

Article

Deriving Input Variables through Applied Machine Learning for Short-Term Electric Load Forecasting in Eskilstuna, Sweden [†]

Pontus Netzell ^{1,*} , Hussain Kazmi ²  and Konstantinos Kyprianidis ¹ ¹ Future Energy Center, Mälardalen University, 722 20 Västerås, Sweden² Department of Electrical Engineering, KU Leuven, 3001 Leuven, Belgium

* Correspondence: pontus.netzell@gmail.com

[†] This paper is an extended version of our paper published in 64th International Conference of Scandinavian Simulation Society, SIMS 2023, Västerås, Sweden, 25–28 September 2023.

Abstract: As the demand for electricity, electrification, and renewable energy rises, accurate forecasting and flexible energy management become imperative. Distribution network operators face capacity limits set by regional grids, risking economic penalties if exceeded. This study examined data-driven approaches of load forecasting to address these challenges on a city scale through a use case study of Eskilstuna, Sweden. Multiple Linear Regression was used to model electric load data, identifying key calendar and meteorological variables through a rolling origin validation process, using three years of historical data. Despite its low cost, Multiple Linear Regression outperforms the more expensive non-linear Light Gradient Boosting Machine, and both outperform the “weekly Naïve” benchmark with a relative Root Mean Square Errors of 32–34% and 39–40%, respectively. Best-practice hyperparameter settings were derived, and they emphasize frequent re-training, maximizing the training data size, and setting a lag size larger than or equal to the forecast horizon for improved accuracy. Combining both models into an ensemble could enhance accuracy. This paper demonstrates that robust load forecasts can be achieved by leveraging domain knowledge and statistical analysis, utilizing readily available machine learning libraries. The methodology for achieving this is presented within the paper. These models have the potential for economic optimization and load-shifting strategies, offering valuable insights into sustainable energy management.



Citation: Netzell, P.; Kazmi, H.; Kyprianidis, K. Deriving Input Variables through Applied Machine Learning for Short-Term Electric Load Forecasting in Eskilstuna, Sweden. *Energies* **2024**, *17*, 2246. <https://doi.org/10.3390/en17102246>

Academic Editor: Ahmed Abu-Siada

Received: 8 April 2024

Revised: 29 April 2024

Accepted: 2 May 2024

Published: 7 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: short-term load forecasting; electrical grid; machine learning; multiple linear regression; light gradient boosting machine; explanatory variables

1. Introduction

A part of the solution to reach the global climate goals is to use renewable energy sources, which are volatile, intermittent, and non-dispatchable by nature [1]. This poses several questions about continued grid stability, and conventional power plants need to adapt to this reality by operating more flexibly, ramping up and down at a pace not traditionally seen [2]. Uncertainty and volatility in electricity production from variable renewable energy sources could be handled through demand response [3], e.g., the utilization of energy storage for load shifting [4].

In Sweden, the electrification of the transport and industry sectors is crucial for carbon emission reduction, leading to significant growth in the electricity demand. Two outstanding examples of industrial growth are the HYBRIT green steel project in the northern parts [5] and the southern Mälardalen region due to its dense population and the addition of new electricity intense industries. Electricity has traditionally been transferred through the national grid from northern hydropower plants and, more recently, from offshore wind turbines to the energy-intense southern half of the country. However, due to the rapid growth and end of life for several southern nuclear power plants, there are short-term issues in the transfer capabilities, meaning that the southern demand cannot be

sustainably fulfilled with northern electricity. Intense reinforcement and the expansion of the high-voltage grid may eventually make it possible to supply the additional demand. In the long-term, however, the increase in electricity usage in the northern areas could lead to a shortage of energy to transfer to the south. Therefore, there is a need for increased local production in the energy-intensive southern cities and regions for a robust and resilient local energy system [6].

In Eskilstuna, a city located in the Mälardalen region, the dispatchable local electricity production currently makes up a small portion of the total demand, and the rest is imported. The addition of several megawatt-size (MW-size) photovoltaic (PV) parks and a wind turbine park will increase the yearly energy self-usage ratio. However, it does not resolve the issue on an hourly and seasonal basis, as there is no substantial electricity generation from PVs in the evenings or during winter in Sweden. With the growing demand, electrification, and renewable proliferation, the necessity of being able to forecast future demand in combination with flexible energy usage is tangible.

Reliable forecasts can enable system operators and utilities to better manage the demand and supply balance in real time, as well as control energy storage units for shifting the load from high- to low-production periods, i.e., from day to night, or summer to winter. Use cases for forecasts range from long-term world trends and national changes to medium- and short-term changes on a regional or city-scale level [7]. Forecasting is essential for the energy and power sector, and the area has attracted attention for many decades, and with increasing computational power and new advanced models, the area continues to be focused on. Individual investigations are necessary, as each dataset is unique, and more complex models do not equal increased accuracy. Managing energy assets based on bad forecasts can lead to higher operating costs and, in a worst-case scenario, blackouts in the power grid.

Forecasting can be divided into three main parts using a system engineering perspective: *Input*, *Model*, and *Output* [8]. The size of historical data for training and the selection of both dependent and independent variables are examples of *Input*. If the data are disaggregated by geographical location, then hierarchical forecasting can be chosen as the *Model* technique [7]. Other *Model* variants are the selection of, e.g., non-linear or linear and black-box or non-black-box models, and their respective parameters. The predictions (*Output*) can be combined into ensembles, which is usually considered the best practice from an accuracy perspective [9]. The application of the forecasts matters; peak prediction generally demands an approach that is different from forecasts used for the operational optimization of energy units [10]. While numerous forecasting techniques have been proposed, there is no one-size-fits-all technique [7], and a detailed analysis of the specific case is needed for maximizing the forecast accuracy.

This paper focuses on forecasts and their usage on the urban electricity demand level in a city by studying the use case of the Eskilstuna Strängnäs Energi och Miljö (ESEM) electrical grid and energy system. Short-Term Load Forecasting (STLF) was applied to the hourly average electric load. This study aimed to create and explore a framework while analyzing and evaluating forecasting models and input variables. This study is summarized in three specific objectives:

1. Derive which calendar and meteorological variables improve the accuracy of the STLF in this city;
2. Evaluate the models: Multiple Linear Regression (MLR), Light Gradient Boosting Machine (LGBM), and the benchmark “weekly Naïve”;
3. Determine suitable hyperparameters for the models.

MLR and LGBM are compared to the benchmark “weekly Naïve” to determine whether advanced methods provide additional value compared to simpler benchmarks. The implementation of these forecasts for the control of energy storage units and other flexible assets is discussed, and the possible strengths and weaknesses of the two models are emphasized.

The rest of the paper is structured as follows: In Section 2: Methods, the data acquisition, algorithm creation, and model selection are presented. In Section 3: Results, the choice of explanatory variables and model results are presented, and these are criticized in Section 3: Discussion. This paper is concluded in Section 4: Conclusions.

2. Methods

Modeling and simulation, specifically Machine Learning (ML) in the context of the STLF of the electrical load, were the selected methods for this study. The focus lies on finding cause-and-effect relationships in the system under study to inductively generate and generalize knowledge. Knowledge is based on both an observable reality and empirically observable impressions, as discussions and guidance from domain experts are an important element of the progression. The selected methods are suitable for answering the research questions, as the data collected and created were quantitative. The approach is oriented toward problem solving, aiming to address the challenges in this specific context, i.e., the shortage of electricity and meeting the future electricity demands. However, studying this specific cross-sectional case will generate knowledge that can be transferred beyond its spatiotemporal boundaries. The detailed methodology of the data collection, model creation and execution, and validation and testing are as follows.

In general, the knowledge of the subject and the models is found through a literature review. Variables are added to the model, and the change in accuracy is tracked through a validation process. If a variable increases the model accuracy, it is kept, and disregarded if not. This validation is an iterative manual process and leads to a set of explanatory variables that are then tested against another unseen year of data. A manual process was selected over an automated hyperparameter tuning methodology due to the sheer size of the variables and to maintain control and a good understanding of the system dynamics. This study was finalized by analyzing and presenting the results in detail, and the conclusions were drawn from these.

2.1. Data Collection and Analysis

The dataset used herein comprised the hourly average electrical load in MW from 1 January 2020 to 31 October 2023 (3 years and 10 months) and was collected for all entry points (transformer stations) between the regional and the local grid in Eskilstuna, Sweden. Local electricity production, e.g., small-scale hydropower generation, was accounted for according to which transformer station they it was connected to. The summation of the seven transformer station loads, together with the generation from all large (63 and 80 Ampere) PV installations, made up the total energy usage of the city, denoted as the “total energy usage” in this paper. The electricity generation from smaller local PV installations, such as private households, was not included in the total energy usage.

To build an accurate forecast model, several meteorological and calendar explanatory variables needed to be evaluated in terms of correlation with the total energy usage. Some of the meteorological variables analyzed were reanalysis data of the wet and dry temperatures, wind speed, rain, and global irradiance from SMHI [11]. Other data were the relative and specific humidity from NASA [12]. Measured in situ, temperature from the central power plant was also used, including smoothed variants, i.e., moving averages with different window sizes. Cross effects are commonly used in forecasting and can be calculated by multiplying meteorological and calendar variables [13]. Degree days and hours for heating and cooling, which are the temperature differences below or above a certain threshold multiplied by time [14], are examples of cross effects used in this research. Additionally, historical regional loads and regional electricity prices were evaluated as explanatory variables, such as the actual load of electricity area SE3 of Sweden [15], as well as the actual production and forecasts of the regional solar power.

