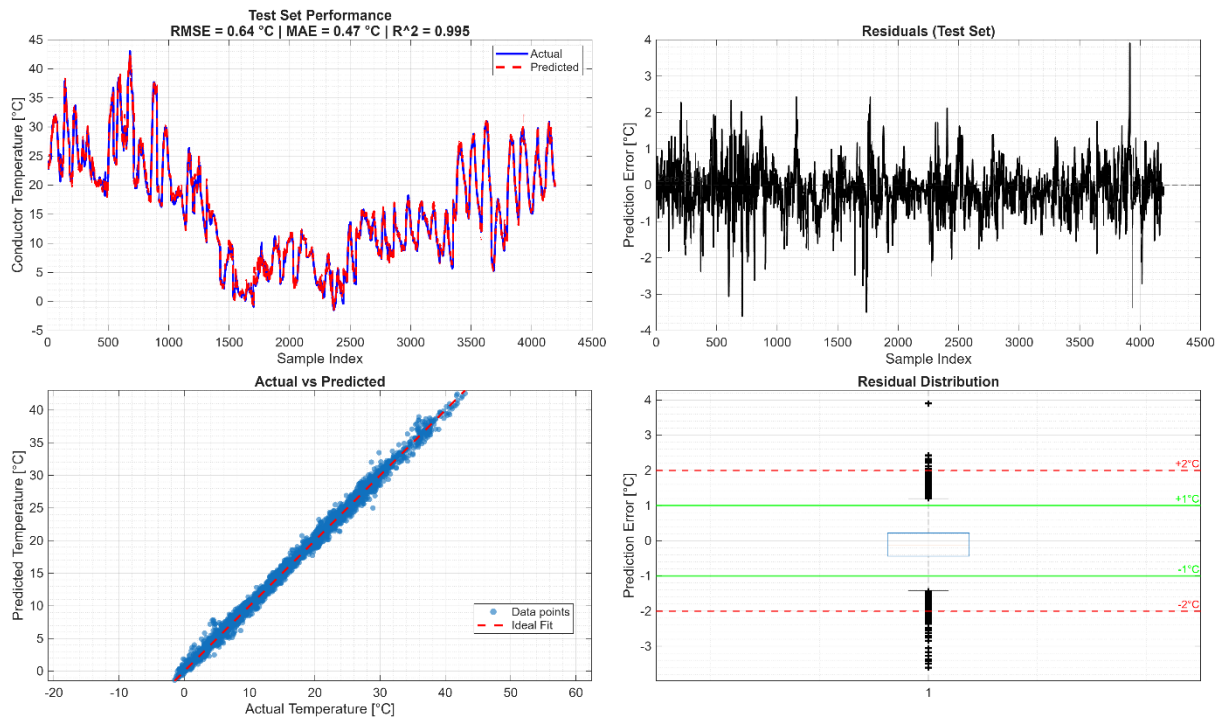


Figure 30 Plots for model C1_6m with weekday split

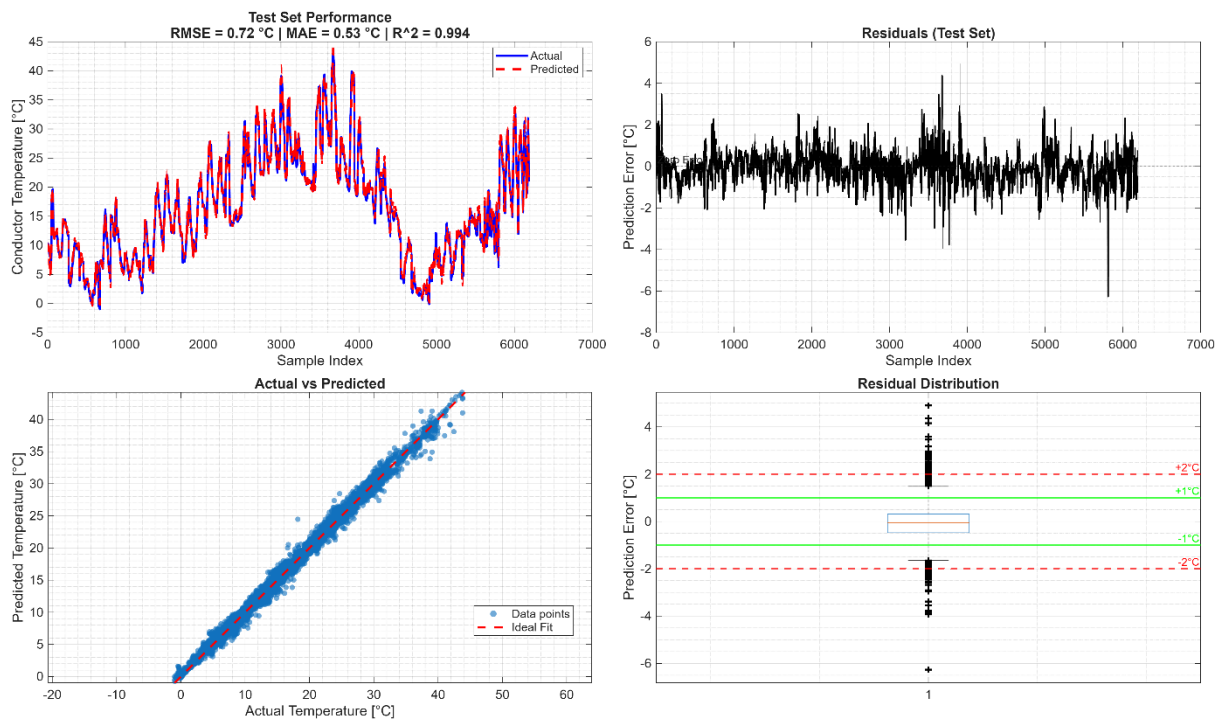


Figure 31 Plots for model C2_12m with weekday split

5.3.4 Discussion of results for the E.ON grid

The results obtained for E.ON are even more promising than those achieved for MAVIR. Overall, the models developed for the E.ON grid demonstrate the best performance among all evaluated grids. The E.ON models outperform both the MAVIR and APG models across all investigated split methods. This behaviour was anticipated during the early stages of development, although it is noteworthy given that the E.ON dataset exhibited the lowest availability rate among the three grids. The better performance is most likely attributable to the higher quality of the available data. The missing data in the E.ON datasets were primarily due to temporary communication and data acquisition issues affecting the line sensors, weather stations, and SCADA systems. In some cases, measurements were unavailable due to communication failures, and inconsistencies in timestamps across data sources also led to records being excluded during data preparation. In addition, frozen sensor values and physically unrealistic measurements were identified and removed where they could not be considered representative of the actual operating conditions. Importantly, temporary communication failures did not necessarily result in permanent data loss, as historical measurements stored in local sensor memory could often be recovered and subsequently integrated into the dataset.

The higher data quality is also reflected in the relatively small differences between the data-splitting methods investigated. The tendency of the weekday split method to achieve the best mathematical performance is also evident in the E.ON case; however, the differences in accuracy between the split methods are considerably smaller than those observed with APG and MAVIR. It is also worth noting that, while the time-series split was clearly the worst-performing method for both APG and MAVIR, in the case of E.ON, it achieved results comparable to those of the weekday and seasonal split methods. Among the three methods, the seasonal split appears to yield slightly lower accuracy. Nevertheless, although the weekday split achieved the best overall numerical performance, the seasonal split remains particularly important for model generalizability. Since conductor temperature prediction is strongly influenced by seasonal meteorological conditions, the seasonal split provides a more representative evaluation of the models' ability to generalize across unseen weather regimes and different parts of the year. Therefore, while the weekday split achieves better performance mathematically, with an RMSE of 0.691 °C and an MAE of 0.505 °C for model C1_12m, with an RMSE of 0.724 °C and an MAE of 0.530 °C for C2_12m, the final method will be the seasonal split.

The effect of training duration differs between the E.ON dataset and APG and MAVIR. In the case of E.ON, extending the training period does not necessarily improve accuracy; in some configurations, it even results in a slight decrease in performance. However, examination of the reliability metrics indicates that the longer training period, despite slightly reducing accuracy, improves model reliability in certain cases.

As the seasonal split was selected as the final splitting strategy from a grid reliability perspective, the corresponding models were chosen as the final models for the E.ON grid. The results indicate that the models achieve very high accuracy at both measurement points. For measurement point C1, model C1_12m achieves an RMSE of 0.826 °C and an MAE of 0.551 °C. Similarly, model C2_12m achieves an RMSE of 0.791 °C and an MAE of 0.556 °C. These results indicate that the prediction accuracy is approximately 0.5-0.7 °C, which is comparable to the measurement uncertainty of the conductor temperature sensors themselves. Consequently, achieving substantially higher accuracy is challenging, as the model performance approaches the intrinsic accuracy limits of the measurements.

The models' reliability is also exceptionally high. For model C1_12m, the outlier percentage is 3.076%, and the P95 value is 1.650 °C. For model C2_12m, the corresponding metrics are 2.524 % and 1.538 °C, respectively. These results demonstrate that the models achieve a very high level of reliability. Model C2_12m is also a prime example of the phenomenon in which a longer training period improves both reliability and accuracy: model C2_6m achieves marginally lower accuracy than model C2_12m, with lower reliability.

