Comparing Deep Learning Approaches for Weather Forecasting: Insights from the PRECEDE Project

Maira Aracne[1,*], Tommaso Ruga[2,*], Camilla Lops[3,*], Deborah Federico[4], Luciano Caroprese[3], Ester Zumpano[2], Sergio Montelpare[3], Mariano Pierantozzi[3], Francesco Dattola[4], Pasquale Iaquinta[4], Miriam Iusi[4], Raffaele Greco[4], Marco Talerico[4], Valentina Coscarella[4], Luca Legato[4], Ivana Pellegrino[4], Sonia Bergamaschi[5], Mirko Orsini[5], Riccardo Martoglia[5], Andrea Livaldi[5], Abeer Jelali[5] and Simone Sbreglia[5]

[1]*University Leonardo da Vinci, Piazza San Rocco, 2, Torrevecchia Teatina, Italy*
[2]*DIMES Department, University of Calabria, Via Ponte Pietro Bucci, Rende (CS), Italy*
[3]*Engineering and Geology Department, University G. d'Annunzio of Chieti-Pescara, Viale Pindaro 42, Pescara, Italy*
[4]*e way Enterprise Business Solutions, Via Francesco de Francesco 19, Cosenza, Italy*
[5]*DataRiver Srl, Via Emilia Est, 985, Modena, Italy*

Abstract

Accurately predicting weather conditions in advance is crucial across various sectors. It informs decision-making in agriculture, enables preparation for potential natural disasters, optimizes renewable energy management, and helps reduce energy waste and inefficiencies. In the current historical context, achieving maximum forecast precision has become more critical than ever. Artificial intelligence is transforming the methods and tools we use to reach this goal, paving the way for unprecedented advancements. Its main advantage lies in the ability to manage and evaluate huge amounts of data identifying complex patterns and correlations that could escape from human analysis. The system's fast processing capabilities constitute another fundamental aspect, producing territory-specific forecasts that consider both local micro-climates and distinctive geographical features. The present work aims to compare different deep learning approaches applied to weather forecasting, conducted as part of the PRECEDE Project. Recurrent neural networks are analyzed, in particular Gated Recurrent Units, and Temporal Convolutional Networks, known to be two architectures specialized in modelling data sequences over time horizons. The study highlights the performance of neural networks in enhancing the outputs of the MM5 weather model, a regional mesoscale model, over one-, two-, and three-day time horizons. Furthermore, the work explores the strengths and limitations of each approach, providing insights into their effectiveness, which serve as a foundation for guiding future research and practical applications of deep learning in weather forecasting.

Keywords

Weather forecasting, Deep Learning, Machine Learning, Mesoscale Model 5, Renewable energy prediction

1. Introduction

Weather forecasting plays a crucial role in decision-making across various sectors, including agriculture, disaster management, energy optimization and urban planning. Reliable predictions reduce risks, enhance efficiency and support the transition to sustainable energy systems through better management of renewable sources. As global energy demand rises, advanced forecasting methods help optimize wind and solar energy usage within decentralized models like energy communities - local networks where individuals both produce and consume energy, essential for reducing fossil fuel dependency.

Weather forecasting is challenging due to dynamic patterns. Deep Learning has improved accuracy by analyzing large datasets and identifying complex temporal patterns. These advances help optimize energy systems while maintaining grid stability and meeting prosumer needs.

DL methods, including Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs), and Temporal Convolutional Networks (TCNs), excel at modeling temporal data sequences and capturing climate data trends. Additionally, traditional models like the Fifth-Generation NCAR-/Penn State Mesoscale Model (MM5) maintain their importance due to operational dependability, offering solutions across multiple sectors. Despite these advancements, two significant challenges persist in the pursuit of precise weather forecasting and its integration with energy systems. The first challenge lies in developing a robust platform capable of efficiently managing, integrating, and analysing vast amounts of heterogeneous data from diverse sources. This is essential given the intricate relationship between climatic variables and renewable energy production. The second challenge focuses on leveraging this integrated data to design advanced models and services that can accurately predict energy production and meet evolving management requirements, ensuring both reliability and sustainability.

In this context, the present work aims to address the limitations of current Regional Climate Models (RCMs), including MM5's occasional inaccuracies, through a comparative analysis of deep learning methodologies applied to weather forecasting. Insights are drawn from the PRECEDE project, which explores innovative approaches to enhance prediction accuracy and optimize energy-related applications by integrating the strengths of DL models with traditional fore-

Published in the Proceedings of the Workshops of the EDBT/ICDT 2025 Joint Conference (March 25-28, 2025), Barcelona, Spain
*Corresponding author.
m.aracne@unidav.it (M. Aracne); tommaso.ruga@dimes.unical.it (T. Ruga); camilla.lops@unich.it (C. Lops); dfederico@eway-solutions.it (D. Federico); luciano.caroprese@unich.it (L. Caroprese); e.zumpano@dimes.unical.it (E. Zumpano); sergio.montelpare@unich.it (S. Montelpare); mariano.pierantozzi@unich.it (M. Pierantozzi); fdattola@eway-solutions.it (F. Dattola); piaquinta@eway-solutions.it (P. Iaquinta); miusi@eway-solutions.it (M. Iusi); rgreco@eway-solutions.it (R. Greco); mtalerico@eway-solutions.it (M. Talerico); vcoscarella@eway-solutions.it (V. Coscarella); llegato@eway-solutions.it (L. Legato); ipellegrino@eway-solutions.it (I. Pellegrino); sonia.bergamaschi@unimore.it (S. Bergamaschi); mirko.orsini@datariver.it (M. Orsini); riccardo.martoglia@datariver.it (R. Martoglia); andrea.livaldi@datariver.it (A. Livaldi); abeer.jelali@gmail.com (A. Jelali); simo.sbreglia@gmail.com (S. Sbreglia)

casting frameworks. Additional details about the project can be found in [1].