2.1.1. Correlation and Data Preparation

The correlation matrix of 27 interesting variables is shown in Figure 1, shortened from the original size of 70 to fit the paper. As an example, the correlation coefficient between the total energy usage and the reanalyzed dry temperature is -0.54 . Further, the forecasted and actual load for electricity area SE3 (from the ENTSO-E database [15]) of Sweden are highly correlated with the total energy usage of Eskilstuna and highly correlated with each other (correlation coefficient of 0.99). This correlation analysis is a part of the data preparation and analysis methodology employed to strengthen the selection of which explanatory variables to keep or not. As the correlation suggests similarities in the data, using several correlated variables as input to a model would mean an excess of information. Therefore, multicollinearity was considered during the entire study, and the final selection of the explanatory variables was conducted by avoiding collinear variables.

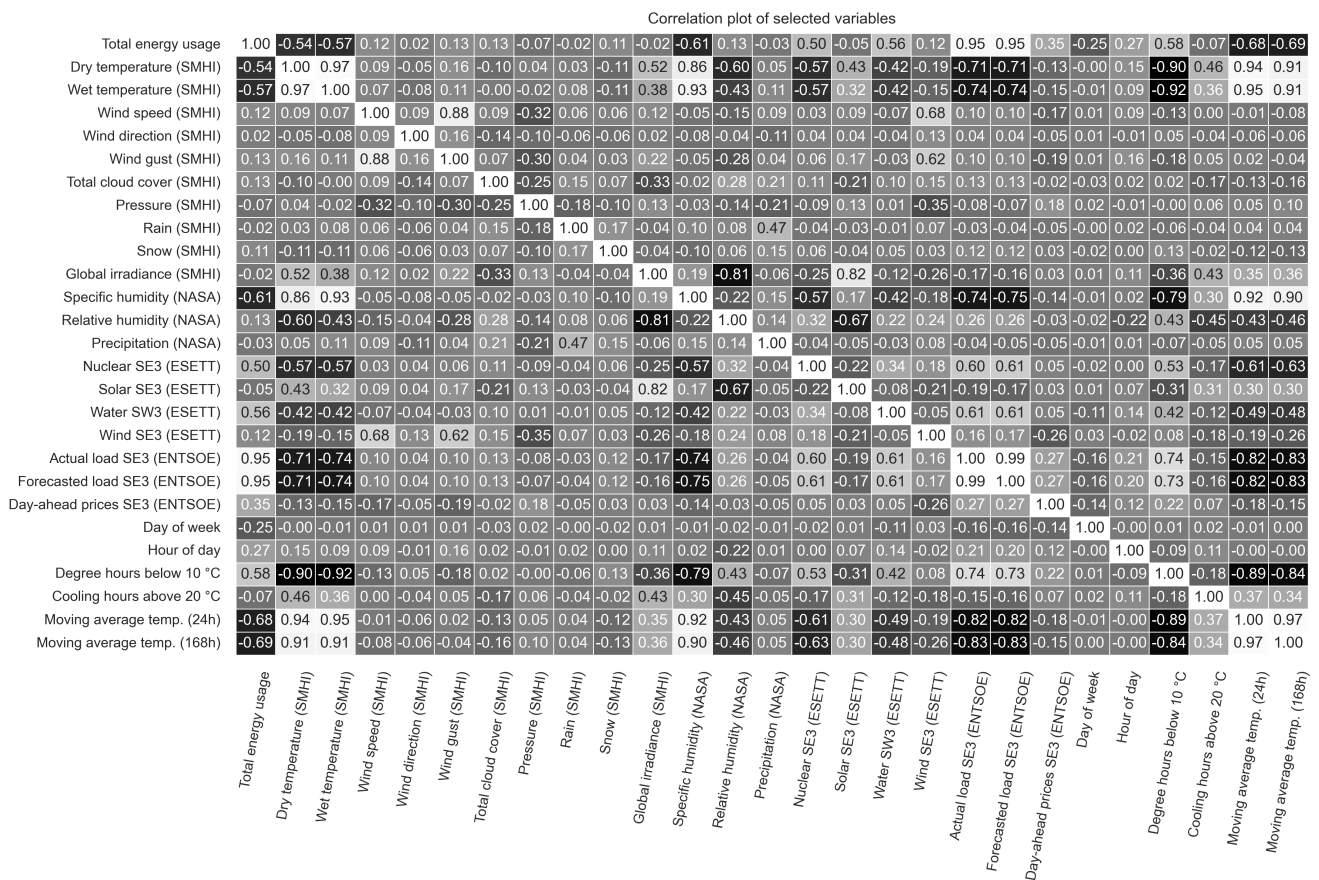


Figure 1. Correlation plot of 27 selected variables.

2.1.2. Load Decomposition

Public holidays are considered non-typical days [16], where the load is significantly lower. The additive decomposition of the trend, seasonal, and residual components [17] was applied using the Python library Statsmodels [18]. Similar to Işık et al. [19], the MSTL (Multi-Seasonal Trend decomposition using LOESS (Locally Estimated Scatterplot Smoothing)) revealed daily and weekly seasonality, as shown in Figure 2. Through this plot, the calendar variables Hour of day and Day of week were cemented as important explanatory variables for the total energy usage. The bottom residuals graph of Figure 2 shows a negative correlation between the MSTL residual and temperature for this winter example, which suggests that temperature is important for explaining the total energy usage. When not explained by temperature, large peaks in the residuals from the bottom

graph of Figure 2 can be explained with knowledge of public holidays (Christmas and New Year).

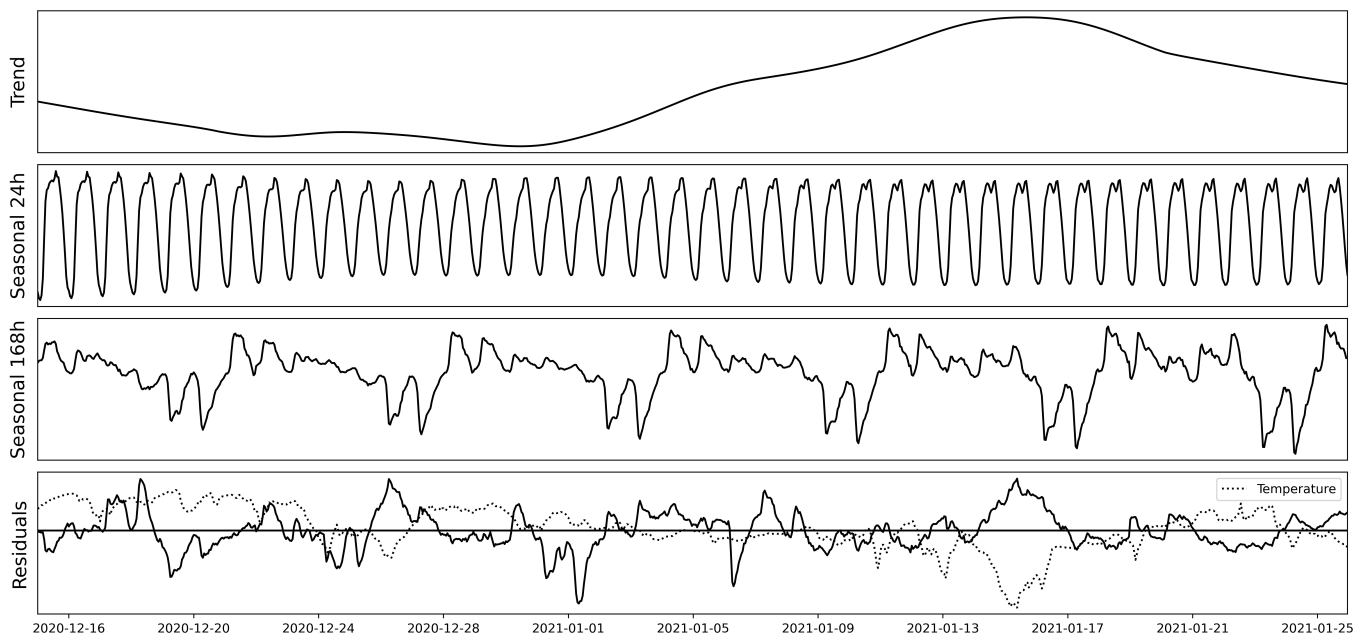


Figure 2. Additive load MSTL with seasonality periods of 24 and 168. Temperature is included in the bottom residual graph.

MSTL was applied, and the residuals were plotted against the outdoor temperature in Figure 3, with public holidays plotted separately. A portion of the residuals were significantly lower than the rest of the residuals during public holidays. Excluding public holidays and adding a LOESS line of the best fit gives a curve explaining how the residuals vary with temperature, depicted as “Smoothed” in Figure 3. This shows that residuals are negatively correlated with temperatures below 10 °C, while positively correlated with temperatures above 20 °C. This information was used to select and cement the existence of explanatory variables.