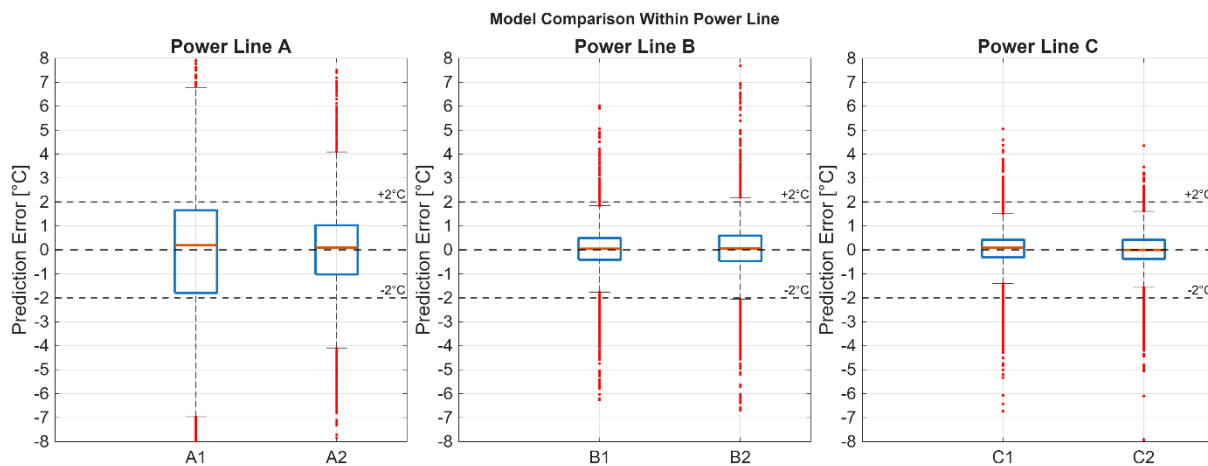
5.4 Discussion of the results

Overall, the results across the three investigated transmission grids demonstrate that the quality and availability of local measurement data are key to the accuracy and reliability of conductor temperature prediction models. Although the involved grids benefited from the applied neural network optimization and model development process, substantial differences can be observed between the APG, MAVIR, and E.ON datasets. The differences in the results of APG, MAVIR, and E.ON clearly indicate that improved data quality and locally measured meteorological information significantly enhance model performance.

As stated previously, the seasonal split method was selected as the final splitting strategy. This decision was made because the seasonal split provides the most generalized and representative evaluation framework for the investigated models. In the seasonal split, both the training and test datasets contain samples from all seasons of the year, ensuring that the models are evaluated across a wide range of meteorological conditions and operating regimes. This is particularly important for conductor temperature prediction, where environmental conditions strongly influence the target variable. Although the weekday split method generally achieved slightly better mathematical performance, it introduces limitations in terms of generalizability. While the weekday split also preserves seasonal representativeness, the test dataset consists exclusively of Sundays from each week. Consequently, the test data may not fully reflect the variability of the full weekly operational behaviour. Even though the models do not explicitly use date or time information as input features, hidden temporal trends may still exist within the data, especially in the load measurements. Because electrical load profiles often differ between weekdays and weekends, restricting the test set to Sundays may unintentionally simplify the prediction task and yield overly optimistic evaluation results. In contrast, the seasonal split uses randomized test days within each season; therefore, the selected testing samples are not associated with any specific recurring weekly pattern or operational characteristic. As a result, the seasonal split provides a more unbiased evaluation of the models' ability to generalize across varying environmental and operational conditions. For these reasons, the seasonal split was considered more appropriate for the final evaluation, as it better reflects the practical deployment scenario in which models must generalize reliably across varying weather conditions, seasonal effects, and operational states throughout the year. Consequently, the final selected models were the six trained on 12 months of data using the seasonal split method. The summarized performance metrics of these final models are presented in the following Figures and Table.

Table 24 Metrics for the final models

Model number	Architecture	Normalization	RMSE [°C]	MAE [°C]	R2 [-]	P95 [°C]	Over ± 2 °C [%]
A1	FFN	Robust	2.927	2.190	0.907	5.579	43.421
A2	FFN	Min-Max	1.899	1.385	0.955	3.828	23.755
B1	FFN	Min-Max	0.962	0.662	0.989	1.946	4.622
B2	CFN	Robust	1.087	0.755	0.984	2.161	6.200
C1	CFN	Min-Max	0.826	0.551	0.991	1.650	3.076
C2	CFN	Robust	0.791	0.556	0.992	1.538	2.524

*Figure 32 Comparison of boxplots between the three grids*

Among the three investigated grids, the APG models achieved the weakest overall performance. This outcome was expected based on the preliminary results and the absence of local weather measurements in the APG dataset. During the early stages of the project, it became evident that conductor temperature prediction is highly sensitive to local environmental conditions; therefore, the lack of directly measured weather data greatly limits the achievable accuracy and reliability of the predictions. Nevertheless, the optimization process and the application of more advanced neural network configurations improved the results considerably compared to the initial baseline models.

Within the APG dataset, the models for measurement point A2 consistently outperformed those for A1 across all evaluated configurations, most likely due to differences in data quality. From a generalizability standpoint, the seasonal split was considered the most important validation strategy, as it assesses the models' ability to generalize across different seasons and meteorological regimes. In conductor temperature prediction, where weather conditions strongly influence the target variable, this aspect is particularly important. The best-performing seasonal model, A2_12m, achieved moderate prediction accuracy with an RMSE of 1.899 °C and an MAE of 1.385 °C; however, the reliability metrics indicated considerable uncertainty, with a P95 of 3.828 °C and nearly 25% of predictions falling outside the ± 2 °C target range. In comparison, the A1 models produced substantially

weaker results, with higher prediction errors and outlier percentages approaching 45%, indicating limited reliability in practical applications.

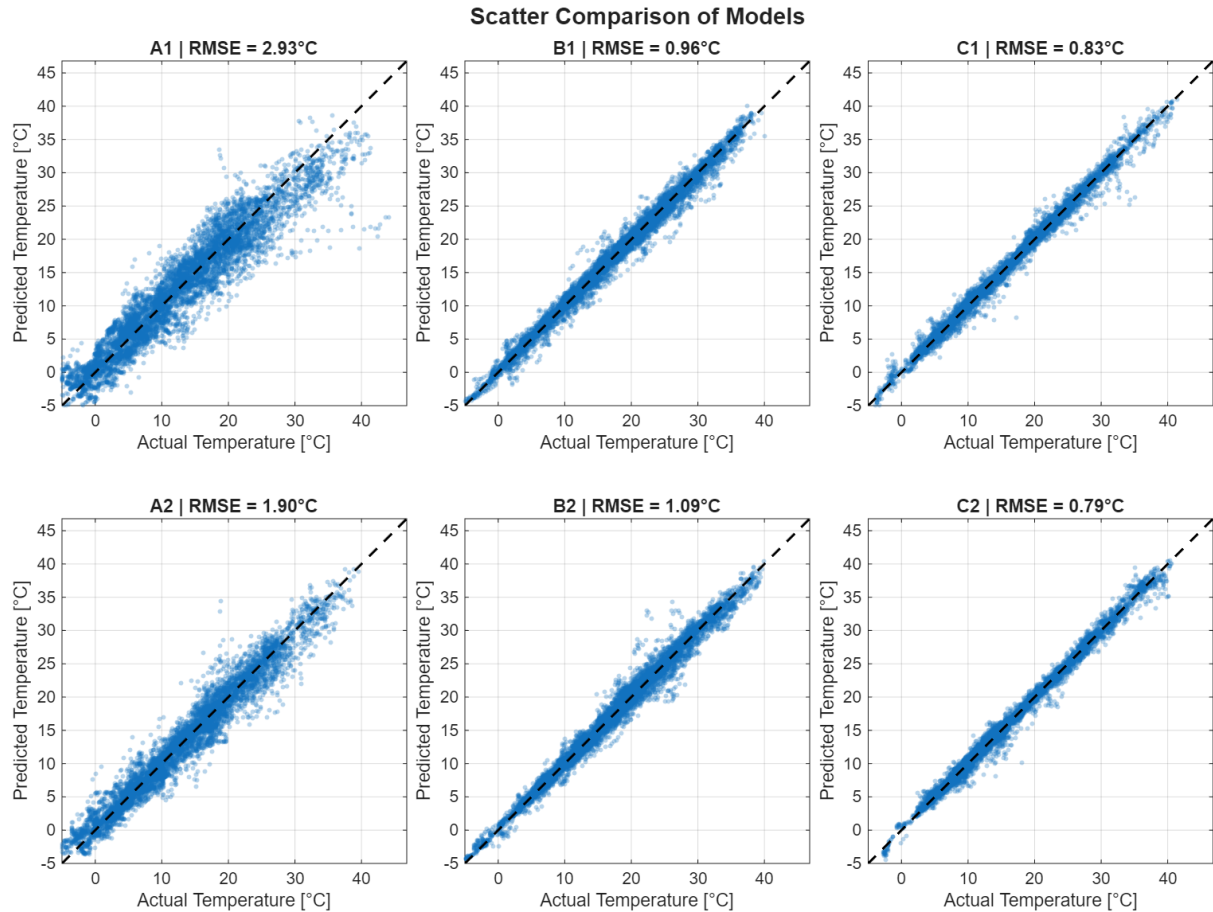


Figure 33 Scatter plots for the final models

The MAVIR models showed a significant improvement over APG. These results strongly support the assumption that local weather measurements have a greater influence on model performance than temporal resolution alone. In the MAVIR case, local meteorological data were available at 10-minute resolution, enabling significantly more accurate predictions. Unlike APG, the two MAVIR measurement points, B1 and B2, produced highly similar results, suggesting a more consistent data quality across the dataset.

For MAVIR, the weekday split method provided the best overall performance, while the seasonal split produced comparable results, and the time-series split showed lower accuracy. The best-performing MAVIR models achieved RMSE values of 0.9 °C and MAE values of 0.7 °C, representing much higher accuracies than those of the APG models. The reliability metrics are also very promising: the outlier percentages were 4.622% and 6.200% for MAVIR, compared to 25-45% for APG. The P95 values, approximately 1.9-2.1 °C, further confirm the significantly better stability and reliability of the predictions. In addition, the results consistently showed that extending the training duration improved model performance in nearly all MAVIR configurations, indicating that the larger training dataset enabled the neural networks to generalize more effectively.

