EPH-International Journal of Applied Science

Volume - 12, Issue - 02
April- 2026
DOI -https://doi.org/10.69980/enxr1t24

DATA-DRIVEN FORECASTING OF STEEL INDUSTRY ENERGY CONSUMPTION

[1]Dr.S.ALAGU

[1]ASSISTANT PROFESSOR; Department of Computer Science ; Hindustan College of Arts & Science, Chennai - 603103 Tamilnadu; ORCID ID: 0000-0002-1632-5032; Email ID: sivaalagu30@gmail.com

Received:- 11-03-2026, Revised:- 15-04-2026, Accepted:-23-04-2026, Published:-30-04-2026

Abstract
The steel industry is energy-intensive and requires accurate short-term electricity forecasting for efficient production scheduling, managing peak demand, and energy management. Data-driven modelling has a lot of potential for increasing the accuracy of forecasts at the operational level, especially with high-frequency data. This study developed and evaluated an interpretable forecasting framework for short-term steel industry energy consumption using temporal and electrical operating variables. A dataset containing 35,040 observations recorded at 15-minute intervals over one year was analysed. Following data-quality assessment, leakage-prone variables were excluded, and lagged consumption, rolling statistics, electrical measurements, and cyclical temporal features were generated. Persistence, Seasonal Naïve, Linear Regression, Random Forest, and XGBoost models were evaluated using chronological training, validation, and independent test partitions. Forecasting performance was assessed using MAE, RMSE, MAPE, sMAPE, and R2, while permutation importance was used to interpret predictor contributions. Random Forest achieved the best overall performance with an RMSE of 8.4778 kWh and an R2of 0.9265, whereas XGBoost produced the lowest MAE of 4.5443 kWh. Both ensemble models outperformed Linear Regression and baseline approaches. Lagged energy consumption, lagged reactive power, and one-hour rolling statistics were identified as the most influential predictors. The proposed framework offers accurate and interpretable short-term energy forecasts and proves to be of practical use for industrial load scheduling, production planning and energy-efficiency management purposes.

Keywords: *steel industry, energy forecasting, machine learning, Random Forest, industrial energy management*

1. Introduction

The steel sector plays a critical role in the overall economic growth of the world, as it provides a basic material for all infrastructure, transportation, construction, automotive production and energy systems. Although steelmaking is a strategically important industry, it is one of the most energy-intensive industrial processes, as electricity is used in the preparation of raw materials, melting, casting, rolling and finishing processes. Therefore, effective energy management has become a key factor in the production cost, efficiency and competitiveness of industries. With the spread of modern manufacturing strategies, the optimization of production processes is more and more combined with the planning of energy-aware operations to achieve better production performance and reduce resource use (Xu et al., 2021). At the same time, the development of manufacturing technologies, automation and process monitoring has allowed steel producers to increase the efficiency of their operations by making more intelligent manufacturing systems that will adapt to the varying manufacturing conditions (Backman et al., 2019).

Technologies of Industry 4.0 have been adopted very fast in the steel industry, leading to a rapid digital transformation. Operational data is continually collected from interconnected sensors, supervisory control systems, and industrial monitoring platforms at high frequencies in modern production facilities. In modern production facilities, a large amount of data is generated at high frequencies and is produced continuously from the operations. These environments, filled with data, enable predictive analytics, which helps manufacturers make decisions based on data and evidence instead of relying on traditional reactive methods and optimizes production schedules, resource utilization and equipment performance. Consequently, digitalization has emerged as a key factor for intelligent energy use in steel production, where the ability to accurately predict future electricity consumption has become more and more vital for sustainable industry operations (Branca et al., 2020).

The iron/steel industry has been a major focus of research on energy consumption due to the fact that the reduction of electricity consumption not only decreases production costs but also reduces the environmental impacts. The energy-saving technologies and the optimization of manufacturing processes have proved to have great potential to cut the energy usage of industry without compromising production capacity (He & Wang, 2017). With the growth of the world's electric power consumption due to industrialization, the economy and electrification, this challenge is becoming more important all the time. Therefore, it is crucial for manufacturing industries to have intelligent forecasting systems that can help them manage energy allocation effectively in the face of changing electricity supply dynamics (Ugli et al., 2025). Other estimates suggest that industrial electricity usage will steadily increase over the next few decades, highlighting the critical need for accurate forecasts to ensure operational resilience and sustainable manufacturing (Batra et al., 2025).