In detail, machine and deep learning models are implemented to combine real-time measurements with RCM outputs. The models analyze climate data streams to identify seasonal patterns and short-term trends, improving the prediction of key weather parameters - including solar radiation, temperature, relative humidity, and atmospheric pressure - which in turn leads to more accurate forecasts of renewable energy generation. This comprehensive approach optimizes energy storage and distribution for energy communities, while improving the efficiency and sustainability of renewable energy systems at both individual and community scales.

The paper is structured as follows: after this brief introduction, Section 2 reviews the state-of-the-art artificial intelligence techniques currently used for predicting climatic parameters. Section 3 provides an overview of the MM5 model, the weather datasets and the chosen artificial neural networks utilized in this study. Finally, the results are presented in Section 4, followed by a discussion of the main findings in Section 5.

2. Related Works

Predicting climate variables is notoriously difficult due to their dynamic nature, thus considerable effort has been made to apply Artificial Intelligence (AI) to this challenge. Consequently, a new field, Deep Learning for Weather Prediction (DLWP), has born and it has demonstrated impressive results, as shown in [2, 3]. The ability of Neural Networks to learn complex nonlinear relationships and to process vast amounts of data simultaneously enables their application in different fields, such as in solar radiation prediction (at both daily and hourly scales), short-term and long-term wind resource estimation, and in the forecasting of various meteorological parameters such as temperature, precipitation, cyclones, and humidity [2, 4, 5, 6, 7]

Among the available architectures, we chose the GRUs and the TCNs due to their specialization in modeling data sequences. LSTM and GRU are the two main RNN variants that handle long sequences better than vanilla RNNs. After conducting a comparative evaluation between GRU and LSTM on a sample dataset, our analysis revealed comparable performance between the two models. GRU has been chosen for its simpler design, faster training, and more efficient memory use.

The TCNs are well known to outperform RNNs across a broad range of sequence modeling tasks [8]. However, in [9] , the GRUs show better prediction capability than TCNs, but the problem of the correct tuning hyperparameters is opened. Also, the study raises the possibility that results may differ if the lengths of the input or output changes. The full potential of GRUs and TCNs in climate variable prediction remains to be explored, as their application in this domain is still an emerging area of research.

Different solutions exist in the field of energy, or power, forecasting. Shaikh et al. [10] demonstrated that TCNs typically outperform LSTM models. On the other hand, the review conducted in [11] studies the possible advantages and disadvantages of different neural networks for Photovoltaic (PV) power prediction and it finds that the Multilayer Perceptron, RNNs, Convolutional Neural Network, and Graph Neural Network architectures have different forecasting advantages that depend on its specific application scenario.

For weather forecasting applications, studies in [12] and [13] demonstrated TCN models' effectiveness in predicting Global Horizontal Irradiation (GHI) and ten weather parameters, respectively. However, while these works addressed forecasting horizons ranging from minutes (5, 10, 15, and 20) to several hours (up to 9), they operate on different time scales than the here presented research.

Despite notable advances in energy management and forecasting research, most approaches remain fragmented rather than converging into comprehensive solutions. While machine learning has shown promising results in weather prediction [2, 3, 4, 5], solar radiation estimation [6, 14, 15, 16] and temperature forecasting [7, 17, 18], these advances have not been fully integrated into comprehensive energy management systems. Furthermore, existing energy optimization strategies [19, 20, 21, 22] tend to focus on specific components, such as battery management or demand response, without addressing the complex, interconnected nature of energy communities. To overcome these limitations, the PRECEDE project introduces an integrated framework that leverages multiple AI techniques through a modular architecture, comprehensively addressing the energy management pipeline from data integration to community-scale optimization. Unlike previous approaches limited to specific community data, PRECEDE's architecture transcends these limitations offering a generalizable framework that adapts to diverse settings and environmental conditions while bridging the gap between climate forecasting and energy optimization.

3. Background

This section explores both the meteorological and AI aspects of our datasets and proposed models.

3.1. The Fifth Mesoscale Model (MM5)

The Fifth-Generation Penn State/NCAR Mesoscale Model (MM5) is a widely used numerical weather prediction system, developed collaboratively by the Penn State University and the National Center for Atmospheric Research (NCAR). It is designed to simulate mesoscale and regional atmospheric phenomena for both research and operational forecasting.

MM5 is highly adaptable, offering configurable grid resolutions, physical parameterizations, and boundary conditions to suit various meteorological applications. It employs a σ-coordinate system based on hydrostatic pressure and finite-difference numerical schemes, specifically the Arakawa-Lamb B-staggering technique, enabling detailed simulations of convection, radiation, cloud microphysics, and surface-atmosphere interactions.

The non-hydrostatic model relies on conservation equations for momentum and energy, incorporating a tendency equation for perturbation pressure. MM5 predicts meteorological variables like temperature, pressure, wind, solar radiation, and cloud cover, making it suitable for diverse applications, including short-term weather forecasting, climate studies, air quality management, water resource planning, and severe weather analysis.