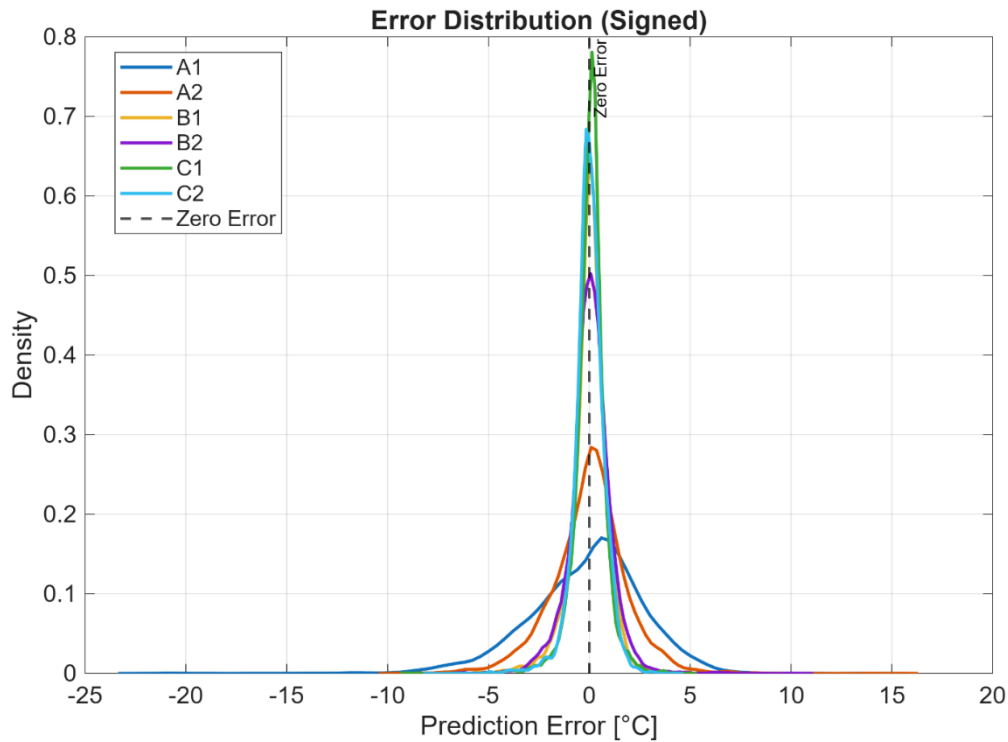


Figure 34 Signed error distribution of the final models

The best overall performance was achieved for the E.ON grid. Despite the E.ON dataset exhibiting the lowest availability rate among the investigated grids, the developed models surpassed both the APG and MAVIR models in terms of both accuracy and reliability. This result highlights that data quality is ultimately more important than dataset size or availability alone. The E.ON models achieved prediction accuracy approaching the uncertainty limits of the conductor temperature sensors themselves. As observed for APG and MAVIR, the weekday split method again provided the best overall results for E.ON, although the differences between the split methods were considerably smaller. Interestingly, unlike the previous grids, the time-series split achieved performance comparable to the weekday and seasonal approaches, while the seasonal split produced slightly weaker results overall. As in the case of APG and MAVIR before, the seasonal split was selected for the final models in case of E.ON as well. The final selected E.ON models achieved exceptionally high accuracy, with RMSE values of 0.826 °C and 0.791 °C, and MAE values close to 0.55 °C for measurement points C1 and C2, respectively. These values indicate that the average prediction error is approximately equal to the uncertainty of the physical temperature measurements themselves, suggesting that higher predictive accuracy may not be realistically achievable using the available data.

The reliability metrics for E.ON further emphasize the quality of the developed models. Outlier percentages remained below 5% for both measurement points, with them being 3.076 and 2.524%. The P95 values were 1.650 and 1.538°C for measurement points C1 and C2. The outlier percentage being approximately 3% means that the predictions have an error of less than $\pm 2^{\circ}\text{C}$ with approximately 97% confidence. These results represent a substantial improvement over both MAVIR and APG and demonstrate that the E.ON models provide highly stable and dependable predictions. An additional noteworthy observation is the influence of training duration in the E.ON case. Unlike APG and MAVIR, extending the training period did not always improve prediction accuracy and occasionally produced

a slight decrease in RMSE and MAE performance. However, the reliability metrics often improved with longer training periods, indicating that the additional data enhanced model robustness even when average prediction accuracy remained nearly unchanged.

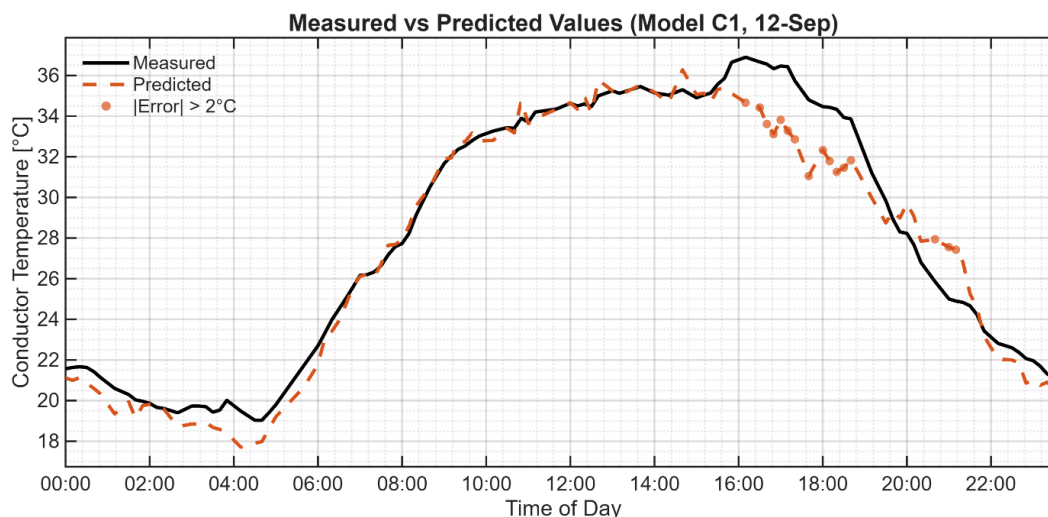


Figure 35 Time-series plot for model C1 during one whole day

In summary, the work conducted demonstrates that accurate conductor temperature prediction using neural networks is strongly dependent on the quality and locality of the available input data. The APG results reveal the limitations caused by the absence of local meteorological measurements, while the MAVIR and especially the E.ON results confirm that high-quality local measurements enable highly accurate and reliable predictions. Across all the grids investigated, the weekday split method proved to be the most accurate mathematically, but the seasonal method was chosen for its representative validation strategy. Both seasonal and weekday methods greatly outperformed the time-series method, which confirms that in conductor temperature prediction, where weather data is important, the application of a split method representative of the seasons is essential. Furthermore, the results confirm that increasing training duration generally improves model performance when sufficient data quality is available, although in very high-quality datasets such as the E.ON dataset, improvements may primarily manifest in enhanced reliability rather than further reductions in average errors.

5.5 DLR calculation algorithm

The transmission capacity of overhead lines is conventionally determined using worst-case environmental scenarios to avoid conductor overheating under all operating conditions. By calculating the heating effect of the current (Joule-heating) and solar radiation, and the cooling effect of radiation and moving air (natural and forced convection), the heat balance equation of the conductors results in the maximum current flow of the phase conductors that leads to the maximum operating temperature (MOT) during the environmental conditions taken into account. This so-called static line rating (SLR) is determined based on the worst-case environmental factors, e.g., 35 °C ambient temperature, 0.61 m/s perpendicular wind, and 1000 W/m² solar radiation. The concept is that, even on a hot, sunny day, the conductors can be safely operated within these current flow limitations. The

SLR can also be calculated for summer and winter periods, yielding seasonal SLR values that provide power flow constraints for the summer and winter half-year periods. Based on the calculation philosophy, it can be seen that both the SLR and seasonal SLR are wasteful capacity allocation methodologies, as most of the time the environmental conditions are more favorable from the conductor cooling effect point of view; accordingly, there is a thermal reserve that changes over time according to the changing of the environmental factors that can be utilized over time.

To exploit this thermal reserve of the conductors, the dynamic line rating (DLR) technology has been developed. The main scope of the application of the DLR technology on power lines is twofold:

- During environmental conditions which result in a higher cooling effect on the conductor than the parameters used for the SLR calculation, the transmission capacity can be increased,
- While the environmental conditions result in a lower cooling effect on the conductor than the parameters used for the SLR calculation, the operational safety can be increased by limiting the transmission capacity, and this way avoiding the thermal overload of the conductors.

The former case is expected to occur in 95-99% of cases, while the latter can reach up to 5% in operational practice, as shown in Figure 36.

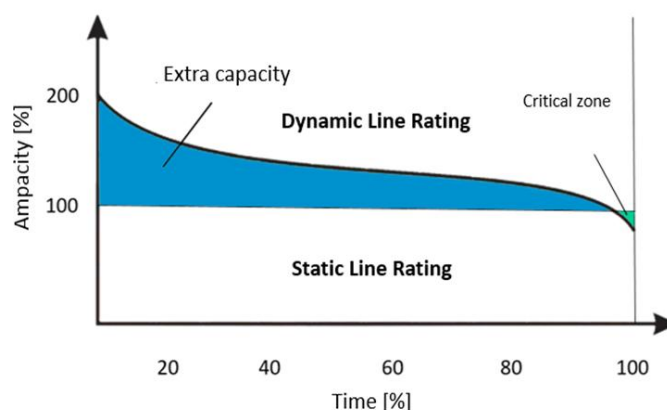


Figure 36 Capacity gain can be achieved by the application of DLR technology

CIGRE and IEEE are two major international organizations that have developed their own models for DLR calculation [16], [21]. The calculation methodology for these models is the same as during the SLR calculation; however, the measurement of environmental parameters is quasi-real-time (5-15-minute frequency). The advantage of the models is that real-time capacity allocation can be realized. However, it is also necessary to note that, since these are indirect models – they calculate based on environmental parameters – their accuracy is questionable. Therefore, a new model is developed by the DLR reach group of BME and F4STER, a direct model that calculates based on measured conductor temperature [22].

The phase conductor can be modelled as a single heat-storage system; therefore, its transient thermal equation is given by (9).

$$\frac{dQ}{dt} = m \cdot C_p \cdot \frac{dT}{dt} \quad (9)$$

where Q is the heat energy per unit length of the conductor [W/m^3], m is the specific mass of the conductor [kg/m], C_p is the heat capacity of the conductor [$\text{J}/\text{kg}\cdot\text{K}$], dT is the temperature change of the conductor [$^{\circ}\text{C}$] and dt is the measurement frequency [s].

By accounting for the fact that the dQ/dt thermal energy change in the conductor is generated by fluctuations in the load between two measured temperature values, the thermal energy balance equation of the conductor can be written as (10).

$$\frac{dQ}{dt} = I_{\text{surplus}}^2 \cdot R_{AC}(T) \quad (10)$$

where I_{surplus} is the surplus current that is permissible beside the actual load of the conductor, not to exceed its maximal continuous operational temperature [A], and $R_{AC}(T)$ is the temperature-dependent AC resistance of the conductor [Ω/m].