With the possibility of continuous operational measurements, data-driven forecasting methods now have a high profile in industrial environments. Data-driven approaches rely on historical observations to infer complex nonlinear interactions, rather than relying on mostly pre-defined physical relationships as is done in conventional mathematical modeling, and are capable of adapting to the changing operating conditions. Machine learning is especially appropriate for industrial applications with numerous interacting operational variables and rapidly changing production schedules, as it is able to learn the dynamic behaviour of their production processes (Zhao et al., 2018). Short-term electricity forecasts can be used in manufacturing systems to provide reliable estimates of electricity demand before decisions are made regarding operations, production scheduling, maintenance planning, demand-side management and energy procurement.

It has been recently shown considerable success in improving the accuracy of electricity forecasting in industry using sophisticated machine learning approaches. It has been recently shown that by adopting sophisticated machine learning approaches, the accuracy of electricity forecasting in industry has improved significantly. Under changing production scenarios, models that are able to model nonlinear temporal relations between variables outperform many traditional statistical models for short-term electricity forecasting, helping to make better operational planning and energy management decisions (Khan et al., 2022). In addition, alternative forecasting models based on component estimation have been found to offer superior stability and better forecasting accuracy in capturing simple temporal patterns in electricity demand for a wide variety of electricity consumption patterns (Shah et al., 2019). Comparative investigations further show that the incorporation of operational variables describing the behaviour of the physical system greatly improves the predictive performance, and that machine learning algorithms outperforming conventional regression-based methods can be used in industrial environments (Kapp et al., 2023). In addition to prediction accuracy, efficient use of energy in the industrial sector is recognized as an integral part of sustainable manufacturing, with enhanced energy efficiency directly impacting resource conservation and environmental sustainability in the long run (Chen et al., 2021).

There is significant variability in manufacturing processes because electricity consumption of industries is affected by equipment efficiency, mechanical losses, production intensity and scheduling. These interactions are intricate and need models they can rely on that represent the operational dynamics and not simply on the measurements of electricity in the past. The introduction of emerging digital twin technologies has reinforced this capacity along with the ability to simulate manufacturing behaviour in real-time to optimise operational decisions before they are implemented (Holmberg & Erdemir, 2017; Yu et al., 2022).

Although significant progress has been made, there are some methodological issues that still need to be addressed. Previous studies have examined prediction accuracy in isolation, and few studies have examined the problem in the context of strict, chronological validation with independent test sets. Additionally, many existing works have focused on using only a few temporal features, while the total effect of several temporal features and rolling statistical features that affect high-frequency industrial electricity demand have been neglected. Moreover, few studies have paid attention to prediction model interpretability, which, in turn, limits the applicability of forecasting systems for industrial decision-makers who need clear and easily understandable information about which variables affect prediction accuracy.

In the present study, these limitations are overcome by creating a data-driven forecasting model of energy consumption for the short term in the steel industry based on high-resolution operational data collected for an entire production year. The strategy consists of temporal feature engineering, multiple forecasting models (baseline models, Linear Regression, Random Forest and Extreme Gradient Boosting), and a chronological training, validation and independent testing approach. Model performance is assessed with complementary statistical indicators, and the variables that are the most important for the model to predict well are identified using a feature importance analysis based on permutations. The integrated approach offers a practical and meaningful approach to forecasting that is accurate, interpretable and suitable for today's steel manufacturing environment for intelligent energy management and operational decision-making.

2. Methodology

2.1 Research Design

The study was designed using a quantitative methodology based on data-driven forecasting to forecast the energy use for the steel industry in the short term with the application of supervised machine learning techniques. The analytical methodology was composed of a series of steps: data preparation, exploratory analysis, feature engineering, model development, model validation, and independent performance evaluation. To avoid information leakage within and between model development and evaluation phases, and to maintain temporal characteristics of the observations, a chronological (temporal) forecasting method was used. Different forecasting methods such as statistical learning, linear programming and ensemble learning methods were assessed and compared to find the most effective forecasting method for real operational conditions of these industrial energy consumptions.

2.2 Data Source

The data used for this study came from the study of Sathishkumar et al. (2021). It includes operational measurements for electricity consumption, reactive power, power factor, temporal characteristics and variables related to electrical production from a steel manufacturing facility. The observations were gathered at 15-minute time steps for a calendar year, which is adequate temporal resolution for the development and evaluation of a short-term forecasting model (Sathishkumar et al., 2021).

