Benefit of Weather Integration for Industrial Electric Load Forecasting - A Benchmark Study

David Gekeler
Fraunhofer IPA
Stuttgart, Germany
0009-0001-6609-7402

Lukas Baur*
Fraunhofer IPA
University of Stuttgart
Stuttgart, Germany
0000-0002-6265-9877
*Corresponding author

Felix Schnell
Fraunhofer IPA
University of Stuttgart
Stuttgart, Germany

Philipp Hennig
Tübingen AI Center
University of Tübingen
Tübingen, Germany
0000-0001-7293-6092

Alexander Sauer
Fraunhofer IPA
University of Stuttgart
Stuttgart, Germany
0000-0003-3822-1514

***Abstract*—This paper investigates weather data integration into industrial electric load forecasting, specifically focusing on manufacturing companies in Germany without onsite generation. To assess the impact of weather factors such as temperature, humidity, wind speed, and precipitation on forecasting accuracy, day-ahead forecasts were conducted for ten German manufacturing companies using three different weather data sources and three model architectures (XGBoost, Regression, and Gaussian Process Regressor), all compared to a baseline. The analysis indicates that weather plays a relatively minor role in forecasting accuracy. The reason may be that industrial electricity consumption does not very often strongly depend on the weather, and extreme weather fluctuations in Germany are rare. Furthermore, other features in the models implicitly capture most seasonal weather trends.**

***Index Terms*—Short-term load forecasting, Weather, Feature Selection, Accuracy,**

I. Introduction

Weather is an important factor for several kinds of electricity forecasts, including price forecasts, feed-in forecasts, grid congestion management, and load forecasts [1] [2] [3]. Many industrial processes are also weather-dependent, especially for cold and heat processes, air conditioning, and electric cooling/heating [4]. For aggregated load forecasts on a grid operator level, the usage of weather variables is well established, and their influence is well researched. In [2], the authors report the use of weather features for 63%, and in [5] for 50% of their reviewed scientific works, respectively. Integrating weather dependencies into load forecasting can significantly improve predictions [6]. This can help stabilize the grid, for instance, by enabling peak shaving, i.e., moving flexible loads from demand peaks to off-peak hours. For companies, load prediction can help lower their electricity spending [7], especially if dynamic electricity contracts based on spot-market prices are used. However, acquiring commercial weather data may outweigh these savings, preventing the widespread adoption of this approach.

This work examines weather features' dependence on industrial load forecasts, i.e., day-ahead forecasts for electricity demand for manufacturing companies in Germany. It addresses the following research questions:

- What added value does the integration of weather data for forecasting electrical loads have for individual companies in the manufacturing industry?
- How does the forecasting performance differ when using different weather data sources?

A. Related Work

The integration of weather data for forecasting electrical loads at the grid level is well-studied. Overall, depending on the source [2], [8], 41-63% of the papers studied integrate weather somehow. However, there is a significant bias towards studies on higher aggregation levels, as 87% of the papers forecast some kind of aggregated load, e.g., utilities or full grids [2]. This is likely due to a lack of publicly available industrial load datasets [9].

Focusing on the integration of weather data for industrial electric load forecasting, little research has been done. In [7], an Italian manufacturing company's load has been forecasted using several linear model configurations. In a previous analysis, weather variables, above all ambient temperature, were examined. It was concluded that the weather does not provide any advantage for this load forecast

A similar finding was made by [10], which predicted the load of a printed circuit board industry facility using an ensemble hidden Markov model. The temperature had a small influence on the load.

Several studies integrate meteorological features into the forecast for singular industrial facilities (as in [10], [11]), but do not measure the effect of these features.

Accordingly, different from forecasting residential and aggregated loads, the influence of weather on industrial customers has only been examined to a limited extent and, if at all, for a single company.

B. Contributions

There are two main contributions in this paper that address the questions above. First, the weather dependency on demand forecasting is assessed based on an explorative analysis for ten German companies. Second, three different weather data sources are evaluated with regard to their quality. To do so, they are integrated into a machine-learning forecasting benchmark and compared based on their model performance.

The rest of the paper is structured as follows: Section II describes the experiments carried out before the findings are discussed in Section III. The latter section is divided into an exploratory part in which the weather influences on the companies' loads are examined. The second part then compares three different weather sources for the relevant companies. The work concludes with a discussion, listing limitations, and future work.