With the rearrangement of (9) and (10), the transmission capacity of the conductors can be expressed as (11).

$$I = I_{\text{load}} + \sqrt{\frac{m \cdot C_p \cdot dT}{R_{AC}(T) \cdot dt}} \quad (11)$$

where I_{load} is the actual load of the conductor [A]. In the case of bundled conductors, the result of (8) should be multiplied by the number of bundles to obtain the line rating of a phase.

The main advantage of this DLR calculation methodology is that it is based on direct temperature measurement; accordingly, the thermal inertia of the conductors is included in the calculation, which results in less fluctuating ampacity constraint, while the possibility of thermal overloads is minimized. The ampacity determined based on different models is shown in Figure 37.

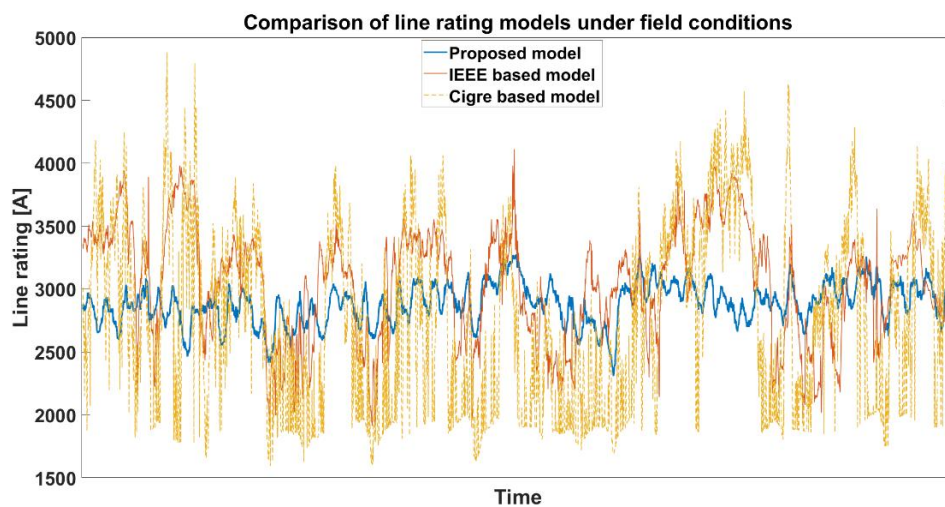


Figure 37 DLR calculation results with different models

The developed model also fits into the scope of the TwinEU project. As the ampacity calculation in the proposed model uses real-time conductor temperature values, it cannot be provided solely from DLR sensor measurements, but also from the virtual sensors developed within this project.

5.5.1 DLR values calculated for the various grids

After presenting the theoretical background, the following sections demonstrate the benefits of using DLR technology. For presentation purposes, one year of data was taken from all three grids to demonstrate the benefits achievable with the use of DLR technology. The benefits will be examined from the angle of the average capacity gained compared to the SLR value. The average capacity gain is presented as the ratio of the available DLR surplus to the corresponding SLR value. Consequently, the results are expressed as percentage increases in transmission capacity relative to the conventional static line rating. This representation allows straightforward comparison across different periods, operating conditions, and transmission lines, while providing a clear measure of the practical benefits achievable through DLR implementation.

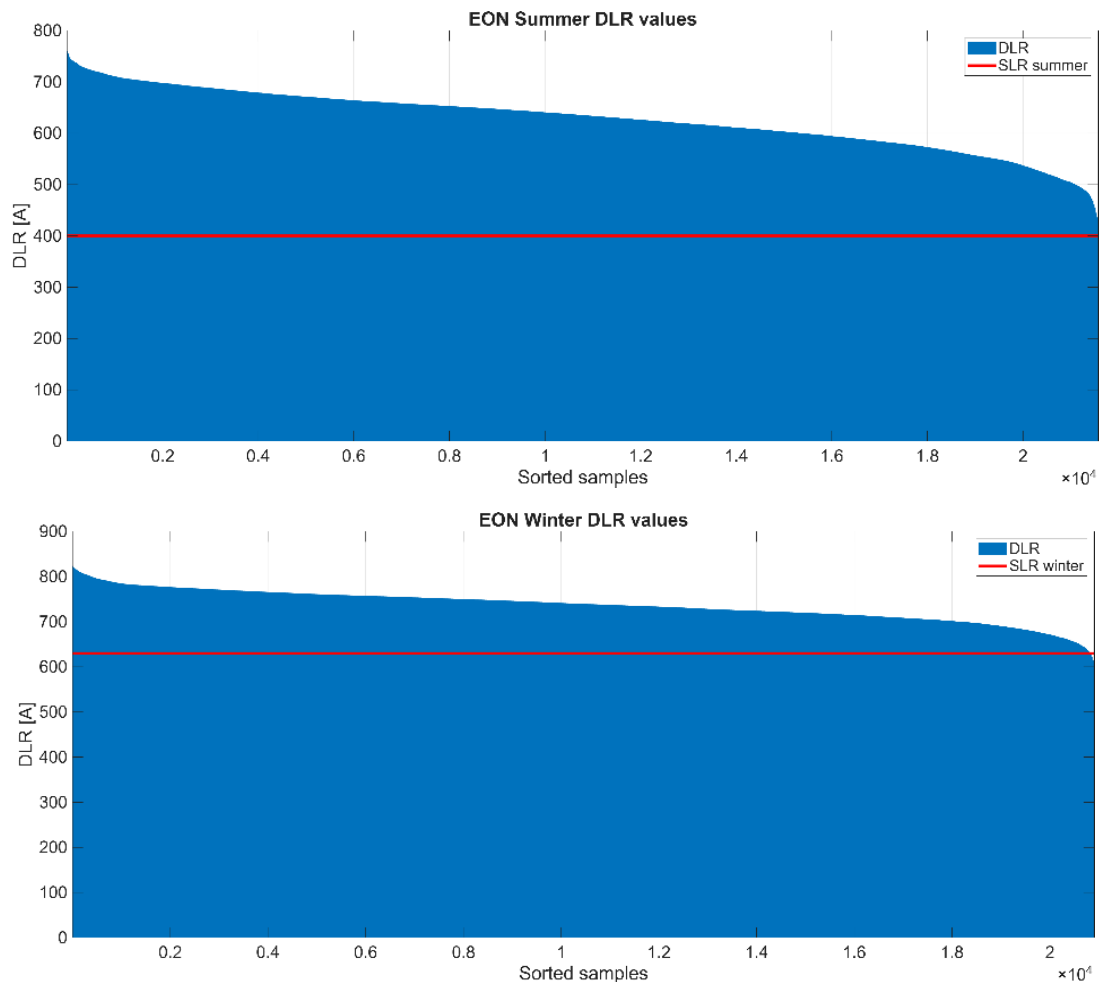


Figure 38 Capacity gain for one year on the E.ON line (summer and winter)

Both MAVIR and E.ON use the seasonal line rating principle, meaning the SLR value differs between summer and winter. Therefore, the capacity calculation has to address that difference. In the case of APG, only one value was available; the calculation is simpler.

To demonstrate the benefits, one full year of data was collected from the line sensors for all three grids. Then, the DLR values were calculated according to the equations presented in section 5.5.

For the E.ON grid, the benefits are substantial, as the DLR achieves an average capacity gain of 57% in the summer and 17% in the winter period. In this case, the DLR values are larger than the SLR values in more than 99.5% of the year. The benefits of the capacity of the line can also be seen in Figure 38, which was made according to Figure 36.

For MAVIR, the capacity gained is lower, but the benefits of dynamic line rating is clear. The MAVIR line gains averagely 13% more capacity in the summer, and 5% in the winter. In summer, the DLR value exceeds the SLR value 97% of the time, while in winter the ratio is 90%. This also presents another benefit of the DLR technology, with the increase of operational safety.

APG's case is special because of the lack of summer/winter split in the SLR ratings. Therefore, the average capacity gain is calculated for the whole year. In the case of APG, the gain is about 28% for a year. Interestingly, in the case of the selected power line, the DLR values are exceeding the SLR value throughout the whole year, which results in a ratio of a 100%.

The calculations can also be made with the virtual sensors. Below is an example of the DLR calculated from the measured and predicted values for a selected day for the E.ON grid.

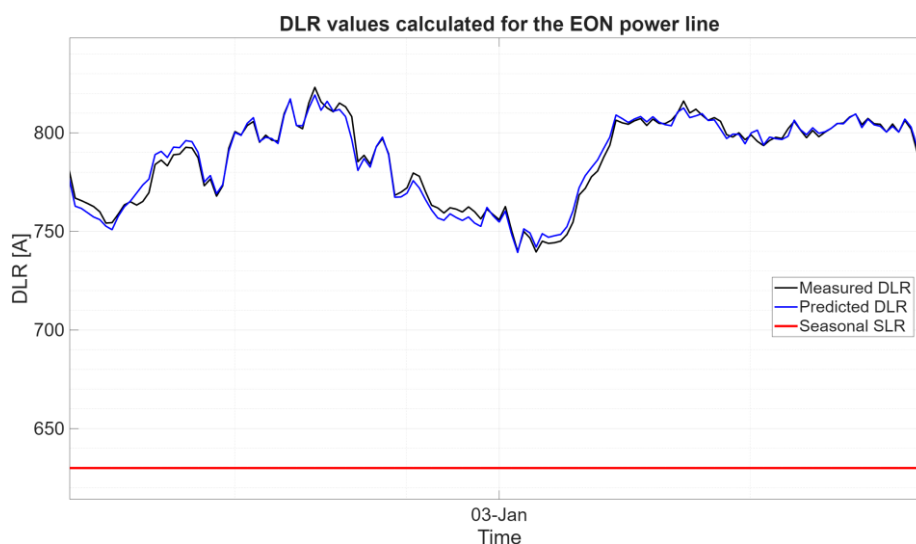


Figure 39 DLR values calculated from the real and virtual sensors

It can be observed that the DLR values calculated using the virtual sensor estimates closely match those obtained from the measured conductor temperatures. This result is consistent with the previously presented conductor temperature prediction accuracies. Since the DLR calculation is based directly on the conductor temperature, accurate temperature predictions naturally lead to accurate ampacity estimates. The close agreement between the DLR values derived from real and virtual sensors demonstrates that the developed models are capable of reproducing the thermal behaviour of the conductor with sufficient accuracy for practical DLR applications. Consequently, the results

confirm that virtual sensors can provide a reliable alternative to physical conductor temperature measurements for dynamic line rating calculations.