2.3 Data Preparation and Feature Engineering

The dataset was initially analysed to ensure temporal continuity, duplications, missing data and inconsistent data for recorded variables. The variable timestamp was converted to a regular datetime format, and it was used to ensure the observations were kept in the right order in time throughout the analysis. To avoid target leakage while building the model, variables that could directly provide information about the future targets were eliminated. To address the periodic consumption patterns, additional temporal descriptors were created such as hour of day, month, day of week and cyclical sine and cosine representations. To account for the short-term and seasonal temporal dependencies, the forecasting models, features based on historical lags of 15 minutes, one hour, one day and one week, as well as features based on rolling means and rolling standard deviations, were created.

2.4 Forecasting Model Development

The engineered data set was split into a training set (70%), a validation set (15%) and an independent testing set (15%) and future observations were never used in the model training. Four prediction methods were tested: Persistence Baseline, Seasonal Naïve Baseline, Linear Regression, Random Forest and Extreme Gradient Boosting (XGBoost). Model selection was done with the validation observations, and the final assessment was done only with the independent test observations after retraining the selected models with the training and validation observations. This approach gave an unbiased assessment of forecasting accuracy with unpredictable operating conditions.

2.5 Model Evaluation and Interpretation

The mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), symmetric mean absolute percentage error (sMAPE) and coefficient of determination (R^2) were used to assess the accuracy of the forecasts. Complementary metrics measured the accuracy of prediction, the size of the error, and the ability to explain, for all forecasting models. The final model was further explained using the permutation-based feature importance analysis in order to determine the contribution of each predictor of the model to the forecasting performance. Residual analysis, actual versus predicted comparisons and statistical comparison of forecast errors followed to test for stability of the forecasts, assessment of systematic bias and comparison of the predictive behaviour of the machine learning models.

3. Results

3.1 Descriptive Characteristics of the Steel Industry Energy Dataset

The data were obtained from 35,040 observations at 15-minute intervals for the entire calendar year, giving analysts a continuous time series that spans one year. A data quality evaluation showed that there were no missing values, duplicate observations, or irregular sampling intervals, which meant no missing data or temporal reconstruction had to be done prior to model development. There was a wide range of variation in the numerical variables by operational conditions. The range of energy consumption was from a minimum of 0.00 kWh to a maximum of 157.18 kWh, with a mean consumption of 27.39 kWh, showing significant differences between idle and peak production times. Reactive power variables also showed significant variations and the power factor values were relatively stable at high operating efficiencies. It can be noted that the energy consumption and reactive power distributions are skewed positively, indicating that the periods of low consumption are more frequent, with relatively lesser amounts but higher energy peaks. Table 1 shows the descriptive statistics of the numerical variables used for forecasting.

Table 1. Descriptive Statistics of Numerical Variables

Variable	Observations	Mean	Standard Deviation	Minimum	Median	Maximum	Skewness
Usage (kWh)	35,040	27.387	33.444	0.00	4.57	157.18	1.197

Lagging Current Reactive Power (kVarh)	35,040	13.035	16.306	0.00	5.00	96.91	1.438
Leading Current Reactive Power (kVarh)	35,040	3.871	7.424	0.00	0.00	27.76	1.734
CO_2 Emissions (tCO_2)	35,040	0.012	0.016	0.00	0.00	0.07	1.149
Lagging Current Power Factor	35,040	80.578	18.921	0.00	87.96	100.00	−0.606
Leading Current Power Factor	35,040	84.368	30.457	0.00	100.00	100.00	−1.512
Seconds from Midnight	35,040	42,750.000	24,940.534	0.00	42,750.00	85,500.00	0.000

3.2 Temporal Patterns of Industrial Energy Consumption

The results of the daily and hourly analyses showed different temporal variations in electricity demand over the observation period. Daily mean energy consumption showed a high variation throughout the year in line with the different production schedules and operational needs. There were a number of high withdrawal periods in the initial months, after which there were interspersed periods of high and low demand. There was little energy use throughout the day, but the use started to rise significantly at about 08:00. The electricity demand was high during regular production time and was found to be decreasing slowly during the evening period, which suggests that there is a good relationship between electricity usage and industrial production schedules. The daily temporal trend of mean energy consumption over the study period is illustrated in Figure 1(A). The average energy consumption profile over time is presented in Figure 1(B) and is typical of a production cycle for a single day with low and high energy demands for industry.