II. Experiment Setup

A forecasting pipeline was set up to analyze the influence of weather on the performance of load forecasting models. This section briefly describes the forecasting experiments conducted, the datasets and models used, and the required preprocessing steps.

A. Load Datasets

The analysis was carried out on ten datasets from the German manufacturing industry, anonymized to D01 to D10. As there is a major weather dependency for companies with behind-the-meter solar or wind generation, these companies were excluded from this study. An overview of the load dataset descriptions can be found in Table I, a subset of load curves is shown in Figure 1. The input resolutions varied between 15 seconds and 15 minutes and have been remapped to 15 minutes for all datasets to enhance comparability. The recording periods range from June 2018 to August 2024, and the sets differ in length between 32 and 66 months.

TABLE I: Load Dataset Overview

Load Dataset	Properties				
	Sector	*State*[a]	*Size*[b]	*Peak / Total*[c]	
D01	Wood Processing	RP	32	7.7	55.2
D02	Furniture Manufacturing	BW	50	0.53	1.14
D03	Sheet Metal Processing	BY	42	0.40	1.40
D04	Glass Processing	BW	60	1.05	3.29
D05	Metal Processing	BW	48	3.70	18.10
D06	Metal Processing	BW	48	6.40	29.60
D07	Metal Processing	BW	54	2.00	9.90
D08	Metal Processing	SH	30	0.67	2.24
D09	Dairy Products	BY	36	10.26	50.80
D10	Rubber Manufacturing	SH	66	1.00	4.29

[a]Locations: German subdivision codes
[b]Dataset length in months
[c]Loads: yearly peak (in MW) / yearly consumption (in GWh)

During data preprocessing, the following steps were carried out: standardization to the same unit and magnitude, removal of extreme outliers (five standard deviations of the mean), manual check for long temporal gaps and erroneous values, resampling, and imputation of smaller gaps.

B. Weather Datasets

Three different weather sources were examined. The weather API from the provider Metomatics [12] (**Meteo**) was integrated as a commercial product. The data from Germany's National Meteorological Service (**DWD**), requested over the

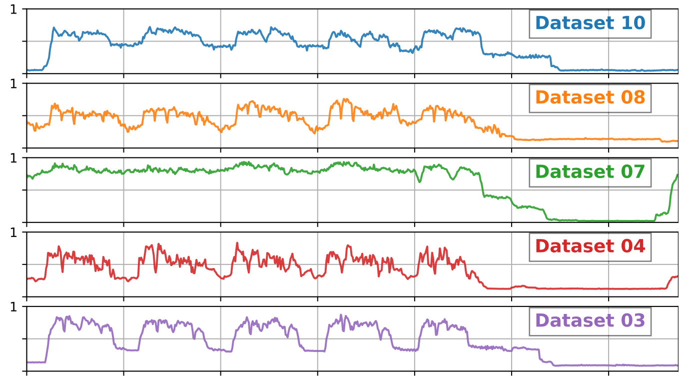

Fig. 1: Load samples: typical week (normalized to [0, 1])

wetterdienst [13] library, and the weather data from **NASA** POWER [14], [15] were used as freely available open data APIs. Different weather features were available depending on the provider. The features shown in Table II were identified. For better comparability, all weather data were retrieved at the lowest common resolution, i.e., hourly resolution.

TABLE II: Weather Sources Overview

Weather Feature	Sources / APIs		
	DWD	*NASA*	*Meteo*
Temperature in °C	mean 2m air temp	mean 2m air temp	2m air temp
Humidity in %	rel. humidity	rel. humidity at 2m	rel. humidity at 2m
Wind Speed in m/s	Mean wind speed	Mean wind speed at 10m	Wind Speed at 10m
Precipitation in mm	hourly	hourly	hourly

C. External Features

In addition to the weather features, calendar information, time variables, and lagged values were integrated as additional external features. Binary features were introduced to distinguish between working days and weekends and to identify public holidays. The temporal context was encoded as follows: hour, day of the week, and month as *one-hot*, time relative to day, week, and year as *fractional* encoding. The values of the previous day and the previous week were used as lagged values, as identified in a prior autocorrelation analysis.

D. Forecasting-Setting and Used Models

The forecasting setting was a multivariate day-ahead multi-step forecast, i.e., a daily 96-time-step prediction horizon (15-minute interval). The resolution was chosen based on that of the electricity market. Four types of forecasting architectures were integrated:

- Baseline (**Lag**): due to the weekly periodicity, last week's value was used as the current day's prediction. This resulting time series serves as a reference for the machine-learning-based approaches.