5.6 Connection to Task 7.3 in Work Package 7

5.6.1 Concept of the joint work

The DLR sensor digital twin provides the forecasting layer that connects power line monitoring with the Hungarian pilot's flow-based capacity calculation and market co-optimization workflow, which is developed in T7.3. The objective is to estimate short-term dynamic line ratings from weather and loading forecasts and to use these ratings as updated transmission constraints in the flow-based capacity calculation (FBCC) process. In this way, the digital twin does not replace the existing capacity calculation methodology but enhances it by providing more realistic and time-dependent ampacity values for selected critical transmission lines.

The concept builds on the observation that conventional static line ratings are typically conservative because they are calculated under fixed assumptions about ambient conditions and conductor cooling. In practice, the actual thermal capability of an overhead line depends on meteorological and operational variables, including ambient temperature, wind speed, wind direction, solar radiation, loading current and the resulting conductor temperature. The ANN-based digital twin learns the relationship between these inputs and the conductor's thermal state from historical sensor measurements. After training and validation, the model can forecast dynamic ampacity for future market timeframes using weather and loading forecasts.

The primary output of the digital twin is a time series of ampacity improvement factors for the relevant monitored lines and time intervals. These outputs are then transferred to the capacity calculation module, where they are used to adjust the maximum admissible flow limits. In practical terms, the DLR forecast modifies the available transmission margin by scaling the static rating-based limit with a dynamic rating ratio. The resulting DLR-enhanced Remaining Available Margin (RAM) is then used together with the corresponding Power Transfer Distribution Factor (PTDF) values in the flow-based domain.

This approach is especially relevant for the Hungarian pilot because the market co-optimization framework jointly considers energy and balancing capacity products across day-ahead and intraday timeframes. Four sequential capacity calculation runs are foreseen before the respective market gate closures: one day-ahead capacity calculation and three intraday capacity calculations. The sensor digital twin supports this sequential structure by providing updated ampacity forecasts as delivery approaches. As weather and loading forecasts become more accurate closer to real time, the DLR-enhanced capacity calculation can better reflect the actual transmission capability at each auction stage.

The resulting DLR-enhanced flow-based domains are used as network constraints in the market co-optimization module. The market module clears energy and balancing capacity bids while respecting the PTDF-RAM constraints produced by the capacity calculation layer. Compared to a static-line rating reference case, the ANN-based DLR extension can expand the feasible transmission boundary under favorable weather and loading conditions. This additional margin can then be allocated by the co-optimization algorithm to the market product where it has the highest economic value, whether day-

ahead energy exchange, intraday energy exchange, or balancing capacity procurement. The combined framework, therefore, separates two effects: DLR expands the technically feasible network capacity, while co-optimization determines how this capacity is economically allocated between competing market products.

5.6.2 Data forwarded to Task 7.3

In support of Task 7.3, this Demo input data was provided to us to examine the effect of DLR's increased line rating on the co-optimization algorithm's output and market processes. Given that the above processes are not important in real time, but in forecasting, we do not use real-time data in this part of the project. The presented DLR model is also suitable for line rating forecasting if the load and weather forecasts are provided as inputs. The difficulty of the latter is that in each case, a weather agency service is required for the following parameters: ambient temperature, wind speed, wind direction, solar radiation, and, if available, relative humidity or precipitation. The forecasts used in the project are updated with a 24-hour (120-hour in the case of APG) look-ahead in 3-hour time windows. This also means that forecasts are received from several previous times for a given moment in time, i.e., the line rating value calculated based on them can be tracked in time. This also allows market auctions that close at different times to be run.

During the data provision, the data can be divided into two parts.

There are static data that do not change over time, including the following power line data:

- Country/System operator;
- Power line name;
- Power line ID (same as in the JAO system);
- Nominal voltage [kV];
- Number of circuits [-];
- Static Line Rating (summer) [A];
- Static Line Rating (winter) [A];
- Substation limit [A] (if available).

There are dynamic data as well, that do change over time, including the following power line data:

- For the simulation date: Timestamp [YYYY.MM.DD hh:mm:ss];
- For the forecast date:
 - Timestamp [YYYY.MM.DD hh:mm:ss];
 - Load forecast (if available) [A],
 - Dynamic Line Rating (only for the conductor) [A];
 - Dynamic Line Rating (for the power line including substations) [A];

The forecasted DLR data is presented in the graphs below. For demonstrational purposes, only 3 lead time groups were selected for plotting. The values are presented for a summer period, and for a week in winter. The plots also contain the power line a season specific SLR value.

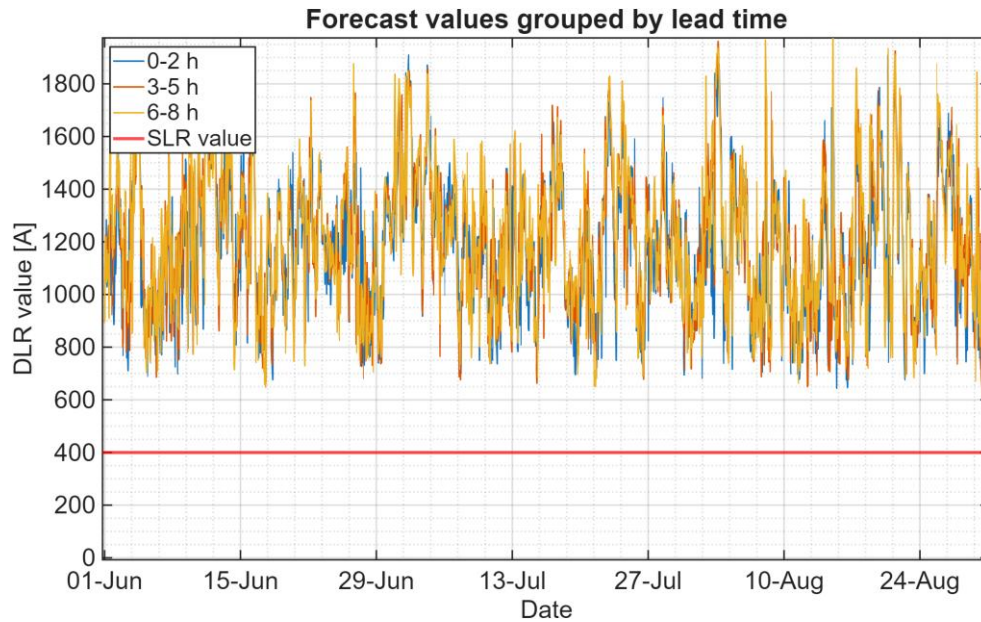


Figure 40 Forecasted DLR values for E.ON grid in summer

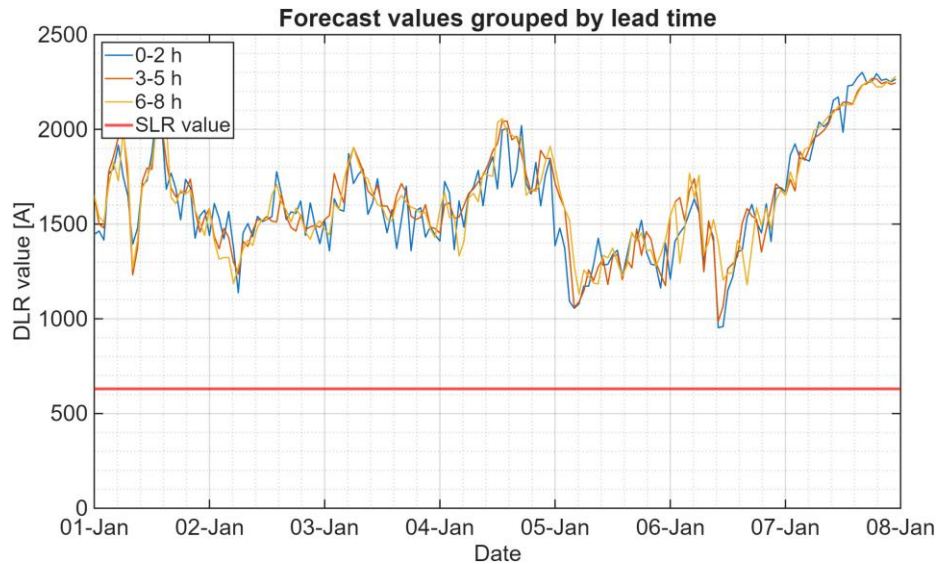


Figure 41 Forecasted DLR values for a week in winter for the E.ON grid

The advantages of DLR forecasting technology can be clearly observed in the presented figures, where the forecasted DLR values remain systematically higher than the corresponding SLR values for most of the investigated period. This additional available capacity represents transmission potential that would remain unused when applying conventional static ratings. By providing a more accurate representation of the actual operating conditions of the transmission line, DLR forecasting enables operators to utilize this surplus capacity while maintaining the required safety margins. The additional

transmission capacity can subsequently be incorporated into market simulations, congestion management, and capacity allocation processes, thereby improving network utilization.

A detailed presentation of the analyses carried out in the work of Task 7.3 is available in the document Deliverable 7.2 TwinEU digital twinning in the balancing context.

5.7 Key Performance Indicators

Subsection 4.3.4 presents the KPI metrics defined at the beginning of the project, their description and target value. Based on the results presented in the Deliverable, Table 24 shows the values achieved by these metrics during the project.