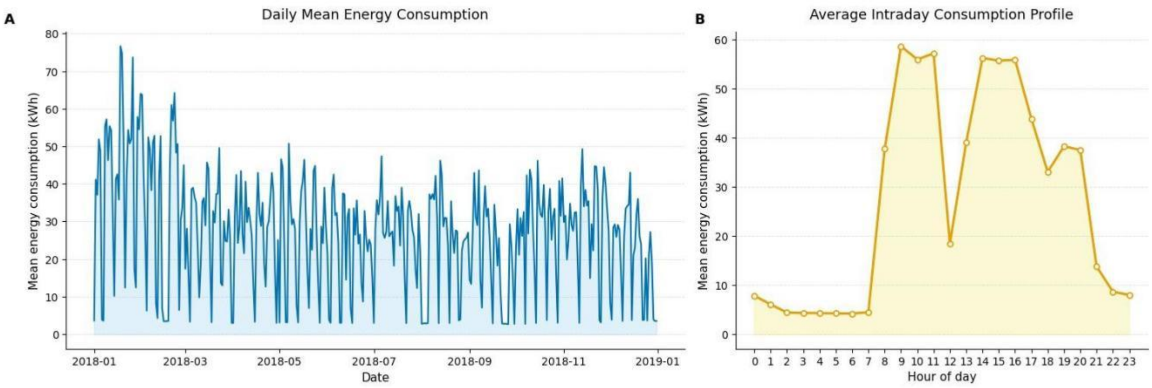

Figure 1. Temporal Patterns of Steel Industry Energy Consumption. (A) Daily Mean Energy

Consumption. (B) Average Intraday Consumption Profile.

3.3 Comparative Performance of Forecasting Models

An independent test set was used to test the prediction capability of 5 forecasting approaches. All the ensemble learning models performed better than the statistical baseline and linear regression methods. Random Forest had the lowest root mean squared error (RMSE) (8.48 kWh) and highest coefficient of determination (0.9265), while XGBoost had the lowest mean absolute error (4.54 kWh). Linear Regression performed well in predicting the data and was still not as accurate as the ensemble models. The forecasts obtained by persistence were a reasonable short-term reference, and the forecasts obtained by the season naïve approach showed large errors. The results showed that the machine learning approaches successfully learned the non-linear temporal patterns among the industrial energy data compared to the traditional forecasting models. Table 2 provides a comparison of the predictive performance of all forecasting models investigated in this paper. Figure 2(A) shows the comparison of mean absolute error and root mean squared error of the forecasting models on the independent test dataset. The coefficient of determination for all models evaluated is shown in Figure 2(B), which indicates that the Random Forest and XGBoost algorithms have the best explanatory power.

Table 2. Independent Test-Set Performance of Forecasting Models

Rank	Model	MAE	RMSE	MAPE (%)	sMAPE (%)	R^2
1	Random Forest	4.5572	8.4778	31.9816	24.2251	0.9265
2	XGBoost	4.5443	8.8726	27.7279	20.8797	0.9195
3	Linear Regression	6.0549	10.5706	53.7480	51.0297	0.8857
4	Persistence Baseline	5.0388	11.6752	19.9630	14.5343	0.8606
5	Seasonal Naïve Baseline	12.8135	24.4028	121.6505	39.2757	0.3911

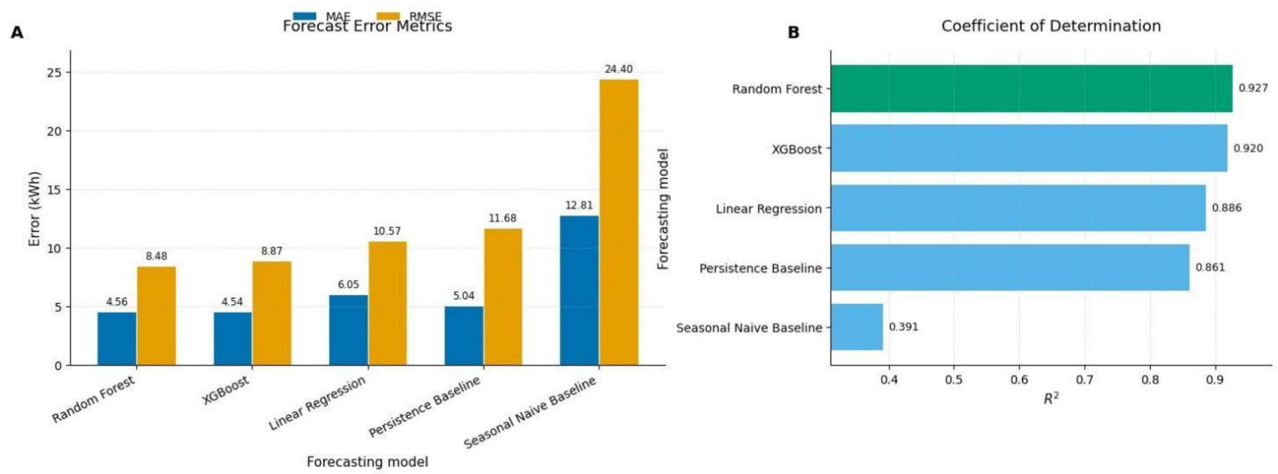

Figure 2. Independent Test-Set Performance of Forecasting Models. (A) Comparison of Mean Absolute Error and Root Mean Squared Error. (B) Comparison of Coefficient of Determination (R^2).