- Multi-Linear Regression model (**MLR**): A simple and interpretable regression that models linear input feature dependencies. Regularization was added to ensure numerical stability and prevent overfitting.
- **XGBoost**, a boosted regression tree; chosen due to its popularity in the time series forecasting community [16] and its powerful results for other load forecasting tasks [17].
- Gaussian Process Regression (**GPR**): Gaussian processes form a non-parametric model class, being able not only to predict a value for each prediction time point, but also to quantify the uncertainty. Besides its implicit interpretability property, it was chosen because other works [18] already use GPR for industrial load forecasting.

The models were each trained and hyperparameter-tuned on two-thirds of the data, with the final third remaining as a test set for model evaluation.

E. Metric

A relative model evaluation metric was chosen. This allows for a comparison across companies. Due to the instability of measured values close to zero (which was the case here), the symmetrical SMAPE (as defined in [19]) was used instead of the MAPE.

III. Results

A. Exploration: Weather Dependencies

The data exploration provides several insights about the given load data. On the one hand, the correlations observed between load and non-weather features in Figure 4 are close to expected, with the possible exception of the binary *is_holiday* variable having a surprisingly small influence. While only D02 was used as an illustrative example, non-weather feature correlations are similar for all datasets. However, only a few minor weather relationships are observable in the data, as indicated by the low correlation coefficients in Figure 3. Particularly, wind speed and precipitation seem to have close to no correlation with load, for any dataset. The most significant correlation is between load and temperature in D09 and can be observed in Figure 2. An explanation for this could be industrial cooling and heating processes, whose energy requirements depend on the outside temperature.

B. Model performance without weather information

The SMAPE test values for each dataset and model are presented in Table III. These results are obtained without utilizing weather data to confirm the feasibility of forecasting based solely on lagged values, calendar information, and time variables. Further, they are used as a baseline for weather feature analysis. While model performances vary per dataset, some tendencies can be observed: First, as expected, the baseline lag model is consistently outperformed by other models. The same holds true in general for MLR, although its performance fluctuates significantly - in many cases, MLR shows large errors, but in some cases it matches the performance of XGBoost and GPR. Those last two models show the lowest SMAPE values, with XGBoost performing best for five of the ten companies, and GPR for four.

TABLE III: Forecasting model results (measured as SMAPE)

	Forecasting Errors (SMAPE)			
Dataset	***Lag***	***MLR***	***XGBoost***	***GPR***
D01	21.95	8.48	10.10	8.85
D02	24.48	48.53	19.37	26.13
D03	23.83	19.89	18.66	17.52
D04	19.91	21.60	18.76	17.09
D05	28.84	32.63	20.31	22.86
D06	35.05	30.71	23.03	27.53
D07	35.89	33.80	25.68	26.78
D08	23.56	26.75	18.32	16.98
D09	14.46	8.95	8.94	8.57
D10	25.39	26.89	17.14	18.66

TABLE IV: Weather Feature Influence

Dataset	***Source***	***Temperature***	***Humidity***	***Wind Speed***	***Precipitation***
D01	DWD	9.03 (0.18)	**8.57 (-0.28)**	**8.79 (-0.06)**	8.86 (0.01)
	Meteo	9.00 (0.16)	**8.63 (-0.21)**	**8.70 (-0.14)**	8.95 (0.11)
	NASA	9.64 (0.79)	9.27 (0.42)	9.46 (0.61)	9.47 (0.62)
D02	DWD	**19.04 (-0.33)**	19.54 (0.17)	19.66 (0.29)	**19.36 (-0.01)**
	Meteo	19.44 (0.07)	19.67 (0.30)	19.70 (0.33)	19.40 (0.03)
	NASA	**19.28 (-0.09)**	**19.31 (-0.06)**	19.72 (0.35)	19.40 (0.03)

C. Contribution of individual weather features

A closer look at the contributions of individual weather features is performed for datasets D01 and D02 - the only two datasets in which forecast results improved more than 1% by including a weather feature. Table IV analyzes SMAPE test values of forecasts with just one weather variable. All values refer to the minimum between XGBoost and GPR. The SMAPE difference to the prediction without weather features is given in brackets behind each value for increased visibility. Bold text implies an improvement by including the corresponding weather feature. Even the two most weather-influenced datasets show limited gains in prediction accuracy from integrating weather features. Dataset D01 shows a significant gain from including humidity, and a marginal gain from wind speed. In the case of D02, only for temperature can a notable performance gain be observed. Large variations are equally present for different weather sources, although this also varies on a case-by-case basis. Overall, no weather source is particularly suitable or unsuitable for load forecasting.

D. Alignment with existing research

Related work has no consensus on the integration of weather data, there are examples where no or little weather influence is measurable [7], [20], while in other examples [21], [22], a model enhancement is measurable. Many prior works however, consider only one or a few companies. Therefore, as a consequence of our simultaneous analysis of ten companies, the inconsistent improvements from integrating weather features are an expected result.

To summarize the results, the following four conclusions can be drawn: First, forecasting accuracy and weather influence on load forecasting are highly case-specific. Second, the impact of weather data on load forecasting tends to be weak for

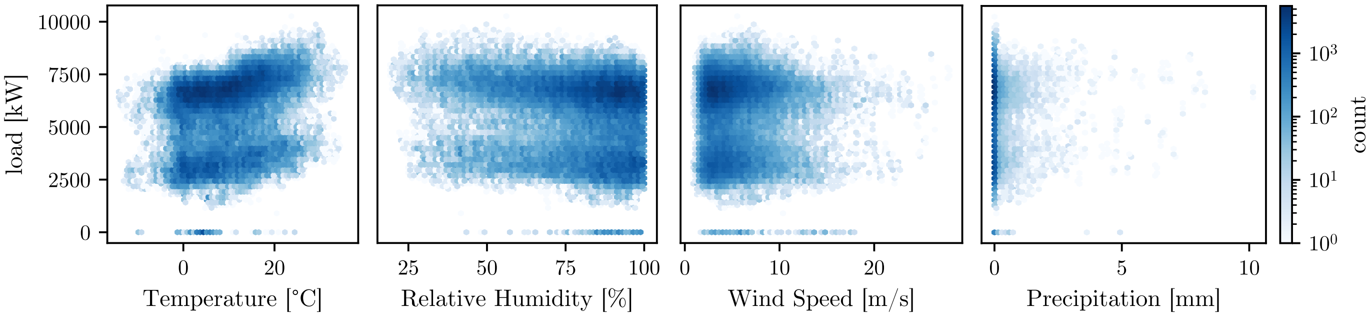

Fig. 2: 2D hexbin histogram of electrical load vs. DWD weather features for D09.

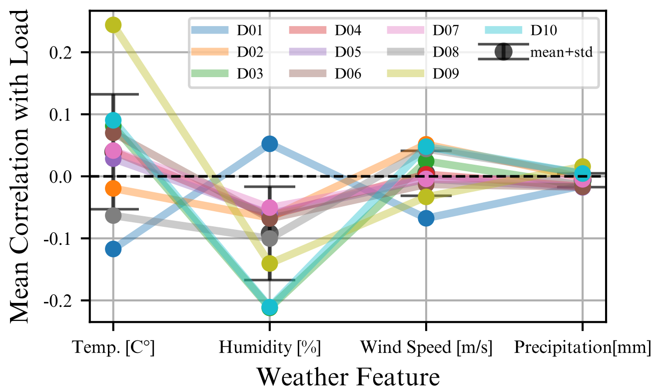

Fig. 3: Mean weather feature correlation coefficient for each dataset, and overall mean and standard deviation over all datasets.

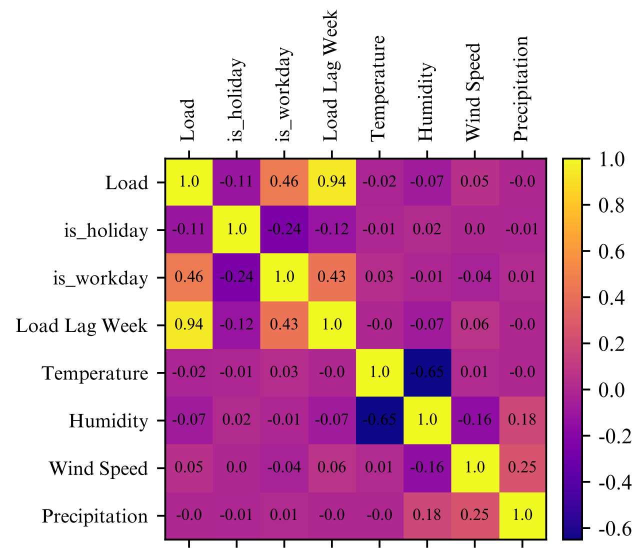

Fig. 4: Mean correlation matrix for D02, where the mean is taken over the three weather sources.

most companies in Germany, assuming no onsite generation. Notably, a correlation in the data does not necessarily translate to accuracy improvements from including that feature, likely as a result of other variables explaining the underlying relation better. For instance, calendar information may explain seasonal load variations more reliably than temperature. Third, where present, the influence of individual weather features is equally case-specific, with some features improving and others worsening forecast performance. Fourth, the weather sources perform similarly well, as their usefulness varies strongly from company to company.

The results thus have the following implications: A weather-based load forecast may not be advantageous in all cases. In cases where they are suitable, careful feature analysis and selection are required to achieve significant weather-related performance improvements. Further, the observed similar performance of different weather sources indicates that the choice of weather source may be secondary, and freely available data sources can be chosen.

IV. Limitations and Future Work

Despite the extensive investigations, some limitations could be addressed in future work. The analysis can be extended in many directions regarding the **scope** of the investigation:

Model selection can be supplemented by additional model classes such as (recurrent) neural networks (e.g., LSTMs, TCNs, transformers), statistical models (ARIMA), SVRs, or ensemble learners to allow more comprehensive statements across models. In addition, hyperparameter tuning can play an important role depending on the model class and has so far only been explored fundamentally Finally, the sample drawn does not allow **representative conclusions** about the entire German industry, as not all sectors are covered, and the sample size is rather small. A repetition of the study would be theoretically conceivable and useful. Still, due to the very limited availability of public industrial load profiles [9], it would only be possible to a limited extent.

Specifically, for industrial company load data, a few industrial processes likely make up a large portion of the load demand. The sheer scale of these processes can overshadow smaller, weather-dependent processes like cooling or heating. To tackle this issue, **decomposing the signal** into weatherindependent and weather-dependent parts could prove an effective approach, motivating more advanced methods like hierarchical load forecasting.

The weather data quality used in this study is better than that in a real-world application, as ground truth data has been used instead of forecasted values. However, this issue may be less severe than is often assumed, as other works [23] report only a 1% deterioration from using forecasts vs. measurements.

V. Conclusion

This work used electric load data of ten German companies without onsite generation to evaluate the impact of temperature, humidity, wind speed, and precipitation on day-ahead load forecasting. Extensive data exploration was performed before developing a forecasting pipeline, including an ablation study regarding weather features.

Several primary findings were made: First, the overall weather feature impact is low for most companies. The rarity of extreme weather events in Germany gives a possible explanation. Another important aspect is that seasonal patterns are already encoded in calendar information. This may prevent correlations between load and weather variables from translating to improvements in forecast accuracy. For some companies, the lack of weather influence may result from a few large consumers diffusing the signal of smaller, weather-dependent processes like cooling or heating. In these cases, hierarchical load forecasting may allow the successful use of weather information.

Second, the observed weather influence varies strongly, on a case-by-case basis - from company to company, between weather sources, and between individual weather variables. In particular, including all weather variables at once is not sufficient - instead, careful feature analysis and selection are necessary to obtain the best performance and prevent adding noise through unrelated weather features. Third, weather forecasting remains indispensable if there is local onsite generation, typically solar or wind.

Acknowledgment

The authors gratefully acknowledge the financial support of the *Kopernikus* project *SynErgie* by the Federal Ministry of Research, Technology, and Space of Germany (BMFTR), the project supervision by the project management organization *Projektträger Jülich* (PtJ), the technical discussion with C. Kaymakci and T. Kuhlmann, and the linguistic feedback from B. Spaeth.

References

[1] Maurizio Titz, Sebastian Pütz, and Dirk Witthaut. Identifying drivers and mitigators for congestion and redispatch in the german electric power system with explainable ai. *Applied Energy*, 356:122351, 2024.

[2] Ramón Christen, Luca Mazzola, Alexander Denzler, and Edy Portmann. Exogenous data for load forecasting: A review. *IJCCI*, pages 489–500, 2020.

[3] Maximilian Auffhammer and Erin T Mansur. Measuring climatic impacts on energy consumption: A review of the empirical literature. *Energy Economics*, 46:522–530, 2014.