Table 25 Final KPIs as a result of the project

ID	Name	Description	Target value	Reached value	Fulfilment
1	No. of physical sensors	How many physical sensors were used for the data collection of the ANN-based DT training?	Over 3	6	Yes
2	No. of DT models	How many ANN-based DT models have been implemented?	Over 3	6	Yes
3	No. of accurate DT models	How many ANN-based DT models have been made where the average achieved accuracy is within plus or minus 2 degrees Celsius?	Over 1	3	Yes
4	Virtual sensor accuracy (max & mean)	What is the maximum and average DT accuracy achievable (one-month exam period)?	Avg: $\pm 1^{\circ}\text{C}$ Max: $\pm 2^{\circ}\text{C}$	See Table 26	Yes
5	Virtual sensor accuracy (outliers)	What is the outlier percentage of the DT model (one-month exam period)?	Below 5%	See Table 26	Yes
6	No. of power lines covered with virtual sensor	How many transmission lines are covered with virtual sensors (DT)?	Over 1	2	Yes
7	No. of potential substitutions	How many physical sensors can be potentially substituted by ANN-based DT models?	Over 1	3	Yes

Based on Table 24, each of the 7 predefined metrics reached or exceeded their preliminary target values, indicating that the DT development and evaluation phase was also successful from a technical perspective.

DT models are considered fully viable solutions for the project if they meet the virtual sensor accuracy metrics defined by KPI IDs 4 and 5. The project's objective was to secure at least one such DT model.

Ultimately, the project produced a total of three such ANN-based virtual sensors, as follows:

- B1_12m;
- C1_12m;
- C2_12m.

For one model, the result is technically acceptable, but based on the preliminary KPI definition, it was classified as not meeting the requirements: B2_12m.

Table 26 Metrics for KPI No. 3 and 4

Grid	Model name	Box range (Average accuracy) [°C]	Whisker range (Maximum accuracy) [°C]	Outlier percentage [%]	KPI No. 4 fulfillment	KPI No. 5 fulfillment
APG	A1_12m	-1.65 – 1.79	-6.82 – 6.96	43.42	No	No
	A2_12m	-1.02 – 1.02	-4.10 – 4.09	23.76	No	No
MAVIR	B1_12m	-0.49 – 0.41	-1.85 – 1.76	4.62	Yes	Yes
	B2_12m	-0.59 – 0.46	-2.19 – 2.06	6.20	No	No
E.ON	C1_12m	-0.42 – 0.31	-1.52 – 1.40	3.08	Yes	Yes
	C2_12m	-0.42 – 0.37	-1.61 – 1.57	2.52	Yes	Yes

6 Lessons learned

Based on the project outcomes, the lessons learned can be collected.

Scaling up is key for line monitoring projects. - The initial assumptions and objectives were valid: there is a genuine and growing demand among system operators for scaling up power line monitoring and Dynamic Line Rating (DLR) solutions. However, the project also demonstrated that challenges differ significantly across voltage levels, even within the same geographical region. On high-voltage transmission networks (220–400 kV), the primary focus is on extending cross-border power flows and improving congestion management. In contrast, on 132 kV networks, the increasing penetration of RES has introduced complex, less transparent power flow patterns that transmission system operators often cannot fully observe from the control room.

This creates a strong need for vertical cooperation between system operators and highlights the importance of closer data exchange, communication, and interoperable IT support systems. The project successfully laid important foundations for such cooperation, even if not every planned element could be fully realized.

An effective DT solution must be multi-layered. - One of the key conclusions is that the development of the Digital Twin (DT) concept proved to be highly promising. The first critical layer is the optimal selection of sensor locations, although this aspect received less emphasis in the project due to constraints imposed by the existing sensor infrastructure. The second layer concerns sensor deployment itself, where the project demonstrated that even installing a limited number of devices requires substantial coordination, scheduling, human resources, and financial investment. Without proper planning, deployment costs can increase significantly.

Data management is a critical layer in an ANN-based DT. - The third and most critical layer is data acquisition and management. The project confirmed that there is no truly maintenance-free monitoring solution. Once the number of sensors exceeds approximately 10, coordinated maintenance strategies and data outage preparation become essential. Redundancy through multiple devices and virtual sensor solutions is therefore indispensable. Since AI-supported systems are fundamentally dependent on both the quantity and quality of data, preprocessing, filtering, anomaly detection, and data cleaning are crucial steps. Sensor faults, frozen values, inconsistent resolutions, missing data, and time zone handling – especially for cross-border lines – must all be addressed systematically. The project also showed that achieving 12 months of high-quality training data may require collecting at least 14 months of measurements in practice.

Data Quality and Data Preparation Lessons. - The demonstration confirmed that data quality is often more important than nominal data availability when developing ANN-based virtual sensors. Although the E.ON datasets had lower overall availability than those from the other demonstration sites, the resulting ANN models achieved the highest prediction accuracy. This finding indicates that consistent, representative, and physically meaningful data contribute more to model performance than the absolute number of measurements collected.

Missing measurements were not artificially reconstructed or interpolated. Instead, the available time-series data were used in chronological order for ANN training to preserve realistic operating conditions and avoid introducing synthetic patterns into the training dataset.

Data preparation also required the identification and removal of frozen sensor values, physically unrealistic measurements, and records with inconsistent timestamps between conductor sensors, weather stations, and SCADA systems. In cases where data quality issues originated from temporary communication failures, historical measurements stored in local sensor memory could often be recovered afterward and integrated into the dataset. This highlighted the importance of local data buffering capabilities for field devices used in Digital Twin applications.

Continuous supervision, backup procedures, and the possibility of recovering historical SCADA information even months after acquisition proved valuable for maintaining the consistency of the final datasets. Overall, the experience confirms that future Digital Twin implementations should place equal emphasis on data governance, quality assurance, and monitoring of data acquisition processes as on ANN model development itself. The demonstration showed that data quality, consistency, and representativeness are ultimately more important for ANN performance than the nominal availability of the collected datasets.

Special attention shall be made to local data sources. - Regarding data sources, SCADA systems provided the most reliable availability. At the same time, the project clearly demonstrated that weather stations are among the most critical components of any DLR or Digital Twin architecture. Local weather measurements consistently outperformed interpolated or downscaled weather agency data, particularly for AI model training, where higher accuracy is required. Even within local measurements, sensor placement and proximity to monitored assets proved highly influential. An important general finding is that while sufficient data is essential, increasing the amount of data beyond a certain threshold does not automatically improve model performance. If the data-splitting strategy effectively captures seasonal behaviour, even six months of properly processed data can yield meaningful results, whereas extending the dataset to 12 months may provide only marginal additional improvement.

Optimization leads to success in ANN development. - The development of the Artificial Neural Network (ANN) also led to several important conclusions. Earlier assumptions suggested that a sufficiently advanced network architecture could provide highly accurate predictions across multiple lines. However, project results showed that each location requires an individually optimized network structure. Furthermore, the mathematically best-performing model is not always the most practical solution from an operational perspective, especially if the objective includes reducing the number of physical sensors over time. The project also confirmed that increasingly complex network architectures do not necessarily yield better performance; instead, an optimal balance between complexity, robustness, and operational usability must be achieved.

In terms of accuracy, the project successfully demonstrated that virtual sensors can achieve temperature estimation accuracy within approximately $\pm 2^{\circ}\text{C}$ under appropriate data conditions. However, accuracy alone is insufficient; the number and magnitude of outliers are equally important. This highlights the necessity for system operators to define acceptable operational thresholds and performance ranges. Importantly, the project succeeded in developing a Digital Twin capable of reaching sufficient reliability to partially replace physical sensors after a certain operational period. Nevertheless, the project strongly recommends maintaining a minimum number of physical sensors on the network to support operational security and continuous retraining and validation of AI models.

Operational Integration of DT-assisted DLR. - The demonstration confirmed that ANN-based virtual sensors and Dynamic Line Rating (DLR) solutions are technically feasible and can provide operationally useful results. However, successful deployment depends not only on model accuracy but also on compliance with regulatory requirements, operational procedures, and safety constraints. A key lesson learned is that DLR outputs can only be utilized within the limits imposed by conductor temperature, sag-clearance requirements, and operational safety margins. Where clearance is the governing constraint, continuous monitoring or reliable estimation of critical spans are essential.

The project demonstrated that ANN-based virtual sensors can provide value beyond simply replacing physical sensors. By increasing the number of observable locations along a transmission corridor, Digital Twin solutions improve network visibility and support a better understanding of conductor thermal behaviour. Virtual sensors can also provide useful information during temporary sensor outages, communication failures, or maintenance activities, thereby increasing monitoring system robustness. The results further showed that DLR can support operational decision-making, outage planning, maintenance activities, and contingency management. Practical implementation nevertheless requires clear validation procedures, operator responsibilities, and well-defined fallback mechanisms. The demonstration also highlighted the importance of data quality, update frequency, and latency. These factors become increasingly critical when network operation approaches dynamic rating limits.

Recommended Criteria for Physical Sensor Substitution. - The project demonstrated that decisions regarding the replacement of physical conductor temperature sensors should be based on validated model performance rather than a predefined training period. The required training and validation duration depends on local climate, seasonal variability, and data quality. The results indicate that virtual sensors should be validated under the full range of expected environmental and loading conditions before operational use. The experience gained during the demonstration also supports the application of a hybrid monitoring concept, where physical and virtual sensors are used together to maintain monitoring reliability and model supervision.

Identifying barriers and enablers for power line monitoring systems. - The project also emphasized that the success of AI-based Digital Twin systems depends not only on technical performance, but also on organizational readiness. The maturity of IT infrastructure and the confidence of control room operators are key prerequisites for large-scale deployment. There are different expectations from the IT side, and different developments are required for each SO. What is common is that it is worth overcoming this as soon as possible, because this is the key issue of scalability. In addition, not only within the company but also across companies, communication and cooperation require appropriate protocols, IT security, and data platforms, which also require a new approach to operations. The results achieved within the project provide a strong basis for building such confidence. Moreover, the developed solutions are not merely academic or experimental outcomes; they respond to clear market needs.

In addition to identifying barriers, the project also highlighted that the data and data collection process for line monitoring systems are not sufficiently supported by guidelines, standards and legislation. Therefore, there is a need to establish a framework that highlights the importance of real-time monitoring and its correct implementation. In the future, more active governance support could fundamentally advance the reliability and accuracy of such systems.

Line rating – the real output value and connector to various applications. - The project moved beyond simple temperature monitoring by developing a proprietary DLR algorithm that calculates line ratings directly from thermal conditions. This functionality is highly valuable for real-time congestion management and market co-optimization applications. The integration with TwinEU WP7 further demonstrated the applicability of these results within broader simulation and forecasting environments. An important advantage of the DT solution is its flexibility: it can operate with both real-time and forecast data inputs.