3.4 Forecasting Interpretation and Predictor Importance

Forecasting results showed a good agreement between the electricity consumption observed from the end of the training period and the one predicted during the independent test period. The overall temporal features of the model were quite well imitated, and most of the high-demand events were well captured, although underestimation was seen for a smaller number of extreme consumption peaks. Further, predictor importance analysis revealed that recent energy use was most important for forecasting accuracy. The previous 15-minute energy observation was found to be the most important predictor, followed by the previous energy observations with short-term rolling averages. The time-of-day cyclical variables were also significant and evidenced that operational scheduling was a key factor in shaping electricity demand. As a whole, these results show that operational experience and the production trend of the past few years are key factors in short-term industrial energy forecasting. The comparison of the observed and predicted energy consumption using the best-performing forecasting model is shown in Figure 3(A) for the independent test period. The permutation importance values for the most important predictors, as shown in Figure 3(B), suggest dominance by recent energy consumption, reactive power measurements and the most recent temporal features for forecasting performance.

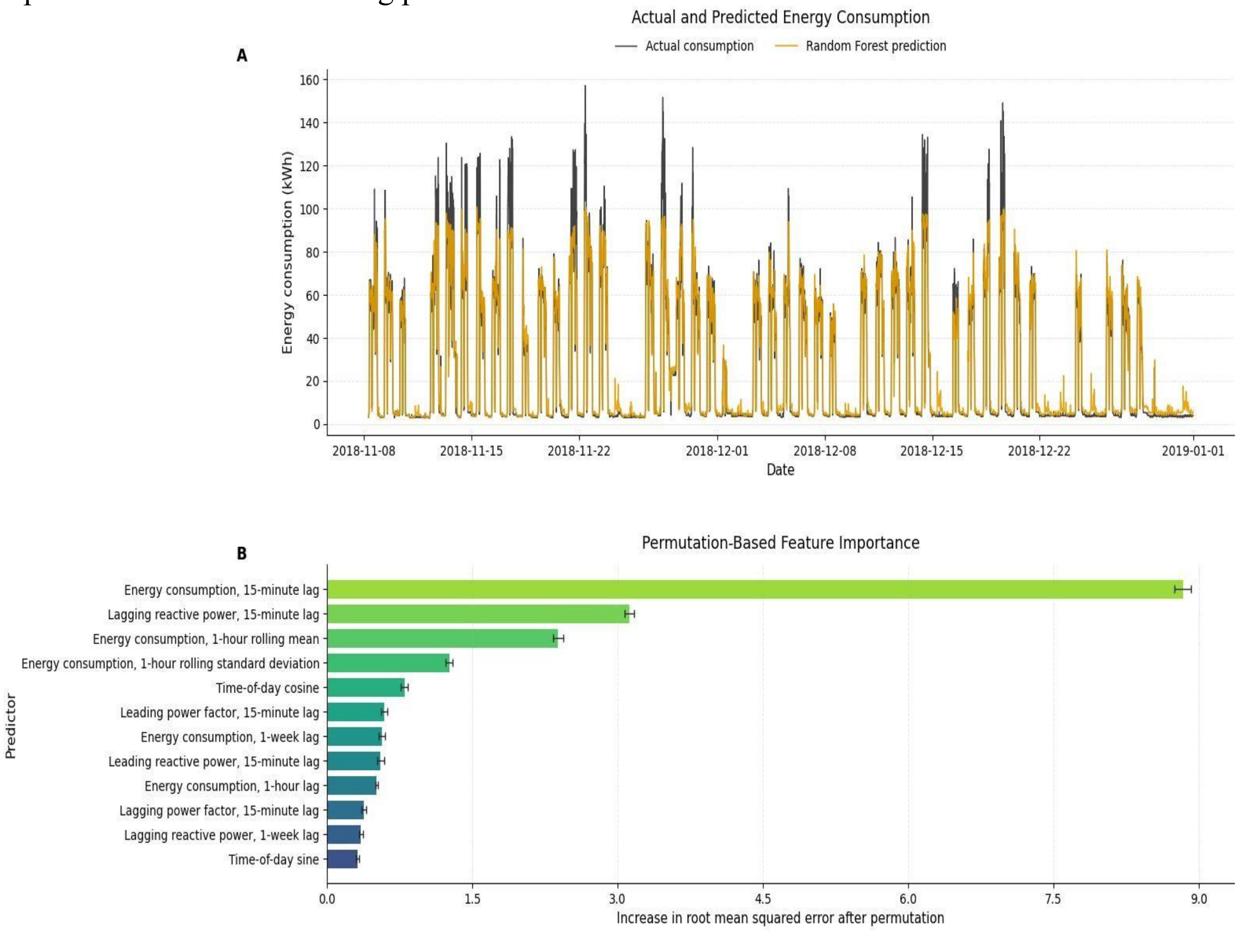

Figure 3. Forecasting Performance and Model Interpretation. (A) Actual versus Predicted Energy Consumption. (B) Permutation-Based Feature Importance of the Best-Performing Forecasting Model.

3.5 Importance of Predictive Variables

Importance analysis based on permutation showed that the forecasts could be attributed to a relatively small set of predictors. The previous 15 minutes of energy use had the largest importance score (8.8366), which was significantly larger than all of the other predictors. Lagged reactive power and one-hour rolling mean energy consumption were the second most influential group and reveal that there is much to be learnt in recent electrical operating conditions. Other power factor parameters, the weekly lagged consumption, and cyclical time-of-day parameters were also found and observed to contribute to the additional contributions, while longer historical parameters lagged contributed comparatively less. The results show that short-term operational memory contributes more to prediction than do the periodic temporal features, which contribute complementary information to the prediction process. The fifteen most important predictors for the Random Forest forecasting model are ranked in Table 3.

Table 3. Top Permutation-Based Predictors of Steel Industry Energy Consumption