Further, there is a significant reduction in load due to the common industry practice of closing operations during the summer vacation period. The industry vacation period can be seen as a significant four-week lowering of the load in the trend of the MSTL plot for the summer (not shown in this paper). Therefore, a binary variable, which was set to zero for those four weeks, was added.

While load decomposition is not a novel concept [20–24], its application and number of publications has increased in the field of electric load forecasting in recent years. Decomposing the electrical load offers valuable insights and data explanations [25], which can prove beneficial for ML practitioners in STLF.

The MSTL algorithm, proposed by Bandara et al. [26], is an extension of the STL decomposition algorithm through extracting multiple seasonalities from a time series. While the utilization of MSTL has been documented in electric load forecasting [19,27–29], employing it to derive explanatory variables, especially the temperature thresholds for cooling and heating degree hours as presented in this manuscript, appears to be unprecedented, to the best of the author’s knowledge, and is therefore an important academic and practical contribution of this work.

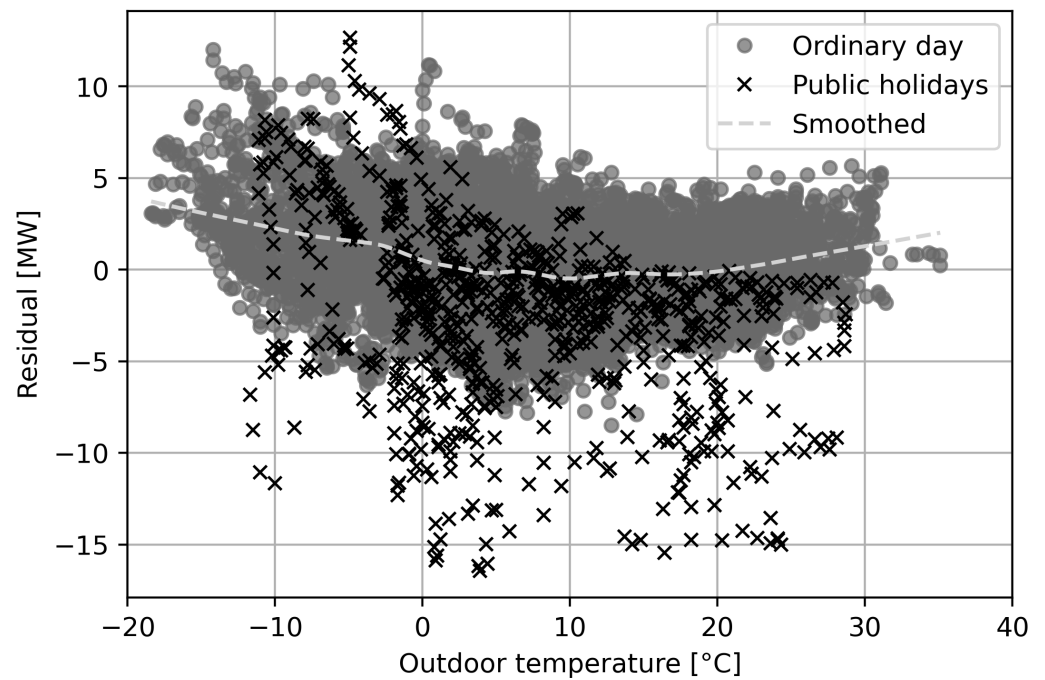


Figure 3. MSTL residuals vs. temperature.

2.1.3. Autocorrelation

Using the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) tools on the total energy usage revealed the temporal dependencies within this time-series data [18]. In Figure 4, the confidence interval is set to 99% ($\alpha = 0.01$), indicating a 99% confidence that the observed autocorrelation values fall within the confidence interval (gray area), and only a 1% chance of them falling outside, due to random chance alone. As most values fall outside the confidence interval, this suggests statistically significant autocorrelations. A seasonal pattern of 24 h is shown in the top ACF plot and possibly a 168 h pattern as well. In analyzing the bottom PACF plot, the first 3 h of lagged data are most significant, followed by the first 24 h, and then the data around multiples of 24 h. In addition, the peaks around 144 h and 168 h are prominent; therefore, lagged values of up to 168 h values are of interest for future analysis.

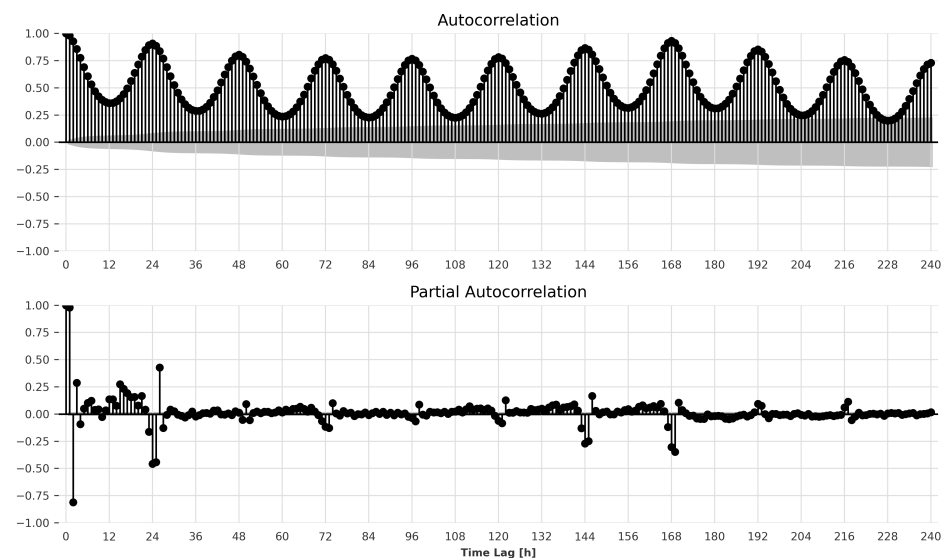


Figure 4. Autocorrelation and partial autocorrelation plot of the total energy usage.

2.2. Forecasting Models and Benchmark

The benchmark model was selected to be the well-known, in energy forecasting, “weekly Naïve” model (copy–paste the previous week’s values as the forecast for the next). It captures the weekly seasonality in the data and therefore outperforms the “daily Naïve” model [30]. A persistence-based benchmark, meaning one that is finding and copying days that are more similar than simply the weekly pattern, was used in a recent forecasting competition [31]. It can lead to a more accurate Naïve benchmark but at a higher cost of implementation and with reduced transferability to other cases; therefore, it was not selected in this study. The statistical modeling technique Multiple Linear Regression (MLR) is commonly utilized in electric load forecasting due to its ability to generate forecasts with minimal computational expense [32]. Equation (1) presents an example of MLR with two independent variables:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e, \quad (1)$$

where Y is the dependent variable, X_1 and X_2 are independent variables, β s are parameters to estimate, and e is the error term [13]. See Supapo et al. [33] for a more detailed explanation of MLR. Even though it cannot capture nonlinear relationships by definition, MLR is used because of its scalability and interpretability, while also achieving state-of-the-art performances in many cases. The ML Light Gradient Boosting Machine (LGBM) model, on the other hand, was highly represented in a recent energy predictor competition [34]. It is recognized as suitable for electric power modeling and explained in more detail in the open literature [35]. One obvious benefit of this model is that it can capture non-linear relationships while still remaining computationally feasible. Their proven track records and community support alongside their simplicity through implementations in open-source libraries, and compatibility (the use of the same past and future covariates), conclude that MLR and LGBM are suitable for this comparative study. The models are accessible and user-friendly, found in the “made for time series” Python library Darts [36], as utilized in this study. If not explicitly specified, the models used default settings.

2.3. Model Creation

The algorithm (referred to as *Historical Forecasts*), which is part of the framework used in this study, is depicted in Figure 5. First, necessary inputs are given to the algorithm: the forecast horizon, size of historical load for training, number of lagged (past) target values to use, how many hours to advance before making a new prediction, how many predictions to make before re-training, and when to stop. Future and past covariates including their lags can also be given to the model, e.g., temperature and day of the week. A prediction start date is given for splitting the data; otherwise, it will start as soon as possible given the size of the training and available dataset. The model was trained, and predictions were made according to the inputs, and at some points, it was re-trained. *Historical Forecasts* uses a rolling window approach for the rolling origin evaluation of the forecast [37]. Each prediction, error, and error metric are saved for further analysis.

2.3.1. Full-Year Run

The models underwent a rolling origin evaluation process using the *Historical Forecasts* algorithm (refer to Figure 5). This involves employing a forecast horizon of 168 h, advancing 17 h between each prediction, and re-training every 100th prediction, resulting in a total of 515 predictions. The forecast horizon was chosen to align with the available weather forecast horizon, thus setting the output chunk length equal to the forecast horizon. Additionally, the forward advancement between predictions was chosen as a prime number to minimize the chance of resonance with any of the seasonal patterns. Past lags for the load and lags for the future and past covariates were set to 168 h. The evaluated period was approximately 1 year (24 September 202X 22:00 to 27 September 202X 10:00), referred to as “the full-year run”. Including a year of training data and predicting a week ahead mean that the model has seen the predicted week once, but adding at least one more week to the training data means that the predicted week has been seen twice. The results from varying the training

size are presented in Section 3. The models were trained four times throughout the full-year run, as the load profile and temperature dependency were known to be different for the four seasons. For every full-year run, 86,520 errors (multiply the forecast horizon by the total number of predictions) were analyzed, together with 515 average errors (one for each prediction made), and a single average error. Which error metric to be used for different datasets can be derived from Hewamalage et al. [37]. The Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) were concluded to be the two most common ones used for the STLF of an electrical load in Nti et al. [38], and the former was used in this study and presented in the MW. The relative RMSE (rRMSE), defined as the RMSE of the MLR and LGBM, respectively, divided by RMSE for the “weekly Naïve” is used to quantify the performance against that of the benchmark.

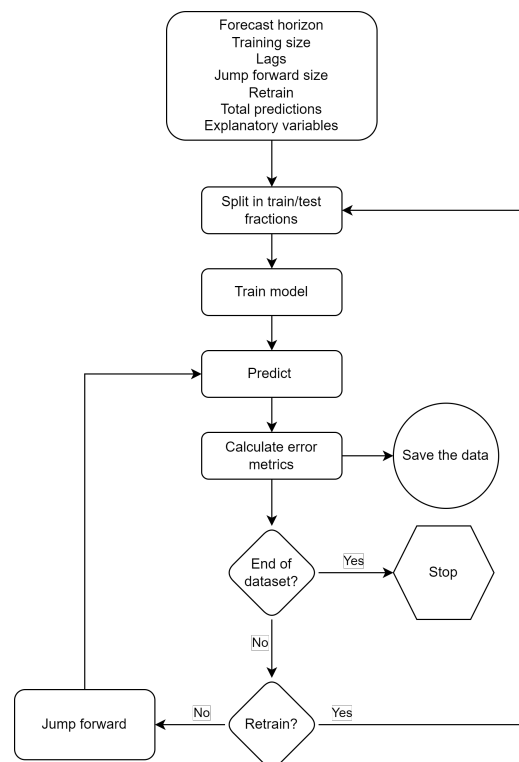


Figure 5. Flowchart of the algorithm Historical Forecasts created for model validation.

2.3.2. Validation

An extensive rolling origin validation was performed, where the least computationally expensive model MLR was used for running hundreds of full-year runs, each generating errors that were manually and iteratively compared. Periods with the largest errors, such as public holidays, the yearly peak, and the summer period, were focused on in greater detail. One explanatory variable after the other was manually added, the model parameters were varied, and the results were evaluated. Through combinations of visual inspections of the animations and plotting the model errors in graphs, the key explanatory variables were concluded. When no significant improvement was achieved with this semi-structured scrutiny of the MLR, the analysis was stopped. The same analysis was not performed with the LGBM due to the computational expense, where only a few selected parameter changes were made to verify the model behavior, e.g., reducing the training size reduces the accuracy.

The rolling origin validation was performed on two years of data, one referred to as 2021–2022 and another as 2022–2023, to determine the selection of explanatory variables and hyperparameters. This usage of two sets of data helped determine whether the explanatory variables and hyperparameters actually enhanced the accuracy consistently over the two years, or if the observed improvements were unique to the first year only.

The starting point for the full-year run during these two validations differed by one year, i.e., 24 September 2021 22:00:00 for 2021–2022 and 24 September 2022 22:00:00 for 2022–2023. The same set of explanatory variables, as concluded useful for the model during the first year, were tested in the second year, and the resulting error metrics were compared.

In addition, the list of explanatory variables was expanded during the process as more knowledge was gained. Several databases were included and tested with the same methodology.

3. Results and Discussion

3.1. Explanatory Variables

Several explanatory variables for STLF in the specific context of this use case were evaluated in terms of usefulness by applying a rolling origin validation (the full-year run). Two separate years of data were used: 2021–2022 and 2022–2023. The change in accuracy of the models when each of the explanatory variables was consecutively added is documented in Table 1. The results for the 2021–2022 year were gathered from previously published work [39].

Table 1. RMSEs when consecutively adding each explanatory variable to the models in a full-year run. The respective best-performing results for a run are underlined, and the subsequent threshold RMSE is framed. Meteorological variables were taken from SMHI [11].

Explanatory Variable	MLR		LGBM	
	2021–2022	2022–2023	2021–2022	2022–2023
Day of week [0–6]	4.54	4.31	4.74	4.59
Hour of day [0–23]	4.56	4.30	4.72	4.56
Holidays [0 OR 1]	4.44	4.12	4.53	4.30
Industry vacation [0 OR 1]	4.31	4.00	4.35	4.18
Heating hours [Kh, <10 °C] ¹	2.20	<u>2.18</u>	2.89	2.99
Global irradiance [W m ^{−2}]	2.12	2.47	2.68	3.00
Cooling hours [Kh, >20 °C] ²	2.06	2.34	<u>2.65</u>	<u>2.89</u>
Wind speed [m s ^{−1}]	<u>2.04</u>	2.41	2.71	2.93

¹ Threshold MLR. ² Threshold LGBM.

Concluded from Table 1, the best set of explanatory variables for the MLR model (only using the SMHI [11] database) is the first five: day of week, hour of day, holidays, industry vacation, and heating hours. Only these five variables improved the model accuracy for both simulated years. The resulting RMSEs of 2.20 (2021–2022) and 2.18 (2022–2023) were used as the thresholds for the rest of the analysis. Several additional databases and explanatory variables were tested against the best-performing models, and these are discussed subsequently, starting with the MLR model.

3.1.1. MLR Additional Databases and Explanatory Variables

The resulting RMSEs when adding a sixth explanatory variable for 2021–2022 are presented in Table 2. Explanatory variables related to wind improved the model for this year. Wind speed also improves the accuracy in the first year (2021–2022), as shown in Table 1. On the contrary, the accuracy worsened when included in the subsequent year (2022–2023). The explanatory variables were only kept if validated as accuracy-improving for both years.

Table 2. RMSEs when adding a sixth explanatory variable to the MLR full-year run 2021–2022, sorted from lowest to highest. The line separates the accuracy-improving variables *.

Variable Added	RMSE	Variable Added	RMSE
Actual load SE3 (ENTSOE)	1.87	Pressure (SMHI)	2.27
Forecasted load SE3 (ENTSOE)	2.14	Dew point temperature (NASA)	2.28
DA wind forecast SE3 (ENTSOE)	2.19	Rain (SMHI)	2.28
Wind SE3 (ESETT)	2.19	Snow (SMHI)	2.28
Wind onshore SE3 (ENTSOE)	2.20	Wind direction (SMHI)	2.28
Gust (SMHI)	2.20	Wind direction 10 m (NASA)	2.29
Wind speed 10 m (NASA)	2.22	Wind direction 50 m (NASA)	2.29
Wind speed 50 m (NASA)	2.22	Wet temperature (NASA)	2.30
Solar SE3 (ESETT)	2.23	Precipitation (NASA)	2.31
DA Wind and Solar SE3 (ENTSOE)	2.24	Total cloud cover (SMHI)	2.34
Relative humidity (NASA)	2.24	Temperature (NASA)	2.35
Nuclear SE3 (ESETT)	2.25	Water SE3 (ESETT)	2.35
Wet temperature (SMHI)	2.25	Unspecified production SE3 (ESETT)	2.41
Relative humidity (NASA)	2.26	CHP production SE3 (ESETT)	2.43
Pressure (NASA)	2.27	Day-ahead prices (ENTSOE)	2.72

* Threshold RMSE = 2.20.

The fact that the explanatory variables are under the threshold does not mean that they are useless for the models when, e.g., running them in a real-time or daily/weekly application. Day-ahead prices are an example where consumers change their consumption patterns depending on the price. Logically, electrical vehicles are likely to be charged during the night when the prices are lower, and electric heating systems could be temporarily shut down during peak prices, which would affect the electrical load on the grid. Still, the day-ahead prices significantly worsen the accuracy. The methodology applied in this study and the results presented in, e.g., Table 2, do not mean that all variables under the threshold RMSEs are useless for the models. The results only mean that these variables are not improving the accuracy in a full-year run performed in this context. Table 2 shows that 18 out of 24 such variables result in errors that are within a 5% difference compared to the threshold. Consequently, they are, on average, not improving the accuracy, but they could theoretically be during specific periods, significantly reducing the error and reducing costs. In addition, periods that are hard to predict can lead to high imbalance costs. Therefore, the inclusion of such variables could be justified economically if the worsening of accuracy comes with no significant cost. To summarize, all explanatory variables could be useful for the models during certain periods. To determine which, and their level of usefulness, they can benefit from being evaluated over a shorter period than a year. An economic assessment where forecasts are used together with, e.g., an electrical battery, can help derive more details about each explanatory variable.

In Table 3, the resulting RMSEs when adding a sixth explanatory variable for 2022–2023 are presented. The actual load of the third electrical area of Sweden (SE3) significantly improved the accuracy for both years. The forecasted load (SE3) did as well, although not to the same extent. Both the actual and forecasted load were gathered from the ENTSO-E database [15]. The correlation coefficient between the actual and the forecasted load is 0.99, as shown in Figure 1. They both correlate with the total energy usage with a coefficient of 0.95. Using the actual load as input means having perfect foresight, which is impossible for real-time predictions. Instead, the forecasted load needs to be used. These results show that using the forecasted load would have meant an accuracy improvement, historically, and could therefore be used as input for real-time STLF in this specific use case. For the rest of the study, however, the actual load and the predetermined five explanatory variables were used in the detailed analysis of the model errors.

Table 3. RMSEs when adding a sixth variable to the MLR full-year run 2022–2023, sorted from lowest to highest. The line separates the accuracy-improving variables *.

Variable Added	RMSE	Variable Added	RMSE
Actual load SE3 (ENTSOE)	1.92	Wind direction (SMHI)	2.28
Forecasted load SE3 (ENTSOE)	2.05	Pressure (SMHI)	2.29
Precipitation (NASA)	2.16	Pressure (NASA)	2.29
Snow (SMHI)	2.25	Wind speed 10 m (NASA)	2.29
Unspecified production SE3 (ESETT)	2.25	DA Wind and Solar SE3 (ENTSOE)	2.31
CHP production SE3 (ESETT)	2.25	Gust (SMHI)	2.32
Wind direction 50 m (NASA)	2.26	Nuclear SE3 (ESETT)	2.33
Wind direction 10 m (NASA)	2.26	Wind speed 50 m (NASA)	2.33
Wet temperature (SMHI)	2.26	Water SE3 (ESETT)	2.33
Relative humidity (NASA)	2.26	Wind onshore SE3 (ENTSOE)	2.41
Dew point temperature (NASA)	2.27	DA wind forecast SE3 (ENTSOE)	2.41
Rain (SMHI)	2.27	Wind SE3 (ESETT)	2.41
Day-ahead prices (ENTSOE)	2.27	Total cloud cover (SMHI)	2.45
Wet temperature (NASA)	2.27	Solar SE3 (ESETT)	2.50
Temperature (NASA)	2.27	Relative humidity (NASA)	2.67

* Threshold RMSE = 2.18.

3.1.2. LGBM Additional Databases and Explanatory Variables

Concluded from Table 1, the best set of explanatory variables for the LGBM model (only using the SMHI [11] database) is the first seven: day of week, hour of day, holidays, industry vacation, heating hours, global irradiance, and cooling hours. These variables improved the model accuracy for both simulated years. The resulting RMSEs of 2.65 (2021–2022) and 2.89 (2022–2023) were used as the thresholds for the rest of the analysis. Several additional databases and explanatory variables were tested against the best-performing models and are discussed subsequently.

In Table 4, the resulting RMSEs when adding an eight explanatory variable for 2021–2022 are presented. The actual and forecasted loads were the only two additional explanatory variables that improved the accuracy.

Table 4. RMSEs when adding an eight variable to the LGBM full-year run 2021–2022, sorted from lowest to highest. The line separates the accuracy-improving variables *.

Variable Added	RMSE	Variable Added	RMSE
Actual load SE3 (ENTSOE)	2.34	Wind speed 50 m (NASA)	2.74
Forecasted load SE3 (ENTSOE)	2.56	Wind speed 10 m (NASA)	2.75
Rain (SMHI)	2.65	Solar SE3 (ESETT)	2.76
Snow (SMHI)	2.65	Wind SE3 (ESETT)	2.77
Wet temperature (SMHI)	2.67	Wind direction 10 m (NASA)	2.77
Wet temperature (NASA)	2.70	Wind direction 50 m (NASA)	2.77
Relative humidity (NASA)	2.71	DA Wind and Solar SE3 (ENTSOE)	2.78
Total cloud cover (SMHI)	2.72	Wind onshore SE3 (ENTSOE)	2.78
Dew point temperature (NASA)	2.72	DA wind forecast SE3 (ENTSOE)	2.79
Temperature (NASA)	2.72	Pressure (NASA)	2.79
Relative humidity (NASA)	2.72	Pressure (SMHI)	2.79
Gust (SMHI)	2.72	Nuclear SE3 (ESETT)	2.88
Precipitation (NASA)	2.73	CHP production SE3 (ESETT)	2.98
Wind direction (SMHI)	2.73	Water SE3 (ESETT)	3.05
Unspecified production SE3 (ESETT)	2.74	Day-ahead prices (ENTSOE)	3.30

* Threshold RMSE = 2.65.

In Table 5, the resulting RMSEs when adding an eight explanatory variable for 2022–2023 are presented. Similarly to Table 4, both the forecasted and actual loads improve the model accuracy.

Table 5. RMSEs when adding an eight variable to the LGBM full-year run 2022–2023, sorted from lowest to highest. The line separates the accuracy-improving variables *.

Variable Added	RMSE	Variable Added	RMSE
Actual load SE3 (ENTSOE)	2.22	Relative humidity (NASA)	2.96
Forecasted load SE3 (ENTSOE)	2.28	Wind speed 10 m (NASA)	2.96
Solar SE3 (ESETT)	2.74	Pressure (SMHI)	2.98
Rain (SMHI)	2.89	Wind onshore SE3 (ENTSOE)	2.98
Wet temperature (SMHI)	2.89	Pressure (NASA)	2.98
Snow (SMHI)	2.90	Wind SE3 (ESETT)	2.98
DA Wind and Solar SE3 (ENTSOE)	2.91	Wind direction (SMHI)	2.98
Total cloud cover (SMHI)	2.92	DA wind forecast SE3 (ENTSOE)	2.98
Temperature (NASA)	2.92	Day-ahead prices (ENTSOE)	2.98
Precipitation (NASA)	2.92	Dew point temperature (NASA)	2.99
CHP production SE3 (ESETT)	2.92	Wind direction 10 m (NASA)	3.00
Nuclear SE3 (ESETT)	2.94	Wind direction 50 m (NASA)	3.00
Gust (SMHI)	2.94	Relative humidity (NASA)	3.00
Wet temperature (NASA)	2.94	Unspecified production SE3 (ESETT)	3.04
Wind speed 50 m (NASA)	2.96	Water SE3 (ESETT)	3.27

* Threshold RMSE = 2.89.

In summary, in first utilizing only a single database of meteorological variables, five explanatory variables were concluded useful for the MLR model, and seven for the LGBM model. The prerequisite for this conclusion was that the variables shall improve the accuracy for both simulated years. In including additional databases, a sixth and, respectively, eighth explanatory variable (the actual load of SE3) were concluded useful. The forecasted load of SE3 was also concluded as useful for the models. The forecasted load is the explanatory variable that needs to be used in a real-time application.

3.2. Detailed MLR, LGBM, and “Weekly Naïve” Full-Year Run Results

The full-year run results for comparing MLR with the LGBM, including the Naïve benchmark, are shown in Figure 6. The best-performing MLR model from Table 2, the best-performing LGBM model from Table 4, and the “weekly Naïve” model’s results are plotted in Figure 6a. The MLR achieved an average RMSE of 1.86 (rRMSE of 32%), and the LGBM achieved an average RMSE of 2.33 (rRMSE of 40%) compared to 5.86 for the “weekly Naïve” model.

The errors of the “weekly Naïve” model depend on both calendar and temperature factors. The Swedish public holidays (Christmas, New Year, and Easter) and the local industry vacation period in June–July give rise to the largest errors as the load is significantly lower during public holidays. When temperature changes from one week to the other during the heating period, the “weekly Naïve” model’s performance is affected, which is expected when copying the previous week’s load as a forecast for the upcoming week. Holidays are still a large source of error in the best-performing MLR and LGBM models, but with the addition of the actual load (SE3) as an explanatory variable, these errors are reduced.

To further increase model accuracy, a detailed simulation using an hourly error scale and analyzing shorter periods (weeks or months) at a time would be useful. By doing so, these periods of interest (e.g., the yearly peak) and periods with high errors (e.g., holidays, industry vacation) can be targeted separately, and the same methodology as applied for the full-year run can be used for evaluating all explanatory variables again. Although such an analysis is an important tool and was used in this study for tuning the hyperparameter and evaluating explanatory variables, the aim was to provide recommendations on a yearly scale.

The best-performing MLR model from Table 3, the best-performing LGBM model from Table 5, and the “weekly Naïve” results are plotted in Figure 6b. The MLR achieves an average RMSE of 1.91 (rRMSE of 34%), and the LGBM achieves an average RMSE of 2.33 (rRMSE of 39%) compared to 5.65 for the “weekly Naïve” model.

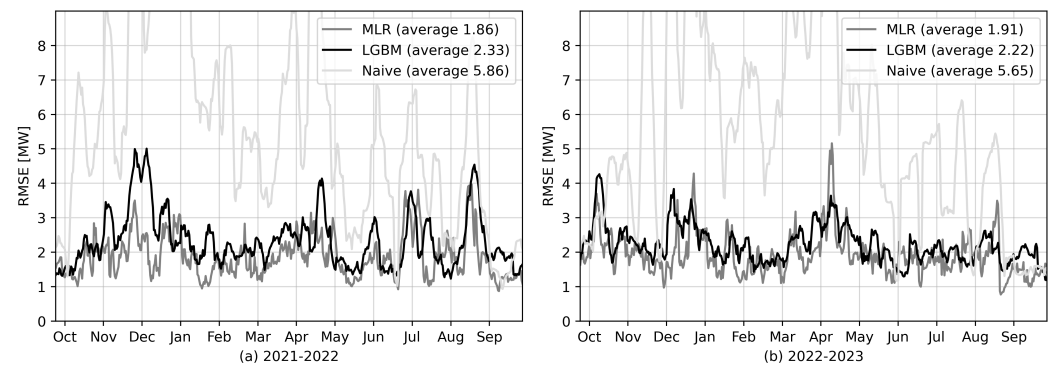


Figure 6. MLR, LGBM, and “weekly Naïve” model’s ((a) 2021–2022 and (b) 2022–2023) full-year run model errors.

Detailed Model Error Analysis

All 86,520 errors for the full-year runs of the two best-performing models are plotted in histograms, as seen on the left in Figure 7a,c. Both models produce errors close to a normal distribution centered close to zero. The centers of the distributions are slightly tilted toward a positive number for MLR (average 0.45), and a negative number for LGBM (average -0.16).

The autocorrelation plots, on the right in Figure 7b,d, are shown for the same prediction, i.e., the 91st out of the 515 made in each full-year run. The confidence interval (alpha) was set to 0.99. The ACF plots show different shapes, different seasonalities, over this 168 h forecast horizon. First, they show that most of the past (lagged) errors are not significantly autocorrelated, except for the first 3–10 errors. This is concluded as serial autocorrelation, meaning that if the model is wrong in one direction for the first time step, it will likely be wrong in the same direction in the next step. Second, seasonal autocorrelation is also observed, meaning that if the prediction is too low one day, it is likely to be too low on the following day, in a seasonal pattern. Third, these plots highlight that both forecast models produce different errors from an autocorrelation perspective and are therefore suitable for combination.

The errors from the two best-performing models are plotted against each other in Figure 8, and they show a weak correlation. This, together with the distribution of the errors in Figure 7, shows good potential for combining them in an ensemble model. Accuracy improvements are expected when combining models according to the literature [9] and forecasting competitions [34], but a deeper analysis of this specific use case is needed.

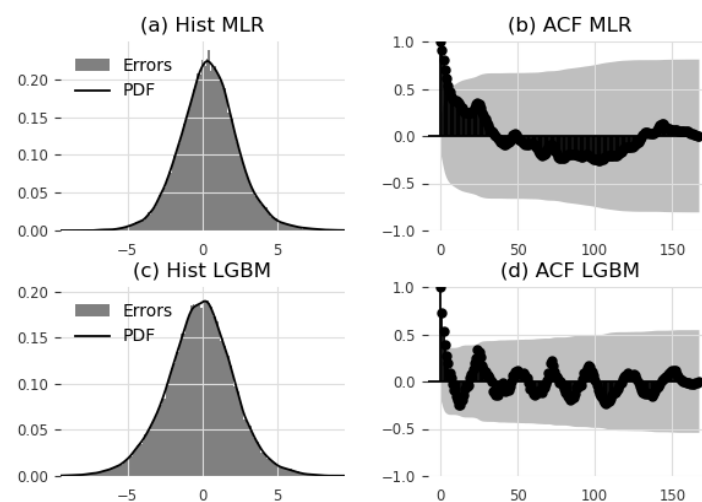


Figure 7. Probability density functions (a,c), and autocorrelation Functions (b,d) of the errors for the best-performing full-year runs.

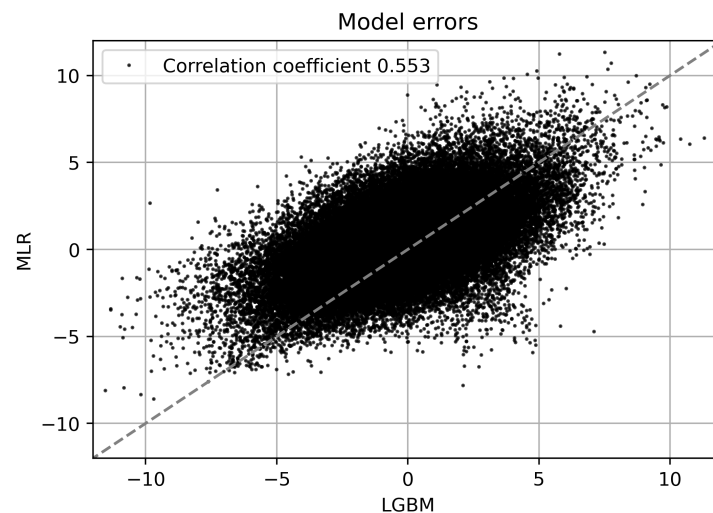


Figure 8. MLR vs. LGBM full-year run model errors.

3.3. Hyperparameter Tuning

3.3.1. Lag Tuning

The effect of varying the size of the included past lags (previous values of the variable) of the load itself, as well as the size of both past and future lags of the explanatory variables, was analyzed. The effect for the MLR model is shown in Figure 9a, and similarly, that for the LGBM is shown in Figure 9b. The forecast horizon determines how many lags need to be included for the best accuracy. Including fewer lags than the forecast horizon significantly reduces the accuracy. Likewise, including more lags than the forecast horizon also affects the accuracy, but not to the same extent. Figure 9 shows that the number of lags should be equal to the forecast horizon while the forecast horizon is 168 h or larger. On the contrary, when the forecast horizon is 72 h, the number of lags that should be included is approximately 120 h, which is valid for both MLR and LGBM. During this analysis, no difference was made between the lag size of each explanatory variable or no difference between past and future lags. All past and future lags, for all variables were set to the same size. An additional detailed analysis for cherry picking only the most important lags, in accordance with the partial autocorrelation plot in Figure 4, is suggested as future work for the re-evaluation and continuous improvement of the models.

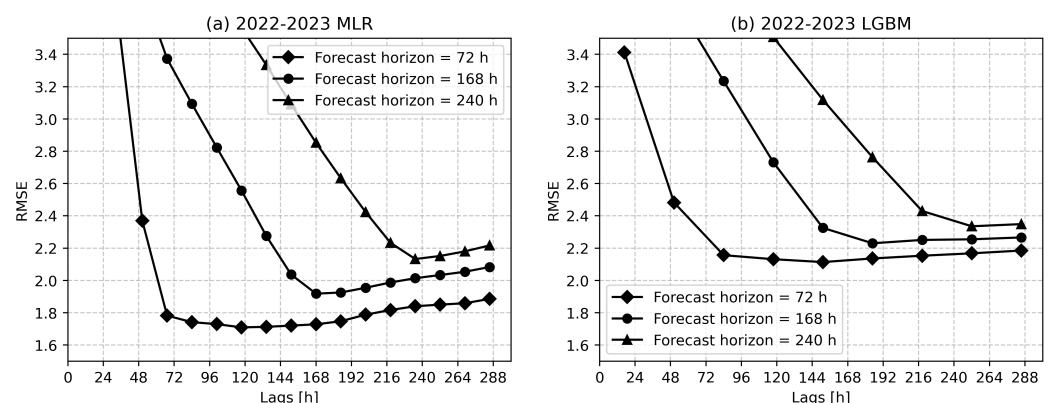


Figure 9. Varying forecast horizon and lag size versus RMSE for the full-year run: (a) MLR and (b) LGBM.

3.3.2. Training Size Tuning

How the error varies when varying the training size is shown in Figure 10. In general, the RMSEs for the full-year runs decrease when the training size increases. The error

is significantly improved when increasing the training size up to 0.8 years, as shown in Figure 10a,c. The curve flattens out to then improve again around 2 years of training data, as shown in Figure 10b,d. The improvement in the full-year run accuracy for the MLR model when including 2 years instead of 1.2 is 5.7%, while it is 7.5% for the LGBM. These graphs show that the training data should be set in intervals of full years. Nonetheless, an increased training size will increase the computational effort and time.

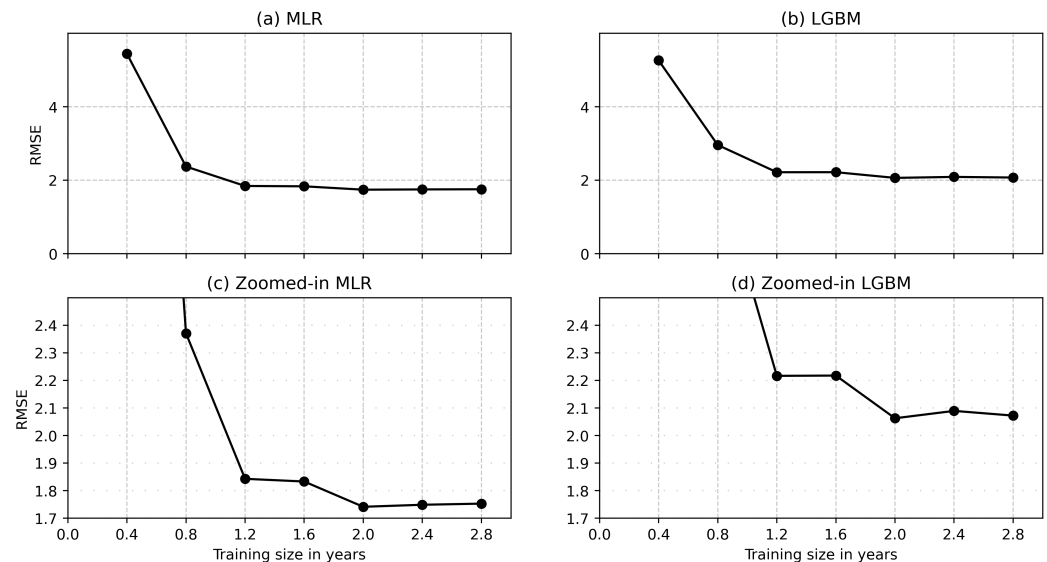


Figure 10. Varying training sizes for both full-year runs: (a) MLR 2021–2022, (b) MLR 2022–2023 (c) MLR 2021–2022 zoomed in, and (d) MLR 2022–2023 zoomed in.

3.3.3. Re-Training Interval Tuning

In Figure 11a, the impact of re-training on the full-year run accuracy is presented. The re-training was tested for the entire dataset. Re-training approximately every week lead to a 5% accuracy improvement for the MLR model compared to not re-training at all over the 2.5 years. For the LGBM in Figure 11b, the same applies but with an 18% improvement. Re-training affects the computational effort.

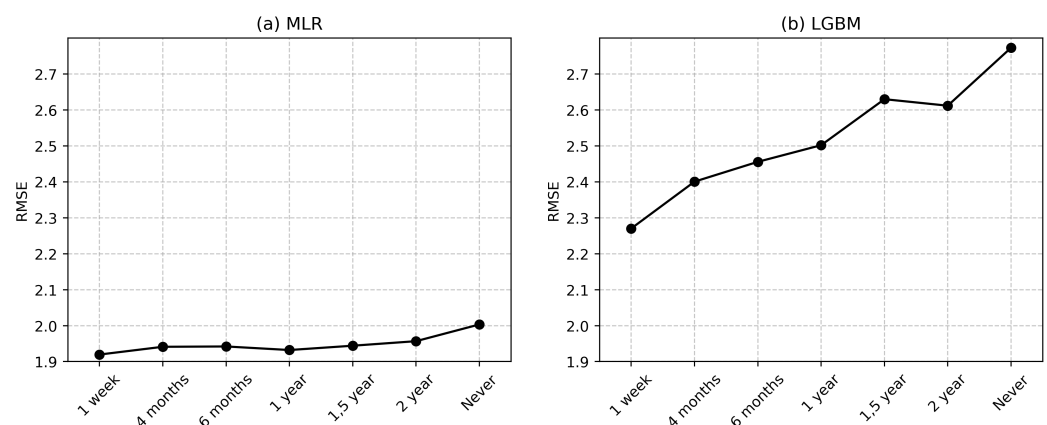


Figure 11. RMSE when varying the re-training interval, (a) MLR and (b) LGBM.

To conclude the remarks on computational effort, the MLR is the winner in terms of computation. It takes about 60 times more time for the LGBM to finish the full-year run. If these models are used in a live decision-making process, e.g., an economic model for the predictive control of grid-scale electrical energy storage, the time constraint is an hour. To clarify, the electrical price settlement period is an hour in Sweden, meaning that the load

prediction needs to be calculated in less than an hour. However, this is about to change as the countries connected to NordPool are changing to a 15 min resolution. In addition, intra-day trading and imbalance markets can be conducted on a smaller timescale, meaning that computational efficiency is important. The LGBM model was trained on three years of data and made a single 168 h forecast, which took about 24 min. This was performed on a laptop with a 2.4 GHz processor and using 16 GB of memory. This means that it is not infeasible that the worst-case scenario model could converge within the limited time of 15 min if run on a higher-performance computer or with several threads. The MLR finishes the same worst-case scenario in 7 s.

3.4. Navigating Future Challenges

The landscape of the electrical grid is rapidly evolving, driven by shifting usage patterns, price fluctuations, and the integration of intermittent renewable energy sources. During this evolution, the adaptability of forecasting models emerges as a critical factor, particularly underscored during the COVID-19 period [31]. The emergence of MW-sized PV and wind parks introduces a dynamic where historical data become less relevant, challenging the usefulness of traditional forecasting approaches. While this study did not directly assess adaptability or robustness, it highlights the necessity for such considerations.

Before using these forecasts in practical applications, such as energy storage planning or control, incorporating an economic dimension through simulation is important. Underestimating the annual peak on the grid can result in significant economic penalties. The expected errors when using these forecasting models are presented in Figures 7a,c, and 8. This level of accuracy may or may not be acceptable, depending on the limit and the size and availability of the electrical energy storage or other flexible energy assets. Despite achieving an accuracy comparable to or better than the overall Swedish national load forecast produced by Svenska kraftnät [40], internal discussions with the local grid operator of Eskilstuna indicate the need for improved precision, especially in anticipation of future energy storage investments.

Moving forward, the integration of forecasting methods with energy storage management and sizing presents a path for further research in this specific use case. Customizing forecasting techniques based on the locations of energy storage systems and detailed information about PV parks holds promise for achieving high-accuracy forecasts [7]. Aligning these forecasts with energy storage management and sizing considerations represents a likely progression for advancing this study's insights.

4. Conclusions

This study explored a framework for analyzing and evaluating forecasting models and explanatory variables. The performances of two models, MLR and LGBM, were assessed using a dataset from the local grid operator of Eskilstuna. Several sets of possible explanatory variables from several open databases were evaluated through a rolling origin validation process. In the same way, model variables were evaluated, with the inclusion of calendar and meteorological variables, industry vacation periods, and cross-effect variables like heating hours below 10 °C. Although public holidays and non-typical periods still contributed to the largest errors, the addition of binary explanatory variables for these factors significantly improved the accuracy. The methodology of using MSTL for deriving explanatory variables, especially the thresholds of the degree hours for heating and cooling, is considered an academic contribution. In parallel, apart from the detailed methodology, incorporating the electric load forecast provided by the Swedish transmission system operator for the SE3 area of Sweden as an explanatory variable notably enhanced the model accuracy, particularly during non-typical periods, and this is considered a practical contribution to the literature.

The best-performing MLR and LGBM models outperformed the “weekly Naïve” benchmark model with rRMSEs of 32–34% and 39–40%, respectively. During the detailed model error analysis, the MLR and LGBM showed varying characteristics and weak

correlations, meaning that they are suitable for combination. However, MLR consistently demonstrated lower errors compared to the more computationally expensive LGBM.

Furthermore, this study demonstrates that larger dataset sizes for training lead to an improved accuracy. This suggests that the inclusion of past and future lags of explanatory variables and past loads should match the forecast horizon when it exceeds 168 h. On the other hand, if the forecast horizon is less than 168 h, lags should be larger than the forecast horizon. Additionally, the findings in this study show the importance of re-training as frequently as possible.

As concluding remarks and in considering future directions, there are several potential paths forward. One possibility is to integrate an economic dimension into the evaluation process by employing the forecasts as control input for electrical energy storage. This approach can aid in determining both a satisfactory level of accuracy and the associated cost of forecasting errors. Additionally, further extensions of this study could involve exploring hierarchical and ensemble forecasting techniques as a means to enhance the model performance. For instance, ensemble modeling, which involves the incorporation of supplementary models (e.g., neural networks, or additional ML and statistical models) holds promise for further improving the model performance.

Author Contributions: Conceptualization, P.N. and H.K.; Methodology, P.N., H.K. and K.K.; Software, P.N. and H.K.; Validation, P.N.; Formal analysis, P.N.; Writing—original draft, P.N.; Writing—review & editing, H.K. and K.K.; Supervision, K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Eskilstuna Strängnäs Energi och Miljö (ESEM) AB, the Knowledge Foundation (KKS), and Mälardalen University under the Industrial Technology (IndTech) Graduate School.

Data Availability Statement: The electrical load data underpinning this study are confidential and proprietary, exclusive to the funding company ESEM, and therefore cannot be shared publicly. Additionally, this study did not involve the creation or analysis of new data that could be made available for sharing.

Acknowledgments: The authors hereby thank ESEM for their support, especially Christer Wiik and Sara Jonsson from the Electrical Grid Department, and Per Örvind from the Energy Department. Additionally, Moksadur Rahman from ABB Corporate Research and Erik Dahlquist from Mälardalen University are acknowledged for their valuable input and fruitful discussions.

Conflicts of Interest: Author Pontus Netzell was employed by the company ESEM. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Huber, M.; Dimkova, D.; Hamacher, T. Integration of wind and solar power in Europe: Assessment of flexibility requirements. *Energy* **2014**, *69*, 236–246. [\[CrossRef\]](#)
2. Beiron, J.; Montañés, R.M.; Normann, F.; Johnsson, F. Combined heat and power operational modes for increased product flexibility in a waste incineration plant. *Energy* **2020**, *202*, 117696. [\[CrossRef\]](#)
3. Meliani, M.; Barkany, A.E.; Abbassi, I.E.; Darcherif, A.M.; Mahmoudi, M. Energy management in the smart grid: State-of-the-art and future trends. *Int. J. Eng. Bus. Manag.* **2021**, *13*, 18479790211032920. [\[CrossRef\]](#)
4. Cebulla, F.; Naegler, T.; Pohl, M. Electrical energy storage in highly renewable European energy systems: Capacity requirements, spatial distribution, and storage dispatch. *J. Energy Storage* **2017**, *14*, 211–223. [\[CrossRef\]](#)
5. Öhman, A.; Karakaya, E.; Urban, F. Enabling the transition to a fossil-free steel sector: The conditions for technology transfer for hydrogen-based steelmaking in Europe. *Energy Res. Soc. Sci.* **2022**, *84*, 102384. [\[CrossRef\]](#)
6. Nik, V.M.; Perera, A.; Chen, D. Towards climate resilient urban energy systems: A review. *Natl. Sci. Rev.* **2021**, *8*, nwaa134. [\[CrossRef\]](#)
7. Hong, T.; Pinson, P.; Wang, Y.; Weron, R.; Yang, D.; Zareipour, H. Energy forecasting: A review and outlook. *IEEE Open Access J. Power Energy* **2020**, *7*, 376–388. [\[CrossRef\]](#)
8. Hong, T.; Fan, S. Probabilistic electric load forecasting: A tutorial review. *Int. J. Forecast.* **2016**, *32*, 914–938. [\[CrossRef\]](#)
9. Wang, Y.; Chen, Q.; Sun, M.; Kang, C.; Xia, Q. An ensemble forecasting method for the aggregated load with subprofiles. *IEEE Trans. Smart Grid* **2018**, *9*, 3906–3908. [\[CrossRef\]](#)

10. Gajowniczek, K.; Ząbkowski, T. Two-stage electricity demand modeling using machine learning algorithms. *Energies* **2017**, *10*, 1547. [\[CrossRef\]](#)
11. SMHI. *Open Data API Docs—Meteorological Forecasts*; SMHI: Norrköping, Sweden, 2023.
12. NASA. *POWER | Data Access Viewer*; NASA: Washington, DC, USA, 2023.
13. Hong, T.; Gui, M.; Baran, M.E.; Willis, H.L. Modeling and forecasting hourly electric load by multiple linear regression with interactions. In Proceedings of the IEEE PES General Meeting, Minneapolis, MN, USA, 25–29 July 2010; pp. 1–8. [\[CrossRef\]](#)
14. Chabouni, N.; Belarbi, Y.; Benhassine, W. Electricity load dynamics, temperature and seasonality Nexus in Algeria. *Energy* **2020**, *200*, 117513. [\[CrossRef\]](#)
15. ESETT. *eSett Open Data*; ESETT: Helsinki, Finland, 2023.
16. Eroshenko, S.A.; Poroshin, V.I.; Senyuk, M.D.; Chunarev, I.V. Expert models for electric load forecasting of power system. In Proceedings of the 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIconRus), St. Petersburg/Moscow, Russia, 1–3 February 2017; pp. 1507–1513. [\[CrossRef\]](#)
17. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, Australia, 2018.
18. Seabold, S.; Perktold, J. Statsmodels: Econometric and statistical modeling with python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; Volume 57, pp. 10–25080. [\[CrossRef\]](#)
19. Işık, G.; Ögüt, H.; Mutlu, M. Deep learning based electricity demand forecasting to minimize the cost of energy imbalance: A real case application with some fortune 500 companies in Türkiye. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105664. [\[CrossRef\]](#)
20. Abu-Shikhah, N.; Elkarmi, F. Medium-term electric load forecasting using singular value decomposition. *Energy* **2011**, *36*, 4259–4271. [\[CrossRef\]](#)
21. Cho, H.; Goude, Y.; Brossat, X.; Yao, Q. Modeling and forecasting daily electricity load curves: A hybrid approach. *J. Am. Stat. Assoc.* **2013**, *108*, 7–21. [\[CrossRef\]](#)
22. Fan, M.; Hu, Y.; Zhang, X.; Yin, H.; Yang, Q.; Fan, L. Short-term load forecasting for distribution network using decomposition with ensemble prediction. In Proceedings of the 2019 Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 152–157. [\[CrossRef\]](#)
23. Bedi, J.; Toshniwal, D. Energy load time-series forecast using decomposition and autoencoder integrated memory network. *Appl. Soft Comput.* **2020**, *93*, 106390. [\[CrossRef\]](#)
24. Zha, W.; Ji, Y.; Liang, C. Short-term load forecasting method based on secondary decomposition and improved hierarchical clustering. *Results Eng.* **2024**, *22*, 101993. [\[CrossRef\]](#)
25. Baur, L.; Ditschuneit, K.; Schambach, M.; Kaymakci, C.; Wollmann, T.; Sauer, A. Explainability and interpretability in electric load forecasting using machine learning techniques—A review. *Energy AI* **2024**, *16*, 100358. [\[CrossRef\]](#)
26. Bandara, K.; Hyndman, R.J.; Bergmeir, C. MSTL: A seasonal-trend decomposition algorithm for time series with multiple seasonal patterns. *arXiv* **2021**, arXiv:2107.13462.
27. Krechiam, A.; Khadir, M.T. Algerian Electricity Consumption Forecasting with Artificial Neural Networks Using a Multiple Seasonal-Trend Decomposition Using LOESS. In Proceedings of the 2023 International Conference on Decision Aid Sciences and Applications (DASA), Annaba, Algeria, 16–17 September 2023; pp. 586–591. [\[CrossRef\]](#)
28. Al Shimmari, M.; Calliess, J.P.; Wallom, D. Load Profile Forecasting of Small and Medium-sized Businesses For Flexibility Programs. In Proceedings of the 2024 4th International Conference on Smart Grid and Renewable Energy (SGRE), Doha, Qatar, 8–10 January 2024; pp. 1–5. [\[CrossRef\]](#)
29. Zhou, S.; Li, Y.; Guo, Y.; Yang, X.; Shahidehpour, M.; Deng, W.; Mei, Y.; Ren, L.; Liu, Y.; Kang, T.; et al. A Load Forecasting Framework Considering Hybrid Ensemble Deep Learning with Two-Stage Load Decomposition. *IEEE Trans. Ind. Appl.* **2024**. [\[CrossRef\]](#)
30. Kolassa, S.; Rostami-Tabar, B.; Siemsen, E. *Demand Forecasting for Executives and Professionals*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2023.
31. Farrokhhabadi, M.; Browell, J.; Wang, Y.; Makonin, S.; Su, W.; Zareipour, H. Day-ahead electricity demand forecasting competition: Post-covid paradigm. *IEEE Open Access J. Power Energy* **2022**, *9*, 185–191. [\[CrossRef\]](#)
32. Kuster, C.; Rezugui, Y.; Mourshed, M. Electrical load forecasting models: A critical systematic review. *Sustain. Cities Soc.* **2017**, *35*, 257–270. [\[CrossRef\]](#)
33. Supapo, K.; Santiago, R.; Pacis, M. Electric load demand forecasting for Aborlan-Narra-Quezon distribution grid in Palawan using multiple linear regression. In Proceedings of the 2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Manila, Philippines, 1–3 December 2017; pp. 1–6. [\[CrossRef\]](#)
34. Miller, C.; Arjunan, P.; Kathirgamanathan, A.; Fu, C.; Roth, J.; Park, J.Y.; Balbach, C.; Gowri, K.; Nagy, Z.; Fontanini, A.D.; et al. The ASHRAE great energy predictor III competition: Overview and results. *Sci. Technol. Built Environ.* **2020**, *26*, 1427–1447. [\[CrossRef\]](#)
35. Tan, Y.; Teng, Z.; Zhang, C.; Zuo, G.; Wang, Z.; Zhao, Z. Long-Term Load Forecasting Based on Feature fusion and LightGBM. In Proceedings of the 2021 IEEE 4th International Conference on Power and Energy Applications (ICPEA), Busan, Republic of Korea, 9–11 October 2021; pp. 104–109. [\[CrossRef\]](#)
36. Herzen, J.; Lässig, F.; Piazzetta, S.G.; Neuer, T.; Tafti, L.; Raille, G.; Van Pottelbergh, T.; Pasieka, M.; Skrodzki, A.; Huguenin, N.; et al. Darts: User-friendly modern machine learning for time series. *J. Mach. Learn. Res.* **2022**, *23*, 5442–5447. [\[CrossRef\]](#)

37. Hewamalage, H.; Ackermann, K.; Bergmeir, C. Forecast evaluation for data scientists: Common pitfalls and best practices. *Data Min. Knowl. Discov.* **2023**, *37*, 788–832. [[CrossRef](#)] [[PubMed](#)]
38. Nti, I.K.; Teimeh, M.; Nyarko-Boateng, O.; Adekoya, A.F. Electricity load forecasting: A systematic review. *J. Electr. Syst. Inf. Technol.* **2020**, *7*, 1–19. [[CrossRef](#)]
39. Netzell, P.; Kazmi, H.; Kyprianidis, K. Applied Machine Learning for Short-Term Electric Load Forecasting in Cities-A Case Study of Eskilstuna, Sweden. *Scand. Simul. Soc.* **2023**, *200*, 29–38.
40. Kazmi, H.; Tao, Z. How good are TSO load and renewable generation forecasts: Learning curves, challenges, and the road ahead. *Appl. Energy* **2022**, *323*, 119565. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.