Another significant achievement is that the proprietary DLR algorithm produces smoother temporal DLR behaviour, making the results easier for control room operators to integrate into operational workflows. Beyond system operation, the Digital Twin concept also opens new possibilities for asset management. Continuously monitored virtual models allow operators to supervise previously unmonitored network sections and simulate critical operating conditions – such as excessive conductor temperature or sag – that would be difficult or unsafe to reproduce physically.

Beyond technological layers - exploitation and usability. - From an economic perspective, the project demonstrated that Digital Twin solutions can reduce both CAPEX and OPEX, depending on the monitoring strategy adopted by the system operator. Maintenance costs can also be significantly reduced. These findings directly support the emergence of new market models, particularly sensor rental services, which offer lower upfront costs, reduced material usage, improved sustainability, and continuously updated hardware platforms. The AI-supported approach enables more efficient lifecycle management with lower maintenance requirements and better technical performance.

Environmental sustainability impact. - The use of dynamic line rating (DLR) system can have a large impact on the sustainability aspirations of the SOs. Due to the extra capacity gained with the DLR system – 20-40% in average – other uprating technologies, e.g. building new power lines or reconductoring with HTLS conductors, can be redeemed, which means a decrease in the high carbon footprint during network development. Moreover, during the implementation of a DLR system, the proposed ANN model can redeem the use of (some of the) physical sensors in medium and long term. Accordingly, the carbon footprint of DLR technology – which already has lower CO₂ emissions compared to the alternative uprating technologies – can be further reduced. Furthermore, DLR technology gives flexibility in the power system due to the changing transmission limits, that also advantageous from the renewable energy integration point of view. As the renewable energy sources feed intermittent energy into the grid, the DLR technology facilitates to minimize dispatching actions by adjusting grid capacity to rapidly changing generation.

7 Future works

With the completion of the project evaluation period, the primary focus of the next phase will be disseminating the project results through scientific publications, technical reports, stakeholder engagement activities, and other communication channels. Beyond dissemination, the knowledge and experience gained during the project also provide a solid foundation for defining future development pathways and opportunities.

Based on the project outcomes, four main directions for future work can be identified. First, the project findings enable the formulation of recommendations for regulatory frameworks and industry standards, reflecting the lessons learned during the development and validation of the Digital Twin (DT). Second, several opportunities exist to further evolve and apply the DT concept across new technical domains and operational scenarios. Third, the future of the monitoring system shall be fitted into the requirements of a modern asset management and communications system by coping with challenges identified in the project. Last but not least, the project has demonstrated significant potential for industrial use of the DT, opening pathways for commercialization and practical deployment in the power sector and related industries.

7.1 Recommendations for regulation

One of the key findings of the project is the critical role of real-time monitoring systems. Continuous collection of local meteorological and transmission line data provides valuable situational awareness and enables more accurate assessment of network conditions. A notable example of this regulatory approach is FERC Order 881 [23], which mandates the partial implementation of dynamic line rating solutions in North America and recognizes the value of real-time operational data for enhancing grid efficiency and reliability. It would be advisable to adopt a similar regulatory framework in the European Union, underscoring the importance of this paradigm shift.

Another important lesson learned is that the successful deployment of next-generation line monitoring systems requires not only advanced sensing technologies but also a high level of digitalization, robust IT support, and access to sufficient quantities of high-quality data. While technological solutions are already available, their effective use is often constrained by the lack of a comprehensive governance framework that would support system operators in implementing and operating these systems consistently. Future regulatory initiatives should therefore aim to establish clear guidelines and standards for line monitoring solutions, covering not only field sensors but also supporting IT infrastructure, data management processes, cybersecurity requirements, and data-centred operational practices.

7.2 Potential future application of the DT

Future DLR-supported operation should incorporate predefined contingency scenarios, including responses to potential N-1-1 events. When transmission lines operate close to their dynamic limits, system operators should have predefined corrective actions available, such as redispatch, generation curtailment, demand limitation, or other security measures.

A practical deployment pathway may consist of four maturity levels:

- Monitoring
- Decision Support
- Operational Use
- Supervised Automation

Digital Twin systems may progressively take a more active role in operational support; however, responsibility for operational decisions should remain with system operators.

In the near term, Digital Twin and DLR technologies should primarily be introduced as decision-support tools, allowing operators to gain confidence in the technology before wider operational integration. Long-term success will depend on the establishment of appropriate operational procedures, governance frameworks, and regulatory acceptance.

The usefulness of the developed Digital Twin has already been validated through the co-optimization process and market simulations, where its line rating outputs contribute directly to operational and market optimization. Beyond these applications, a promising future direction is the utilization of the Digital Twin models for condition monitoring and predictive asset management.

The project demonstrated that Digital Twins equipped with virtual sensors can significantly enhance condition monitoring capabilities compared to conventional physical sensing solutions. One of the most important applications is the assessment of thermal degradation in overhead line conductors. Since conductor ampacity is primarily limited by thermal constraints, accurate estimation of conductor temperature is essential for maintaining both operational reliability and asset integrity. For ACSR conductors, prolonged exposure to temperatures above 90–93 °C may initiate annealing of the aluminum strands, resulting in a gradual reduction in tensile strength and accelerated aging.

A major advantage of the Digital Twin approach is its ability to reconstruct historical thermal exposure using SCADA measurements and environmental data. Unlike physical sensors, which only provide information after installation, Digital Twin models can retrospectively estimate conductor temperatures over extended historical periods. This enables identification of critical overheating events and quantification of their cumulative impact on conductor degradation. By combining temperature histories with laboratory-derived aging models and industry standards, the Digital Twin can estimate remaining mechanical strength and project future asset lifetime.

However, the true value of condition monitoring lies in integrating these insights into existing asset management processes. Digital Twins enable predictive asset management by continuously evaluating asset health and supporting risk-based maintenance decisions. Furthermore, their scalability allows utilities to monitor large portions of the network without extensive sensor deployment, generating spatially distributed condition information along entire transmission corridors.

To fully exploit these capabilities, Digital Twin outputs should be integrated into GIS-based asset management platforms. The combination of spatially resolved monitoring data, predictive analytics, and visualization tools enables the identification of critical conductor sections and emerging degradation patterns. Such an integrated framework supports more targeted maintenance strategies, improves resource allocation, enhances grid resilience, and contributes to the long-term sustainability of transmission infrastructure.

7.3 Future Monitoring, Communication and Asset Management Strategy

Future deployments should maintain a representative subset of physical sensors for continued model validation, retraining, and operational supervision. The optimum number and location of retained sensors should be determined through line-specific risk assessments and sensor placement studies.

An important lesson for future implementation is that there is no truly maintenance-free sensor technology. Regardless of the selected monitoring solution, asset owners must be prepared for equipment ageing, calibration requirements, communication failures, data loss, and component replacement throughout the asset lifecycle. Consequently, transmission asset management organizations should establish a structured strategy for managing monitoring device portfolios, including maintenance planning, spare-part management, replacement policies, and backup solutions to ensure long-term monitoring reliability.

As operational experience increases and ANN models demonstrate stable long-term performance, the number of retained physical sensors may be gradually reduced. Nevertheless, complete elimination of physical reference sensors is not recommended. A minimum set of reference devices should remain available for continuous model validation, performance assessment, and retraining activities.

The communication infrastructure supporting Digital Twin and DLR applications should be designed as an integrated and standardized ecosystem. Communication protocols, data exchange mechanisms, and cybersecurity requirements must operate in a coordinated and harmonized manner across all involved systems. This becomes particularly important when developing DLR platforms that aggregate data from multiple monitoring devices, weather stations, SCADA systems, and external information sources. Consistent communication architectures and unified cybersecurity frameworks are essential prerequisites for reliable and scalable system-wide DLR deployment.

Physical sensors removed from service may be redeployed to other transmission lines, while additional weather stations may be installed to improve ANN performance and enhance network-wide observability. This approach could support the long-term scalability and wider adoption of Digital Twin-assisted Dynamic Line Rating systems.

7.4 Potential industrial utilization of the DT

Beyond its scientific and methodological contributions, the project has demonstrated significant industrial and commercial potential. One of the most important outcomes is the realization that Digital Twin technology can substantially reduce the number of physical monitoring devices required across the transmission network. Fewer permanently installed sensors translate into lower maintenance requirements, reduced operational costs, and greater flexibility for future technological upgrades. This finding reinforces two emerging concepts in the field of line monitoring: the large-scale deployment of virtual sensors and the introduction of sensor rental as a new business model.

The project results have been directly transferred into the industrial solutions branded with the GridGuard system. Unlike many commercially available virtual sensor solutions that rely primarily on

weather agency data, GridGuard's Digital Twin-based virtual sensors are supported by AI models trained and validated using real field measurements. This hybrid approach provides more accurate and reliable estimates of conductor conditions and line ratings, thereby increasing system operators' confidence in operational decision-making.

Furthermore, the experiences gained during the TwinEU Digital Twin development have laid the foundation for introducing a sensor rental concept. Under this approach, utilities can temporarily deploy physical sensors to collect training and validation data, after which Digital Twin models and virtual sensors can continue providing network-wide monitoring with significantly reduced hardware requirements. This approach addresses one of the major challenges of modern grid monitoring—scalability. By combining Digital Twins, AI-based virtual sensors, and flexible deployment models, these solutions offer a practical pathway toward cost-effective, large-scale adoption of dynamic line rating and advanced monitoring technologies.

8 Conclusions

Task 6.8 of the TwinEU project has demonstrated that Dynamic Line Rating (DLR) and Digital Twin (DT) solutions are not merely incremental enhancements to existing line monitoring practices but represent a paradigm shift in how power systems can be observed, managed, and optimized. The results confirm that the initial assumptions were well-founded: there is a strong and increasing demand among system operators for scalable, data-driven monitoring solutions capable of supporting operational efficiency extended with new potential applications such as market optimization and asset management.

A key conclusion is that scalability is not only a technical challenge but a systemic one. While high-voltage transmission networks primarily benefit from DLR through enhanced cross-border power exchange and congestion management, distribution utilities face fundamentally different challenges due to the rapid penetration of renewable energy sources and increasingly complex, less transparent power flows. This divergence highlights the necessity of improved vertical cooperation between TSOs and DSOs, supported by interoperable data platforms and coordinated operational frameworks. The project has shown that such cooperation is feasible, though it requires sustained commitment beyond purely technical implementation.