Rank	Predictor	Mean Permutation	Standard

		Importance	Deviation
1	Energy Consumption, 15-minute Lag	8.8366	0.0873
2	Lagging Reactive Power, 15-minute Lag	3.1213	0.0520
3	Energy Consumption, 1-hour Rolling Mean	2.3854	0.0520
4	Energy Consumption, 1-hour Rolling Standard Deviation	1.2608	0.0341
5	Time-of-Day Cosine	0.7977	0.0333
6	Leading Power Factor, 15-minute Lag	0.5911	0.0335
7	Energy Consumption, 1-week Lag	0.5654	0.0291
8	Leading Reactive Power, 15-minute Lag	0.5530	0.0387
9	Energy Consumption, 1-hour Lag	0.5100	0.0116
10	Lagging Power Factor, 15-minute Lag	0.3805	0.0260
11	Lagging Reactive Power, 1-week Lag	0.3470	0.0202
12	Time-of-Day Sine	0.3158	0.0158
13	Hour of Day	0.2430	0.0194
14	Lagging Reactive Power, 1-hour Lag	0.2367	0.0115
15	Energy Consumption, 1-day Lag	0.1959	0.0215

4. Discussion

The forecasting results show that by taking the recent consumption history, electrical operating conditions and temporal structure into account, the energy demand of the short-term steel industry can be modelled with high accuracy. Random Forest model outperformed the other models in terms of overall test performance, with an RMSE of 8.4778 kWh and an R^2 of 0.9265; while the XGBoost model had the lowest MAE of 4.5443 kWh. The slight discrepancy between the two ensemble techniques suggests that both techniques were able to capture the nonlinear characteristics of the consumption series well. This clear superiority to Linear Regression and the two baseline methods demonstrates that the electricity demand of industry was not being driven by a linear trend. The relatively poor performance of the seasonal naïve benchmark also indicates that this does not account for the observed operational variability due to weekly repetition. Production cycles were apparent but low-frequency variations in the activity of the plant remained important.

This is confirmed by the temporal analysis. Overnight energy consumption was low and it surged higher during the daytime, in particular when active production was in progress. This is representative of the operational characteristics of steel production, in that electricity demand varies in step with the activation and deactivation of equipment, the intensity of the processes, and the scheduled production. Feature engineering for both the recent and cyclical behaviour was thus added to the model results. The 15-minute lag of energy consumption had a much stronger influence, with the next most significant being the 1-hour rolling statistics and lagged reactive power. The overwhelming influence of recent observations suggests that short-term industrial demand is indeed persistent, and the time-of-day features suggest that there is residual value in time-of-day information even if the production schedules are repeated.