[4] Kashif Hesham Khan, Caspar Ryan, and Ermyas Abebe. Optimizing hvac energy usage in industrial processes by scheduling based on weather data. *IEEE Access*, 5:11228–11235, 2017.

[5] Isaac Kofi Nti, Moses Teimeh, Owusu Nyarko-Boateng, and Adebayo Felix Adekoya. Electricity load forecasting: a systematic review. *Journal of Electrical Systems and Information Technology*, 7:1–19, 2020.

[6] Daniel L. Donaldson and Dilan Jayaweera. Integrating geospatial data for weather station selection and enhanced electric load forecasting. In *2022 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, 2022.

[7] Antonio Bracale, Guido Carpinelli, Pasquale De Falco, and Tao Hong. Short-term industrial load forecasting: A case study in an italian factory. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6. IEEE, 2017.

[8] Corentin Kuster, Yacine Rezgui, and Monjur Mourshed. Electrical load forecasting models: A critical systematic review. *Sustainable cities and society*, 35:257–270, 2017.

[9] Lukas Baur, Vignesh Chandramouli, and Alexander Sauer. Publicly available datasets for electric load forecasting–an overview. *ESSN: 2701-6277*, pages 1–12, 2024.

[10] Yuanyuan Wang, Jun Chen, Xiaoqiao Chen, Xiangjun Zeng, Yang Kong, Shanfeng Sun, Yongsheng Guo, and Ying Liu. Short-term load forecasting for industrial customers based on tcn-lightgbm. *IEEE Transactions on Power Systems*, 36(3):1984–1997, 2020.

[11] Stefan Ungureanu, Vasile Topa, and Andrei Cristinel Cziker. Deep learning for short-term load forecasting—industrial consumer case study. *Applied Sciences*, 11(21):10126, 2021.

[12] Meteomatics AG. Weather Parameters — Meteomatics — meteomatics.com. https://www.meteomatics.com/en/api/available-parameters/weather-parameter/, 2025. [Accessed 12-05-2025].

[13] Benjamin Gutzmann and Andreas Motl. wetterdienst, March 2025.

[14] National Aeronautics and Space Administration (NASA). NASA POWER — API Pages — power.larc.nasa.gov. https://power.larc.nasa.gov/api/pages/#/, 2025. [Accessed 12-05-2025].

[15] Paul Stackhouse. NASA POWER — Docs — Tutorials — API Getting Started - NASA POWER — Docs — power.larc.nasa.gov. https://power.larc.nasa.gov/docs/tutorials/api-getting-started/, 2025. [Accessed 12-05-2025].

[16] Rahul Gupta, Anil Kumar Yadav, SK Jha, and Pawan Kumar Pathak. Time series forecasting of solar power generation using facebook prophet and xg boost. In *2022 IEEE Delhi section conference (DELCON)*, pages 1–5. IEEE, 2022.

[17] Raza Abid Abbasi, Nadeem Javaid, Muhammad Nauman Javid Ghuman, Zahoor Ali Khan, Shujat Ur Rehman, and Amanullah. Short term load forecasting using xgboost. In *Web, artificial intelligence and network applications: proceedings of the workshops of the 33rd international conference on advanced information networking and applications (WAINA-2019) 33*, pages 1120–1131. Springer, 2019.

[18] Tianhong Liu, Haikun Wei, Sixing Liu, and Kanjian Zhang. Industrial time series forecasting based on improved gaussian process regression. *Soft Computing*, 24:15853–15869, 2020.

[19] Davide Chicco, Matthijs J Warrens, and Giuseppe Jurman. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *Peerj computer science*, 7:e623, 2021.

[20] John D Hobby and Gabriel H Tucci. Analysis of the residential, commercial and industrial electricity consumption. In *2011 IEEE PES Innovative Smart Grid Technologies*, pages 1–7. IEEE, 2011.

[21] Jingrui Xie, Ying Chen, Tao Hong, and Thomas D Laing. Relative humidity for load forecasting models. *IEEE Transactions on Smart Grid*, 9(1):191–198, 2016.

[22] Alfredo Candela Esclapez, Miguel López García, Sergio Valero Verdú, and Carolina Senabre Blanes. Automatic selection of temperature variables for short-term load forecasting. *Sustainability*, 14(20):13339, 2022.

[23] Damien Fay and John V Ringwood. On the influence of weather forecast errors in short-term load forecasting models. *IEEE transactions on power systems*, 25(3):1751–1758, 2010.