With portability across computational platforms and extensive documentation, MM5 is accessible to users of varying expertise. Additional details about the MM5 system and its key features are available in Grell et al. [15] and Dudhia et al. [16].

3.2. Datasets

Both case studies, Casaccia and Ottana, utilize datasets with measurements collected at 10-minute intervals. The Casaccia database spans three years, beginning January 1^{st}, 2018 at 1:10 am, and includes four physical quantities: GHI, temperature, atmospheric pressure, and relative humidity. The Ottana dataset covers one year, starting from May 31^{st}, 2021 at 17:10, and comprises three physical quantities: GHI, atmospheric pressure, and relative humidity. The difference in dataset duration arises from their availability and reliability from weather stations in each location. Acquiring continuous and high-quality meteorological records remains a challenge, and the selected databases represent the most comprehensive and accurate data accessible for each site. Additionally, the study's methodological approach accounts for these variations by focusing on relative trends and patterns rather than absolute comparisons, thus maintaining the validity of the performance evaluation across locations. For both locations, each measured quantity is paired with its corresponding MM5 system prediction, and all values have been normalized. Distinct datasets were created for each combination of prediction horizon τ (144 for 1 day, 288 for 2 days, and 432 for 3 days) and target variable i (0 for GHI, 1 for temperature, only for Casaccia, 2 for atmospheric pressure and 3 for relative humidity). Each of these datasets has been divided into three subsets training set (60%), validation set (20%) and test set (20%).

3.3. Artificial Intelligence Techniques

3.3.1. Gated Recurrent Unit

The GRU is a specialized recurrent neural network architecture that excels in modeling sequential data and temporal dependencies. At its core, the GRU cell processes sequential information through a sophisticated gating mechanism that selectively retains or discards information at each time step. This mechanism consists of two primary components: the update gate and the reset gate. The update gate balances the integration of new information with historical context, determining how much of the previous hidden state should persist. Meanwhile, the reset gate controls the forgetting mechanism, allowing the model to discard irrelevant past information. This dual-gate architecture allows GRUs to effectively model long-term dependencies while addressing the vanishing gradient problem inherent in traditional RNNs, as both gates work together to precisely control the temporal evolution of the hidden state. By dynamically managing information flow, GRUs maintain an adaptive memory that evolves with the input sequence, making them particularly effective for tasks like time series forecasting and natural language processing.

3.3.2. Temporal Convolutional Networks

The TCN is a neural network specialized in modeling data sequences, drawing inspiration from the operational mechanism of Convolutional Neural Network, which uses filters to recognize patterns in data. However, instead of performing two-dimensional convolution (as with images), it operates

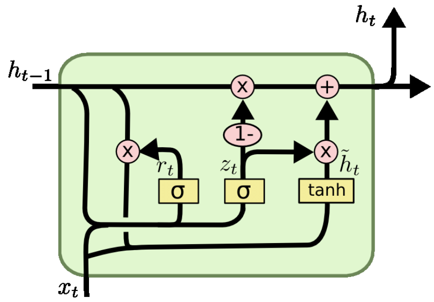


Figure 1: GRU basic building block.

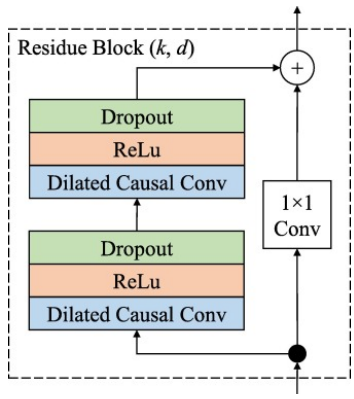


Figure 2: TCN basic building block.

in a single dimension (time series). In this case, the convolution is made causal, meaning the network learns to predict the output at time t by only considering data up to time t, avoiding the use of future information to predict the present. A key requirement for a forecasting model is that each output element should depend on all historically preceding input elements: the TCN adopts dilated convolutions to expand its receptive field without dramatically increasing the number of parameters. This technique ensures coverage of extensive sequence portions without losing resolution or computational efficiency, even with a small convolution kernel. Importantly, expanding the receptive field enhances the network's ability to capture long-term dependencies. To address the vanishing gradient problem, particularly in deep networks, residual connections are employed, creating a direct path between the network's input and output. Currently, TCNs are considered an alternative to RNNs, including their more sophisticated versions, LSTM and GRU. Some advantages of TCNs are their ability to parallelize work, which streamlines the training process, and the absence of a recursive structure, which leads to more stable gradient propagation.

3.3.3. Deep Reinforcement Learning

To manage energy communities, the global PRECEDE framework has adopted a DRL approach, widely considered one of the most effective methodologies in this field. This advanced computational paradigm combines DL architectures with the Reinforcement Learning (RL) framework, enabling robust solutions for complex decision-making processes. However, it is important to note that the DRL approach was not utilized in the analyses conducted here.

RL operates on the principle of sequential decision-making, where an agent interacts with an environment through an iterative process of observation, action, and reward. The environment is typically modeled as a Markov Decision Process (MDP), characterized by a state space S, an action space A, and a reward function R. At each time step t, the agent observes the current state $s_t \in S$ and selects an action $a_t \in A$ based on its policy $\pi(a|s)$, which maps states to action probabilities. Following the action, the environment transitions to a new state s_{t+1} and provides a reward signal r_t. The agent's objective is to learn an optimal policy π^* that maximizes the expected cumulative discounted reward, expressed as $E[\sum \gamma^t r_t]$, where $\gamma \in [0, 1]$ is the discount factor balancing immediate and future rewards.