This research confirms a study that revealed data-driven forecasting can help energy efficiency and maintenance decisions in a smart manufacturing facility. The use of predictive energy models becomes more valuable when they are related to the broader manufacturing management goals, as highlighted by Bermeo-Ayerbe et al. (2022). The present results support that view since variables other than past consumption, such as reactive power and power factor variables, also helped to build a model that is informative. Forecasting in the steel sector is another example of successful evidence reported by combining complementary learning mechanisms in ensemble structures (Mubarak et al., 2023), such as the use of complementary learning mechanisms for forecasting active and reactive energy.

The good performance of the Random Forest and XGBoost is also consistent with previous studies on smart factories where data mining algorithms were shown to be the most effective in capturing the complex consumption pattern as compared to the simpler ones (Sathishkumar et al., 2020). The results build on the previous effort by chronological validation and an independent test period, making it harder to overestimate the results. Studies on high-energy steel production systems have also underlined the significance of machine learning in Industry 4.0 trials, such as when prediction models are implemented to forecast customer demand when production conditions change (Khalid et al., 2024). The current results corroborate that argument and demonstrate that temporal history has a significant impact on prediction accuracy.

These results are also relevant to the practical world, as shown by recent research on the application of machine learning to enhance energy efficiency in steelmaking processes (Zhang et al., 2025). The ability to model more than 92% of the unseen consumption data can help the plant's management team predict short-term demands, effectively schedule equipment and detect and flag unusual demands in advance to prevent disruption to operations. The wide analysis of the iron and steel industry revealed that energy prediction also plays a role in environmental and economic planning, with the ability to create more efficient resource allocation and emissions reduction strategies (Gu et al., 2023). This forecasting framework can thus be applied to work towards both operational and sustainability goals.

Linear Regression had a relatively good performance and was always outperformed by the ensemble models. This result corroborates findings that when explanatory variables are well defined, multiple linear regression could be useful for estimating industry-level demand, but it would fail when the relationship is strongly nonlinear and/or is complex over time (Maaouane et al., 2021). According to the present results, linear models are still more transparent models, while in the case of the high-frequency steel industry data, ensemble methods are more suitable. Previous work on hourly forecasting for the steel industry also showed that machine learning could be effective for learning the varying energy demand patterns (Sathishkumar et al., 2019).

This has operational and strategic consequences. Short-term forecasts can be used to help in production planning, energy purchasing, peak demand control and load scheduling. Forecasting systems should also be related to electrical quality variables, in addition to energy history, due to the importance of reactive power variables. This integration can be useful for recognizing less efficient operating modes and optimise the coordination between production and energy management. The advantages are in line with the increased focus on decarbonizing industrial energy consumption and carbon emissions in the steel industry, more broadly (Holappa, 2020).

There are some caveats to consider. This set of data is from one facility and one year of operation; generalizing beyond this should be done with caution. Weather, production volume, equipment condition and product mix were not available, but each could have an impact on demand. The framework should be studied in several plants, in a longer time frame and with various forecasting periods in the future. It is also possible to consider other ensemble models, probabilistic forecasts, and online updating techniques that evolve with the production situation in the future.

5. Conclusion

A data-driven framework based on temporal feature engineering, chronological validation, and ensemble learning was presented, which enabled accurate short-term forecasting of steel industry energy consumption of the steel industry in the short term. Random Forest performed the best in the overall test accuracy with the lowest RMSE and highest R^2 value, and XGBoost performed the best with the lowest MAE. Industrial energy data was found to be well suited for the use of nonlinear approaches as in both models, they clearly outperformed Linear Regression, persistence forecasting and the seasonal naïve model. These were all short-term operational memory and electrical conditions, with recent energy consumption, lagged reactive power and one-hour rolling statistics most important in predicting performance. The significant intraday periodicity also highlighted the role of production activities on the electricity demand. The framework is useful in scheduling loads, production planning, controlling high demand periods, and sourcing energy. Its clear feature importance analysis also helps to make better decisions about operational parameters and determine which features are important for the prediction of demand. Although only one plant and one year of observations were analysed, the results are a solid basis for future multi-plant analyses, for forecasts that can be longer than one year and for models that include production, equipment and environmental information.