From a technological perspective, the development of the Digital Twin concept proved to be a real option but also emphasized that its effectiveness depends on a well-structured, multi-layered architecture. Sensor placement, deployment logistics, and especially data management emerged as critical components that determine overall system performance. In particular, the project demonstrated that data quality, availability, and consistency are fundamental basis for successful AI-based modelling. Rather than being a secondary concern, data engineering - including preprocessing, anomaly detection, and handling of data gaps - must be treated as a core function of any operational system.

The results also underline that Artificial Neural Network models must be carefully tailored to local conditions. There is no universally optimal model architecture. Our finding is that performance depends on achieving the right balance between accuracy, robustness, and operational usability. Importantly, the project verified that virtual sensing solutions can reach a level of accuracy sufficient to complement, and in some cases partially replace, physical sensors. However, it also reinforced that hybrid approaches remain essential: a minimum set of physical sensors is essential for maintaining secure monitoring, enabling validation, and ensuring long-term reliability.

Another important finding is that the true value of monitoring lies not in single measurements and models, but in actionable outputs. By developing a proprietary DLR algorithm capable of translating thermal conditions into real-time line ratings, the project created a direct link between thermal behaviour and line rating that can be used in operational decision-making. This functionality significantly enhances the usability of the system for congestion management, forecasting, and broader market simulations, while also opening new opportunities in asset management through the ability to simulate non-observable or critical operating conditions.

Beyond technical achievements, the project identified several non-technical barriers that are equally decisive for large-scale deployment. Organizational readiness, IT system maturity, and operator confidence play a central role in determining whether such solutions can be effectively

integrated into daily operations. These aspects are often underestimated, yet they represent the key bottlenecks for scaling up. At the same time, the lack of standardized guidelines and regulatory frameworks for data handling, system validation, and operational use of DLR solutions remains a significant limitation across the sector.

In this context, the project strongly supports the introduction of a comprehensive regulatory framework in Europe, similar in ambition to FERC Order 881. Such a framework should recognize real-time monitoring as a critical infrastructure component and provide clear guidance on sensor deployment, data management, cybersecurity, and system integration. Regulatory support would not only accelerate technical adoption but also ensure consistency, interoperability, and long-term reliability of solutions across different system operators.

Looking forward, the project highlights several potential pathways. DLR-based operation can gradually evolve from monitoring and decision support to operational use and ultimately supervised automation, while predefined contingency scenarios and corrective actions should be established for potential N-1-1 events. In the near term, Digital Twin and DLR technologies should primarily serve as decision-support tools, allowing operators to build confidence in the technology, while their long-term success will depend on the development of appropriate operational procedures, governance frameworks, regulatory acceptance, and the continued responsibility of system operators for final operational decisions.

Future Digital Twin and DLR deployments should rely on a hybrid monitoring architecture, where a representative subset of physical sensors is retained for validation and supervision, while long-term asset management strategies address maintenance, ageing, spare parts, backup solutions, and the gradual optimization of the sensor portfolio. In parallel, scalable DLR implementation requires an integrated and standardized communication ecosystem with harmonized data exchange and cybersecurity practices, enabling reliable integration of monitoring devices, weather stations, SCADA systems, and other operational data sources.

The continued integration of AI-driven Digital Twins into system operation, forecasting, and market applications can significantly enhance grid flexibility and resilience. At the same time, new business models, such as sensor-as-a-service concepts, can reduce implementation barriers and improve cost-efficiency. From a sustainability perspective, the demonstrated ability of DLR systems to unlock additional transmission capacity offers a highly impactful alternative to traditional grid reinforcement, reducing both investment needs and environmental footprint.

In conclusion, the project establishes a solid foundation for the widespread deployment of Digital Twin-based line monitoring systems. While important technical, organizational, and regulatory challenges remain, the results clearly indicate that these solutions provide a strong basis for future large-scale deployment. Their successful adoption will play a critical role in enabling a more flexible, efficient, and sustainable power system.

References

- [1] Farzaneh, M., Farokhi, S., & Chisholm, W. (2012). Electrical design of overhead power transmission lines. McGraw-Hill Education.
https://books.google.hu/books/about/Electrical_Design_of_Overhead_Power_Tran.html?id=PoKD_3KFEzcC&redir_esc=y
- [2] International Council on Large Electric Systems (CIGRE). CIGRE Green Books: Overhead Lines. Springer Nature; 2017. ISBN: 978-3-319-31746-5.
<https://link.springer.com/referencework/10.1007/978-3-319-31747-2>
- [3] Karimi, S., Musilek, P., & Knight, A. M. (2018). Dynamic thermal rating of transmission lines: A review. *Renewable and Sustainable Energy Reviews*, 91, 600-612. Online. Available: <https://www.sciencedirect.com/science/article/pii/S1364032118302119>
- [4] IRENA (2020), Innovation landscape brief: Dynamic line rating, International Renewable Energy Agency, Abu Dhabi. Online. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Jul/IRENA_Dynamic_line_rating_2020.pdf
- [5] Monti, A., Dolley, B., Palensky, P., & Parisot, A. (2024). Building a Digital Twin for European Energy Infrastructure: The role of open source. *IEEE Electrification Magazine*, 12(3), 78-84. Online. Available: <https://ieeexplore.ieee.org/document/10666959>,
<https://palensky.org/pdf/Monti2024.pdf>
- [6] Németh, B., Göcsei, G., Rácz, L., & Szabó, D. (2020). Development and realization of a complex transmission line management system. CIGRE Paris Session: Paris, France. Online. Available: <https://www.e-cigre.org/publications/detail/b2-219-2020-development-and-realization-of-a-complex-transmission-line-management-system.html>
- [7] Szabó, D. et al (2022). Komplex távvezeték menedzsment rendszer alkalmazhatósága a hazai villamosenergia rendszerben. *Elektrotechnika*. Online. Available: https://matarka.hu/cikk_list.php?fusz=188258&nyelv=eng
- [8] Reich, K., Weissnar, R., & Puffer, R. (2020). Investigation of the dynamic rating of tubular busbars in substations. *e & i Elektrotechnik und Informationstechnik*, 137, 455-459. Online. Available: <https://www.e-cigre.org/publications/detail/b3-206-2020-investigation-on-the-dynamic-rating-of-tubular-busbars-in-substations.html>
- [9] Szabó, D., Rácz, L., & Göcsei, G. (2019). A novel approach of critical span analysis. In *The International Symposium on High Voltage Engineering* (pp. 1025-1031). Springer, Cham. Online. Available: https://link.springer.com/chapter/10.1007/978-3-030-31676-1_96
- [10] Rácz, L., Németh, B., & Göcsei, G. (2022). A risk-based, distributed sensor installation concept for high voltage grid monitoring. *Energy Reports*, 8, 266-274. Online. Available: <https://www.sciencedirect.com/science/article/pii/S2352484721012476>
- [11] Official website of “FACilitating Regional CROSS-border Electricity Transmission through Innovation” – FARCROSS project, <https://farcross.eu/> (visited 1 Oct 2022)
- [12] FARCROSS DELIVERABLE D5.5. “Report on final results of demonstration performance” – FARCROSS project, <https://farcross.eu/> (visited 1 Oct 2022)
- [13] Morgan, V. T. (1996). Effect of elevated temperature operation on the tensile strength of overhead conductors. *IEEE Transactions on Power Delivery*, 11(1), 345-352. Online. Available: <https://ieeexplore.ieee.org/document/484034>

- [14] Rácz, L., & Németh, B. (2022). A Novel Concept of Dynamic Line Rating Systems Based on Soft Computing Models. In 2022 10th International Conference on Smart Grid (icSmartGrid) (pp. 131-136). IEEE. Online. Available: <https://ieeexplore.ieee.org/document/9848683>
- [15] Rácz, L. (2023). A complex approach for reliable dynamic line rating systems: Application of DLR systems based on a more accurate analysis of high-voltage networks' real-time thermal conditions (Doctoral dissertation, Budapest University of Technology and Economics (Hungary)). Online. Available: <https://repozitorium.omikk.bme.hu/items/6fd01dec-17f7-4b53-b3fa-4c0934891cc0>
- [16] CIGRE, W. (2014). B2-43, Technical Brochure 601: Guide for Thermal Rating Calculations of overhead lines. Online. Available: <https://www.e-cigre.org/publications/detail/601-guide-for-thermal-rating-calculations-of-overhead-lines.html>
- [17] Conductor Working Group P738. (2013). IEEE 738-2012 Standard for Calculating the Current-Temperature Relationship of Bare Overhead Conductors—Corrigendum 1. Online. Available: <https://standards.ieee.org/ieee/738/4997/>
- [18] ENTSO-E. (n.d.). Common Information Model (CIM) for Grid Models Exchange. Retrieved May 14, 2025, from <https://www.entsoe.eu/digital/common-information-model/cim-for-grid-models-exchange/>
- [19] Zulfiqar, M., Gamage, K. A. A., Kamran, M., & Rasheed, M. B. (2022). Hyperparameter Optimization of Bayesian Neural Network Using Bayesian Optimization and Intelligent Feature Engineering for Load Forecasting. *Sensors*, 22(12), 4446. Online. Available: <https://doi.org/10.3390/s22124446>
- [20] Sultana, N., Hossain, S. M. Z., Almuhaeni, S. H., & Düşteğör, D. (2022). Bayesian Optimization Algorithm-Based Statistical and Machine Learning Approaches for Forecasting Short-Term Electricity Demand. *Energies*, 15(9), 3425. Online. Available: <https://doi.org/10.3390/en15093425>
- [21] IEEE Std 738, IEEE Standard for Calculating the Current-Temperature Relationship of Bare Overhead Conductors, 2023. Online. Available: <https://standards.ieee.org/ieee/738/10207/>
- [22] Szabó, D.; Németh, B. A Novel Approach to Dynamic Line Rating Calculation. *Energies* 2024, 17, 1137. Online. Available: <https://doi.org/10.3390/en17051137>
- [23] United States Of America Federal Energy Regulatory Commission, 18 CFR Part 35, Docket No. RM20-16-000; Order No. 881, "Managing Transmission Line Ratings", (Issued December 16, 2021). Online. Available: https://elibrary.ferc.gov/eLibrary/filelist?accession_num=20211216-3112