DRL represents a sophisticated integration of deep learning architectures with Reinforcement Learning principles, designed to handle complex decision-making tasks in high-dimensional spaces. The integration of deep learning enhances the agent's capability to identify complex patterns and hierarchical representations, leading to more sophisticated decision-making strategies.

4. Discussion

To fully understand the PRECEDE project, a brief introduction strictly connected to the layers which compose the proposed system is resumed in Fig. 3.
The PRECEDE architecture is composed of four different layers, which belong to two different processing steps: the data integration and processing step and the energy production forecasting and managing step. The Data Flow layers, comprising the Data Integration Layer and Climate Variables Broadcasting Layer, handle the acquisition, integration, and processing of heterogeneous data sources, providing reliable climate forecasts through advanced DL models. The Energy Management layers, consisting of the Energy Production Forecasting Layer and Energy Flow Optimization Layer, leverage these forecasts to optimize energy production and distribution within the community through physical models and multi-agent reinforcement learning techniques.

The first step includes the Data Integration and Climate Variables Broadcasting layers, which are involved in the acquisition, integration, and processing of heterogeneous data sources to produce forecasts of climatic variables. The data used in this step come from two different sources, as previously introduced: real data, which belongs to physical weather stations, and MM5 model predictions. To assess this stage, the advanced deep learning models introduced in Fig. 3 are adopted.

These weather forecasts are then integrated into an efficient prediction system for prosumers, using physical models and multi-agent RL to optimize energy production and distribution. Our analysis compares two models for climate variable prediction: GRUs and TCNs. We evaluate their similarities, differences, and relative performance. These architectures were selected for their ability to track temporal data evolution while identifying relevant patterns.

These supervised learning models enhance the prediction of key climatic variables - including GHI, temperature, pressure, and relative humidity. By combining RCMs forecasts with actual observations as inputs, the models produce more accurate predictions and reduce the RCMs' tendency to over- or underestimate values during certain periods [23, 24].

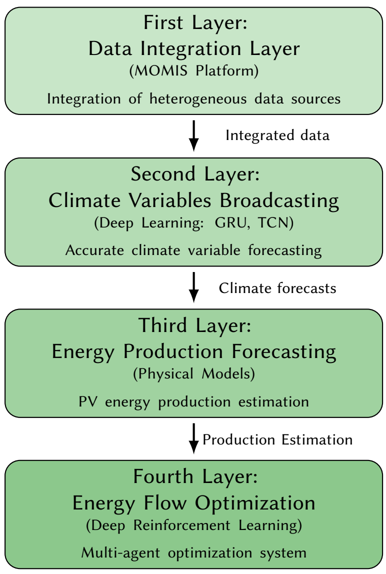


Figure 3: PRECEDE system layers overview.

4.1. Experimental Settings

The model was developed in Python (version 3.11.11) using PyTorch (version 2.5.1+cu121) and a Tesla T4 GPU with 51 GB of RAM and 15 GB of VRAM.

The training process was configured with a maximum of 200 epochs and with the patience parameter in the early stopping mechanism set to 10 iterations.

4.2. Results

To evaluate the comparative performance of the GRU and TCN models, we computed the Mean Absolute Error (MAE), the Coefficient of Determination (R^2) and analyzed their respective Taylor diagrams.

The MAE, reported in Equation (1), measures the average absolute difference between the predicted values and the actual target values by giving equal weight to all errors, regardless of their direction (overestimates or underestimates).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

The Coefficient of Determination (CoD), calculated as defined in Equation (2), measures how well the model is able to predict the variance of the data. The closer its value is to unity, the better the model fits the data. $\bar{y}$ denotes the mean of the observed data, while $\hat{y}$ the predicted value.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2}$$

As presented in [23, 24], the AI model manages to improve the forecasts returned by the MM5 RCM, which exhibits both positive and negative deviations from the real data. In this experiment, good performances are achieved for both the GRUs and the TCNs, as shown in Tables 1 and 2 .

Considering the Casaccia study (see Table 1), it is readily observable that the GRU and TCN models applied to the MM5 model outperform it. The most impressive results

Variable	System	$\tau = 144$ (1 day)		$\tau = 288$ (2 days)		$\tau = 432$ (3 days)	
		MAE	R^2	MAE	R^2	MAE	R^2
GHI	MM5	46.557	0.867	46.472	0.863	45.365	0.867
	GRU	39.140	0.902	39.863	0.892	37.907	**0.944**
	TCN	**37.755**	**0.906**	**38.778**	**0.902**	**37.626**	0.903
Temperature	MM5	2.389	0.841	2.388	0.840	2.361	0.844
	GRU	**1.232**	**0.948**	1.366	**0.940**	**1.313**	**0.942**
	TCN	1.252	0.947	**1.329**	**0.940**	1.363	0.937
Pressure	MM5	173.116	0.863	173.264	0.857	170.598	0.867
	GRU	90.423	0.954	90.062	0.948	90.151	0.952
	TCN	**77.956**	**0.966**	**85.129**	**0.956**	**81.097**	**0.961**
Humidity	MM5	10.720	0.571	10.735	0.570	10.715	0.581
	GRU	**8.313**	**0.733**	**8.498**	**0.724**	**8.480**	**0.728**
	TCN	8.371	**0.733**	8.775	0.706	8.654	0.716

Table 1
Comparison of performance metrics for Casaccia.

Variable	System	$\tau = 144$ (1 day)		$\tau = 288$ (2 days)		$\tau = 432$ (3 days)	
		MAE	R^2	MAE	R^2	MAE	R^2
GHI	MM5	60.230	0.846	59.837	0.844	57.498	0.856
	GRU	52.595	0.887	**53.566**	0.882	**53.105**	0.878
	TCN	**50.282**	**0.898**	53.846	**0.933**	53.471	**0.884**
Pressure	MM5	**60.757**	**0.977**	**64.548**	**0.975**	**64.392**	**0.977**
	GRU	70.594	0.964	72.816	0.968	95.708	0.954
	TCN	67.796	0.967	70.835	0.971	74.626	0.969
Humidity	MM5	10.353	0.680	10.158	0.685	9.982	0.697
	GRU	**7.794**	**0.817**	8.239	**0.794**	9.049	0.750
	TCN	8.088	0.797	**8.196**	0.786	**7.933**	**0.799**

Table 2
Comparison of performance metrics for Ottana.

can be seen in atmospheric pressure forecasting, where the TCN reduce the error of over the 54.97%, 50.86% and 52.46% with 1-day, 2-day and 3-day forecast horizon. At the same time, the CoD is raised by the 11.94%, 11.55% and 10.84%, respectively. Similarly, for the GHI parameter the TCNs exhibit superior performance, although often comparable to that of GRUs. The best results are achieved in the 1-day forecast with a MAE decrease of 18.90% associated to the TCN, and in the 3-day forecast with an R^2 increase of 8.9% for the GRU. Regarding temperature and relative humidity, the GRU shows slightly better results than the TCN, although their results remain very similar.

Concerning Ottana's analysis reported in Table 2, the MM5 model achieves better results than the ANNs regarding pressure forecasting; nevertheless, they still maintain remarkable effectiveness. In summary, a high degree of correspondence is observed between the neural networks' forecasts and the experimental data, validating their predictive capabilities. The best MAE and R^2 performance can be read in the humidity 1-day GRU prediction, with a reduction of 24.73% and an increase of 21% compared to the MM5 forecast.

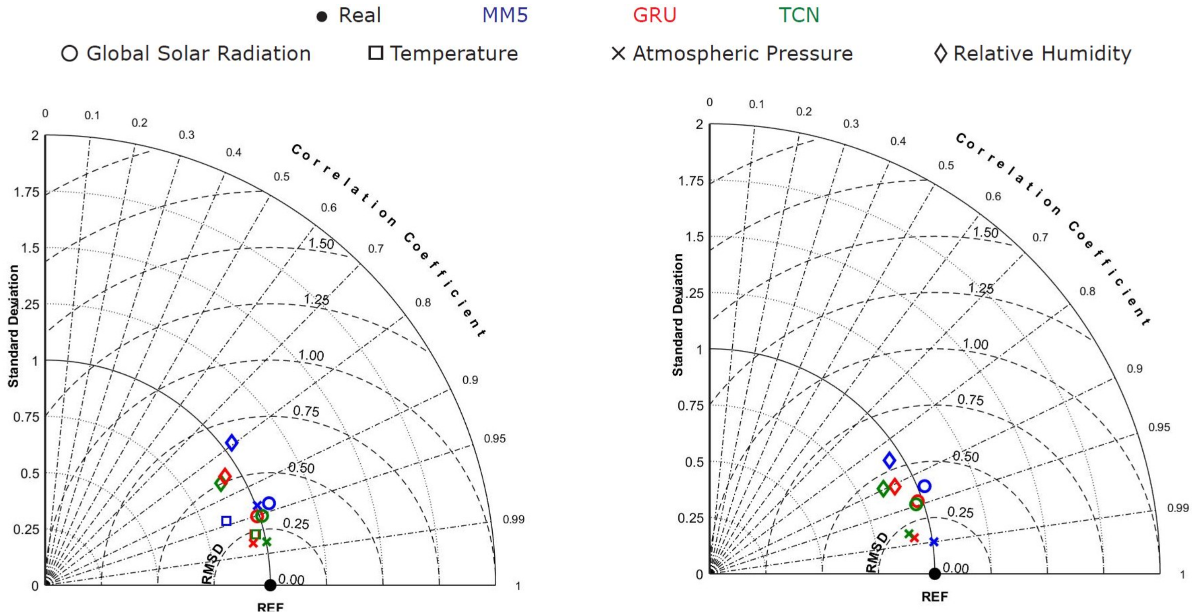
After having discussed the comparison between the neural networks and the MM5 model, the following section presents a comparative assessment of GRUs and TCNs only.

It is important to emphasize that the two datasets include different time intervals: the Casaccia database covers three years, while the Ottana dataset only one. In both cases, the TCN model outperforms the GRU in pressure forecast and it shows significant performance gains. In detail, for the Casaccia study, GRU and TCN models show comparable performance across all parameters except atmospheric pressure. However, in the Ottana analysis, TCN consistently outperforms GRU, though both models maintain high accuracy. The performance difference between the two locations may stem from their different training data durations (three years for Casaccia versus one year for Ottana). Similarly, both the superior performance of MM5 and TCN's enhanced learning capabilities in the Ottana dataset likely reflect the limited one-year training period. Future studies should investigate this relationship by testing model performance across different time spans.

Beyond standard R^2 and MAE statistical measures, the study incorporates Taylor diagrams as visual analytical tools to evaluate how well the method performs in handling multiple variables simultaneously. These diagrams integrate three key statistical measures into a single polar plot: the Standard Deviation (σ), Correlation Coefficient (R), and Centered Root Mean Square Difference (RMSD). The reference observations are positioned at the plot's origin, while predicted values appear as points distributed across the diagram based on their statistical properties.

The standard deviation functions as a measure of the variability of the data, calculating how the values deviate from the mean. The correlation coefficient indicates the strength of the relationship between variables on a scale of -1 to $+1$, where zero shows no relationship, positive values indicate parallel movement and negative values suggest inverse relationships. The root mean square deviation evaluates prediction accuracy by measuring the typical distance between corresponding points in two datasets, with smaller values indicating better alignment. For any comparison between simulated values (f) and reference measurements (r), these metrics are calculated using established statistical formulations as detailed in [25].

Figures 4 through 6 present the Taylor diagrams for Casac-



(a) (b)

Figure 4: Taylor diagrams for 1-day predictions referred to (a) Casaccia and (b) Ottana.

cia and Ottana, corresponding to 1-day, 2-day, and 3-day predictions, respectively. It is important to note that temperature analysis is limited to the Casaccia database, while the remaining climatic parameters are analyzed for both locations.

For Casaccia, the Taylor diagram for 1-day predictions (see Fig. 4a) illustrates that GRU and TCN outperform MM5 across all weather variables. Specifically, both GRU and TCN achieve high correlation coefficients (above 0.9) for global solar radiation and temperature, positioning them close to the reference point and indicating strong predictive accuracy. In the case of atmospheric pressure, GRU and TCN also perform well, with GRU slightly closer to the reference point. For relative humidity, TCN emerges as the most accurate model, exhibiting the highest correlation and the closest match to observed data, while MM5 demonstrates the lowest correlation and the largest deviations across all variables. Overall, TCN and GRU are identified as the most reliable models for weather prediction in Casaccia.

In Ottana (see Fig. 4b), the Taylor diagram highlights the capabilities of GRU and TCN in predicting global solar radiation. For relative humidity, the models show mixed results: one side achieves higher correlation coefficients and lower RMSD, while the other side exhibits poorer standard deviation. In this case, MM5 performs better in terms of standard deviation. Finally, the RCM model surpasses the AI models in estimating atmospheric pressure values.

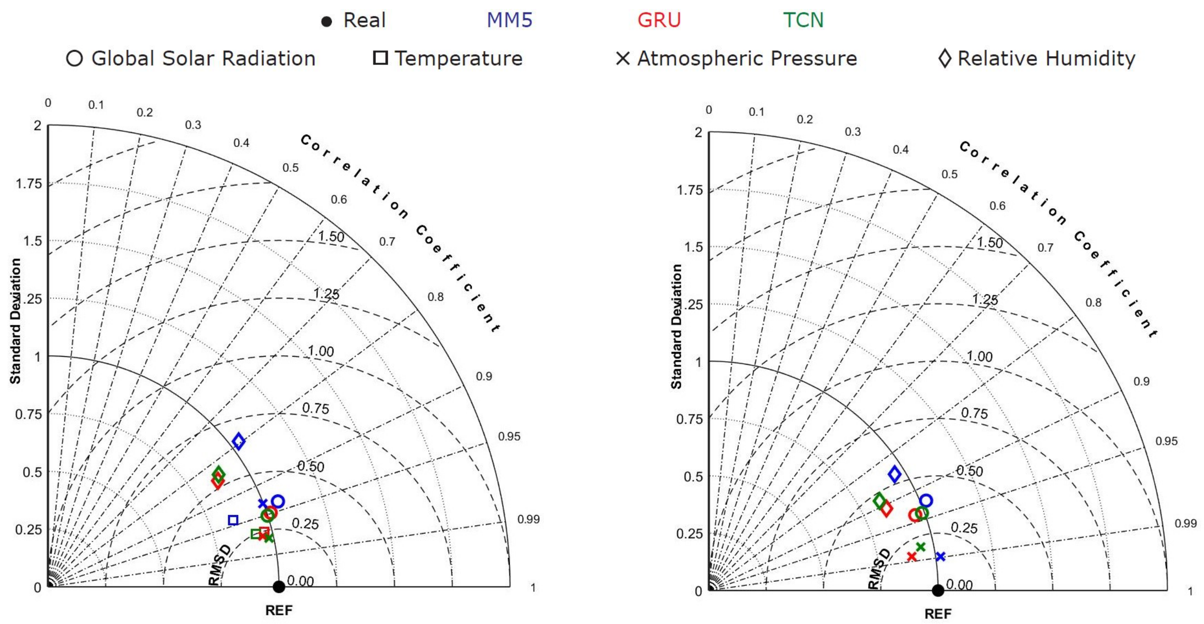
The observed trends for each locality are consistently replicated in the 2-day (Fig. 5) and 3-day (Fig. 6) predictions, underscoring the ability of both TCN and GRU to outperform traditional regional climate models in weather forecasting.

5. Conclusions and Future Perspectives

Conducted within the framework of the PRECEDE project, this study demonstrated the effectiveness of deep learning approaches in enhancing weather forecasting accuracy, particularly through applying GRU and TCN architectures to improve MM5 model predictions. The comparative analysis reveals several key findings that advance the field of weather prediction and its applications in energy management. The results show that both GRU and TCN models generally outperform the traditional MM5 model across multiple weather parameters and time horizons. Notably, TCN demonstrated superior performance in atmospheric pressure forecasting, achieving error reductions of up to 54.97% in one-day forecasts for the Casaccia dataset. While both neural network architectures showed comparable effectiveness in predicting most parameters, TCN exhibited slightly stronger learning capabilities in scenarios with limited training data.

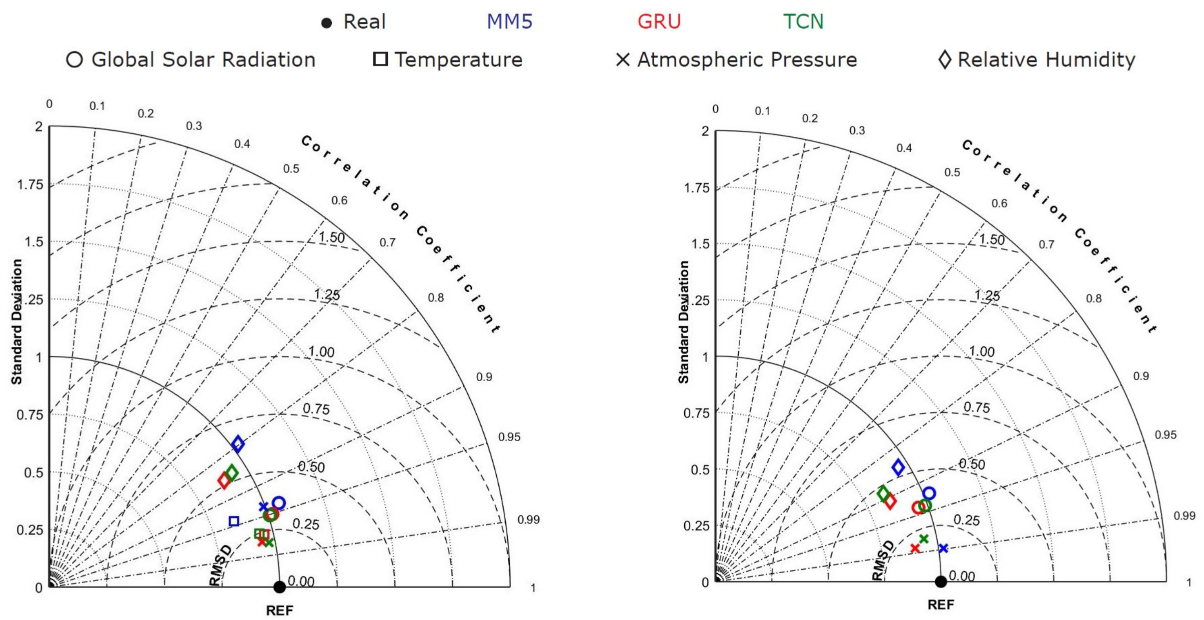
An important finding emerged regarding the impact of training data duration on model performance. The contrasting results between Casaccia and Ottana suggest that the length of the training period significantly influences prediction accuracy. This was particularly evident in the Ottana dataset, where MM5 maintained superior performance in pressure forecasting, highlighting the importance of comprehensive training data for neural network models. The Taylor diagram analysis further validated these findings, demonstrating high correlation coefficients for both GRU and TCN in predicting global solar radiation and temperature. This superior performance was maintained across different prediction horizons (1-, 2-, and 3-day forecasts), confirming the models' reliability and stability.

These analyses serve as the foundation for the PRECEDE project's next phase, which aims to extend these forecasting methodologies to cities in Emilia Romagna. The insights gained from the Casaccia and Ottana studies will inform the implementation of these models across the region, supporting the project's goal of enhancing renewable energy management and community-based energy systems. Moreover, expanding the analysis to diverse climatic conditions and extended temporal scales will provide deeper insights into the robustness and adaptability of the proposed models. We would also like to test the Transformer model.



(a) (b)

Figure 5: Taylor diagrams for 2-day predictions referred to (a) Casaccia and (b) Ottana.



(a) (b)

Figure 6: Taylor diagrams for 3-day predictions referred to (a) Casaccia and (b) Ottana.

Acknowledgments

This work was supported by the PRECEDE project (CUP: E93C22001100001), from the resources of the National Recovery and Resilience Plan (NRRP), funded by the European Union - NextGenerationEU.

Declaration on Generative AI

Either:
The author(s) have not employed any Generative AI tools.

Or (by using the activity taxonomy in ceur-ws.org/genai-tax.html):
During the preparation of this work, the author(s) used X-GPT-4 and Gramby in order to: Grammar and spelling check. Further, the author(s) used X-AI-IMG for figures 3 and 4 in order to: Generate images. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] F. Dattola, P. Iaquinta, M. Iusi, D. Federico, R. Greco, M. Talerico, V. Coscarella, L. Legato, I. Pellegrino,

S. Bergamaschi, et al., Precede: Climate and energy forecasts to support energy communities with deep learning models, in: 2024 IEEE International Conference on Big Data (BigData), IEEE, 2024, pp. 4650–4658.

[2] N. Benti, M. Chaka, A. Semie, Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects, Sustainability 15 (2023). URL: https://www.mdpi.com/2071-1050/15/9/7087.

[3] X. Ren, X. Li, K. Ren, J. Song, Z. Xu, K. Deng, X. Wang, Deep learning-based weather prediction: A survey, Big Data Research 23 (2021). URL: https://www.sciencedirect.com/science/article/pii/S2214579620300460.

[4] J. Fan, L. Wu, F. Zhang, H. Cai, W. Zeng, X. Wang, H. Zou, Empirical and machine learning models for predicting daily global solar radiation from sunshine duration: A review and case study in china, Renewable And Sustainable Energy Reviews 100 (2019) 186–212. URL: https://www.sciencedirect.com/science/article/pii/S1364032118307226.

[5] Q. Yang, C. Lee, M. Tippett, A long short-term memory model for global rapid intensification prediction, Weather And Forecasting 35 (2020) 1203–1220.

[6] M. Mohandes, S. Rehman, T. Halawani, Estimation of global solar radiation using artificial neural networks, Renewable Energy 14 (1998) 179–184. URL: https://www.sciencedirect.com/science/article/pii/S0960148198000652.

[7] M. Ghadiri, A. Marjani, S. Mohammadinia, S. Shirazian, An insight into the estimation of relative humidity of air using artificial intelligence schemes, Environment, Development And Sustainability 23 (2021) 10194–10222. URL: https://doi.org/10.1007/s10668-020-01053-w.

[8] S. Bai, J. Z. Kolter, V. Koltun, An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint arXiv:1803.01271 (2018).

[9] H. Qin, Comparison of deep learning models on time series forecasting: a case study of dissolved oxygen prediction, arXiv preprint arXiv:1911.08414 (2019). doi:10.48550/arXiv.1911.08414.

[10] A. K. Shaikh, A. Nazir, N. Khalique, A. S. Shah, N. Adhikari, A new approach to seasonal energy consumption forecasting using temporal convolutional networks, Results in Engineering 19 (2023) 101296. URL: https://www.sciencedirect.com/science/article/pii/S2590123023004231. doi:https://doi.org/10.1016/j.rineng.2023.101296.

[11] J. Yu, X. Li, L. Yang, L. Li, Z. Huang, K. Shen, X. Yang, Z. Xu, D. Zhang, S. Du, Deep learning models for pv power forecasting: Review, Energies 17 (2024) 3973. doi:10.3390/en17163973.

[12] R. Elmousaid, N. Drioui, R. Elgouri, H. Agueny, Y. Adnani, Accurate short-term ghi forecasting using a novel temporal convolutional network model, e-Prime - Advances in Electrical Engineering, Electronics and Energy 9 (2024) 100667. URL: https://www.sciencedirect.com/science/article/pii/S277267112400247X. doi:https://doi.org/10.1016/j.prime.2024.100667.

[13] P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, Y. Liu, Temporal convolutional neural (tcn) network for an effective weather forecasting using time-series data from the local weather station, Soft Computing 24 (2020) 16453–16482. URL: https://doi.org/10.1007/s00500-020-04954-0. doi:10.1007/s00500-020-04954-0.

[14] M. Alam, F. Al-Ismail, M. Hossain, S. Rahman, Ensemble machine-learning models for accurate prediction of solar irradiation in bangladesh, Processes 11 (2023). URL: https://www.mdpi.com/2227-9717/11/3/908.

[15] A. Mellit, S. Kalogirou, Artificial intelligence techniques for photovoltaic applications: A review, Progress In Energy And Combustion Science 34 (2008) 574–632. URL: https://www.sciencedirect.com/science/article/pii/S0360128508000026.

[16] A. Yadav, S. Chandel, Solar radiation prediction using artificial neural network techniques: A review, Renewable And Sustainable Energy Reviews 33 (2014) 772–781. URL: https://www.sciencedirect.com/science/article/pii/S1364032113005959.

[17] A. Paniagua-Tineo, S. Salcedo-Sanz, C. Casanova-Mateo, E. Ortiz-García, M. Cony, E. Hernández-Martín, Prediction of daily maximum temperature using a support vector regression algorithm, Renewable Energy 36 (2011) 3054–3060. URL: https://www.sciencedirect.com/science/article/pii/S0960148111001443.

[18] D. Kreuzer, M. Munz, S. Schlüter, Short-term temperature forecasts using a convolutional neural network — an application to different weather stations in germany, Machine Learning With Applications 2 (2020) 100007. URL: https://www.sciencedirect.com/science/article/pii/S2666827020300074.

[19] G. Palma, L. Guiducci, M. Stentati, A. Rizzo, S. Paoletti, Reinforcement learning for energy community management: A european-scale study, Energies 17 (2024).

[20] K. X. Perez, M. Baldea, T. F. Edgar, Integrated hvac management and optimal scheduling of smart appliances for community peak load reduction, Energy and Buildings 123 (2016).

[21] W. Jiatong, L. Li, Z. Jiangfeng, Deep reinforcement learning for energy trading and load scheduling in residential peer-to-peer energy trading market, International Journal of Electrical Power & Energy Systems 147 (2023).

[22] Y. Ye, Y. Tang, H. Wang, X.-P. Zhang, G. Strbac, A scalable privacy-preserving multi-agent deep reinforcement learning approach for large-scale peer-to-peer transactive energy trading, IEEE Transactions on Smart Grid 12 (2021).

[23] L. Caroprese, M. Pierantozzi, C. Lops, S. Montelpare, Dl2f: A deep learning model for the local forecasting of renewable sources, Computers & Industrial Engineering 187 (2024). URL: https://www.sciencedirect.com/science/article/pii/S0360835223008094.

[24] C. Lops, M. Pierantozzi, L. Caroprese, S. Montelpare, A deep learning approach for climate parameter estimations and renewable energy sources, in: 2023 IEEE International Conference on Big Data (BigData), 2023, pp. 3942–3951.

[25] K. E. Taylor, Taylor diagram primer, Work Paper (2005) 1–4. doi:doi:10.1029/2000JD900719.