References

1. Backman, J., Kyllönen, V., & Helaakoski, H. (2019). Methods and tools of improving steel manufacturing processes: Current state and future methods. IFAC-PapersOnLine, 52(13), 1174-1179.
2. Batra, L., Harris, D., Katsigiannakis, G., Mackovyak, J., Parmar, H., & Scheller, M. (2025). Rising current: America's growing electricity demand. ICF Report at https://www. ourenergypolicy. org/resources/rising-current-americas-growing-electricitydemand.
3. Bermeo-Ayerbe, M. A., Ocampo-Martinez, C., & Diaz-Rozo, J. (2022). Data-driven energy prediction modeling for both energy efficiency and maintenance in smart manufacturing systems. Energy, 238, 121691.
4. Branca, T. A., Fornai, B., Colla, V., Murri, M. M., Streppa, E., & Schröder, A. J. (2020). The challenge of digitalization in the steel sector. Metals, 10(2), 288.
5. Chen, X., Shuai, C., Wu, Y., & Zhang, Y. (2021). Understanding the sustainable consumption of energy resources in global industrial sector: Evidences from 114 countries. Environmental Impact Assessment Review, 90, 106609.
6. Gu, Y., Liu, W., Wang, B., Tian, B., Yang, X., & Pan, C. (2023). Analysis and prediction of energy, environmental and economic potentials in the iron and steel industry of China. Processes, 11(12), 3258.
7. He, K., & Wang, L. (2017). A review of energy use and energy-efficient technologies for the iron and steel industry. Renewable and Sustainable Energy Reviews, 70, 1022-1039.
8. Holappa, L. (2020). A general vision for reduction of energy consumption and CO2 emissions from the steel industry. Metals, 10(9), 1117.
9. Holmberg, K., & Erdemir, A. (2017). Influence of tribology on global energy consumption, costs and emissions. Friction, 5(3), 263-284.
10. Kapp, S., Choi, J. K., & Hong, T. (2023). Predicting industrial building energy consumption with statistical and machine-learning models informed by physical system parameters. Renewable and Sustainable Energy Reviews, 172, 113045.
11. Khalid, R., Ahmed, N., & Anglani, N. (2024, October). Predictive model for high energy consumption processes in the steel manufacturing industry using machine learning tools: industry 4.0 pathway. In 2024 IEEE International Conference on Artificial Intelligence & Green Energy (ICAIGE) (pp. 1-6). IEEE.
12. Khan, Z. A., Ullah, A., Haq, I. U., Hamdy, M., Mauro, G. M., Muhammad, K., ... & Baik, S. W. (2022). Efficient short-term electricity load forecasting for effective energy management. Sustainable Energy Technologies and Assessments, 53, 102337.
13. Maaouane, M., Zouggar, S., Krajačić, G., & Zahboune, H. (2021). Modelling industry energy demand using multiple linear regression analysis based on consumed quantity of goods. Energy, 225, 120270.
14. Mubarak, H., Sanjari, M. J., Stegen, S., & Abdellatif, A. (2023). Improved active and reactive energy forecasting using a stacking ensemble approach: steel industry case study. Energies, 16(21), 7252.
15. Sathishkumar, V. E., Lee, M. B., Lim, J. H., Shin, C. S., Park, C. W., & Cho, Y. Y. (2019). Hourly steel industry energy consumption prediction using machine learning algorithms.

In Proceedings of the Korea Information Processing Society Conference (pp. 585-588).
16. Sathishkumar, V. E., Lee, M., Lim, J., Kim, Y., Shin, C., Park, J., & Cho, Y. (2020). An energy consumption prediction model for smart factory using data mining algorithms. KIPS Transactions on Software and Data Engineering, 9(5), 153-160.
17. Sathishkumar, V. E., Shin, C., & Cho, Y. (2021). Steel Industry Energy Consumption [Data set]. UCI Machine Learning Repository. https://doi.org/10.24432/C52G8C
18. Shah, I., Iftikhar, H., Ali, S., & Wang, D. (2019). Short-term electricity demand forecasting using components estimation technique. Energies, 12(13), 2532.
19. Ugli, A. T. A., Pastorelli, M., Shuxratovich, K. K., Makhmudovich, M. K., & Qizi, K. Y. U. (2025, October). Rising Of Global Electricity Demand: The Importance Of Industry And Renewable Energy Integration. In 2025 14th International Conference on Renewable Energy Research and Applications (ICRERA) (pp. 1052-1059). IEEE.
20. Xu, Z. J., Zheng, Z., & Gao, X. Q. (2021). Operation optimization of the steel manufacturing process: A brief review. International Journal of Minerals, Metallurgy and Materials, 28(8), 1274-1287.
21. Yu, W., Patros, P., Young, B., Klinac, E., & Walmsley, T. G. (2022). Energy digital twin technology for industrial energy management: Classification, challenges and future. Renewable and Sustainable Energy Reviews, 161, 112407.
22. Zhang, X., Yang, L., & Ma, D. (2025). The importance of energy consumption and the need for efficiency in the steel industry using machine learning. Journal of Electrical Engineering & Tech
23. Zhao, J., Wang, W., & Sheng, C. (2018). Data-driven prediction for industrial processes and their applications. Springer International Publishing: