science, engineering and health studies

Research Article

Science, Engineering and Health Studies
https://li01.tci-thaijo.org/index.php/sehs
ISSN (Online): 2630-0087

Digital twin for decision support system of industrial utility management

Thanatip Satjeenphong, Chanin Panjapornpon, and Santi Bardeeniz*

Department of Chemical Engineering, Center of Excellence on Petrochemical and Materials Technology, Faculty of Engineering, Kasetsart University, Bangkok 10900, Thailand

***Corresponding author**:*
Chanin Panjapornpon
fengcnp@ku.ac.th

Received: *12 May 2023*
Revised: *7 May 2024*
Accepted: *17 May 2024*
Published: *5 November 2024*

***Citation**:*
Satjeenphong, T., Panjapornpon, C., and Bardeeniz, S. (2024). Digital twin for decision support system of industrial utility management. Science, Engineering and Health Studies, 18, 24040002.

ABSTRACT

Manufacturing and industrial operations rely heavily on energy that is generated mostly from fossil fuels, such as coal, oil, and natural gas, which harm the environment and contribute to climate change. Thus, renewable energy is being integrated into industrial processes to lessen environmental effects and reduce fossil fuel usage. However, the renewable source performance process can be greatly affected by disturbances and constraints, such as ambient air temperature and relative humidity, minimum utility consumption, and the total energy required, making effective control difficult. This study proposes a digital twin built with machine learning regression techniques for load demand forecasting as a decision-support system for industrial utility management. From the results, the ensemble tree (ET) model produced the highest accuracy, based on validation and test dataset root mean squared error values of 23.164 and 27.558, respectively, and R^2 values of 0.96 and 0.95, respectively. The digital twin and load demand forecasting approaches effectively created an efficient operating schedule for industrial utility management. The ET model had a total error of 23.86%, substantially lower than the average load demand's total error of 65.29%. Therefore, the ET model with weather conditions in four scenarios could be recommended to optimize energy utilization when creating the operating schedule.

Keywords: digital twin; utility management; machine learning; artificial intelligence

1. INTRODUCTION

Energy is an essential input required for all manufacturing and industrial processes (Rahbar et al., 2017). Overall, most energy production depends on fossil fuels, such as coal, natural gas, and oil, all of which have environmental limitations and drawbacks, causing climate change, and are diminishing energy source options (Nam et al., 2020). The current trend is toward clean energy sources such as renewable energy. On the other hand, renewable energy is considered limited and unstable because the energy produced varies and is difficult to predict due to uncontrollable external natural factors such as weather, daylight hours, night, and season.

The industrial sector consumes a lot of energy and resources and is a major contributor to global greenhouse gas emissions (Panjapornpon et al., 2023a). Effective industrial utility management practices are essential to reduce the environmental impact of industrial operations and to foster sustainability in response to these environmental concerns. Technical know-how, process management skills, and the use of cutting-edge technologies are all necessary for effective utility management. Planning, implementing, and monitoring systems and procedures are essential to optimize the use of resources, such as energy, water, waste, and waste heat, representing essential utility management components that can help to increase efficiency, reduce costs, and minimize environmental impacts (Ahmad et al., 2019).

In recent years, there has been increased interest in the integration of electricity generation sources such as diesel generators and solar energy (Chauhan et al., 2021). During power outages or peak demand periods, diesel generators can be used as a backup power source, while solar energy can be used to meet the base load of electricity demand. Waste heat boilers can also contribute to significant energy savings by utilizing waste heat generated from the processes into usable steam. Aziz et al. (2022) proposed the importance of selecting an appropriate dispatch strategy when designing hybrid energy systems as it has a major effect on the stability, reliability, and the environmental and economic performance of the system. Roslan et al. (2021) proposed an optimized controller for optimizing microgrid energy management to minimize the total operating cost of distributed energy resource units, reduce environmental emissions, and address complex constraint optimization problems. Integrating renewable energy sources and energy management systems are key components of industrial utility management. One major shortcoming of the above-mentioned study was the lack of high-resolution data of real-time disturbances for estimating load demand. This includes hourly, daily, and monthly data on ambient temperature and relative humidity. These inputs are critical in improving the accuracy of load demand forecasts, leading to better management of industrial utility decisions. Obtaining and utilizing these new data inputs to load forecasting models is critical to the improvement of their accuracy and usefulness.

Recently, digital twin technology has gained in popularity due to its numerous advantages in the management of energy resources (Yu et al., 2022). Real-time data and insights into energy production, distribution, and consumption may be made available by a digital twin, as a virtual representation of a physical system. This can facilitate resource optimization and better decision-making. This technology has the potential to significantly increase the efficiency of energy generation and distribution, resulting in cost savings and lower carbon emissions (Panjapornpon et al., 2023b). A robust and comprehensive digital twin architecture has been proposed for future complex power plants to achieve higher reliability, supply availability, and maintenance at lower costs (Sleiti et al., 2022). Their model focused on the importance of using real-world data, physical representation of the entire system, and localized offline simulations provided using the deep learning algorithm and digital twin dynamic system model prediction with an autoregressive vector pattern to detect anomalies. Hasan et al. (2023) proposed an interactive digital twin platform and showed the importance of digital twin technology in offshore wind farm construction. Their platform included real-time data, physics models, and machine-learning regression algorithms to increase production, decrease downtime, and improve maintenance efficiency. The adoption of digital twin technology has the potential to provide more real-time data and insights, resulting in improved decision-making and resource optimization.

Therefore, the aim of the current study was to enhance decision-making in industrial utility management by using machine learning regression models that take real-time disturbances into account such as ambient temperature and relative humidity data from the previous seven days. The models also take other important factors into consideration such as the hour, month, and day of the week. Incorporating these variables into the models will result in more accurate load demand forecasts, which will help industrial utility management make better decisions. To optimize industrial systems, the suggested methodology uses a digital twin, an operating schedule, and artificial intelligence-based regression techniques for load demand forecasting. In contrast, to load demand forecasting, which is used to increase accuracy for the operating schedule, the operating schedule is created to maximize the use of utilities. The digital twin, which represents an industrial system virtually, makes it easier to simulate and analyze system behavior. This allows for the optimization of industrial systems, which lowers costs and improves sustainability.

2. METHODOLOGY

2.1 Description of utility plant

The utility plant studied consisted of various energy sources, both renewable and non-renewable, consisting of diesel generators, photovoltaic (PV) arrays, and a waste heat boiler that generated steam as part of the overall process. The diesel generator acted as the main energy source due to its constant electrical output. PV arrays, which rely on time-dependent solar irradiation (Lan et al., 2019), were run in conjunction with the diesel generator during the day to reduce fuel consumption, emissions, and costs. PV arrays generate electricity only using solar radiation, which means the system must rely heavily on diesel generators at night. This energy system was designed to facilitate stable and sustainable energy production using both renewable and non-renewable energy sources, subject to availability and performance. Making the most of the available sources can reduce dependence on non-renewable energy sources and create a more sustainable and environmental-friendly energy system. Figure 1 shows a schematic diagram of the utility plant energy system design, highlighting the different components and their interconnections.

2.2 Digital twin for industrial utility management

Recently, digital twin technology has emerged as a new approach to improving utility management in manufacturing. A digital twin is used to describe a digital representation of a physical utility plant or system that can be used for performance analysis and improvement. The physical utility plant, including its instruments and machinery, is crucial to the development of a digital twin. Based on prior knowledge, historical data, and real-time information obtained from the physical plant, the virtual utility plant includes models that simulate physical processes and evaluate current and future conditions. Energy sources, such as a diesel generator and PV arrays, can be optimized, saving costs while reducing the environmental impact by using load demand forecasting and machine learning regression techniques. The digital twin can generate an operating schedule that optimizes energy use and further reduces costs based on forecast load demand for utilities such as electricity and steam (Chen et al., 2020).

The forecasting of load demand is one of the key applications of digital twin technology in managing industrial utilities. By assisting utilities in making plans for future energy requirements and enhancing the production and distribution of energy resources, it plays a crucial role in managing energy resources. Figure 2 shows the accurate hourly load demand profile produced by the digital twin load forecasting model using historical load data, weather data, and other relevant resources. By combining real-time data from actual utility plants with data from outside sources, digital twins can simulate a variety of scenarios. For example, based on hourly load demand profiles and weather data, the digital twin can optimize the use of a diesel generator and PV arrays to generate electricity.

2.2.1 Process modeling and simulation

The UniSim Design Suite (Honeywell, United States) was utilized to create a comprehensive model of the studied plant, as shown in Figure 3. The plant was intended for both electricity and steam generation. It consisted of several essential components, including a pump, compressor, reactor, turbine, and waste heat boiler. The process begins by feeding fuel to the pump to increase the pressure based on the operating conditions and then introducing air to the compressor in a 27.5:1 ratio to initiate combustion. The resulting chemical reaction between the fuel and oxygen generates heat energy and exhaust gases, including carbon dioxide and water. Then, the heat energy and exhaust gas are utilized in a turbine to generate electricity and in a waste heat boiler to generate steam for utility purposes via heat exchange with water.

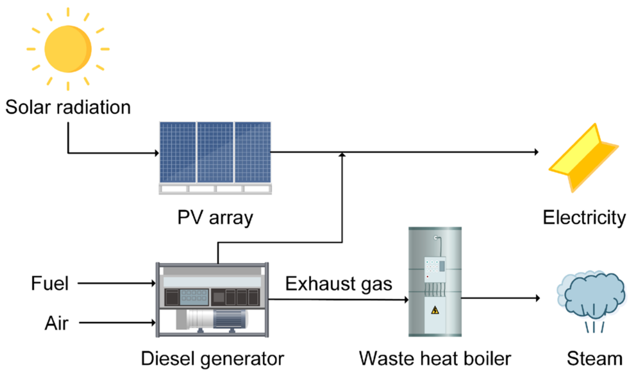

Figure 1. Schematic diagram of utility plant

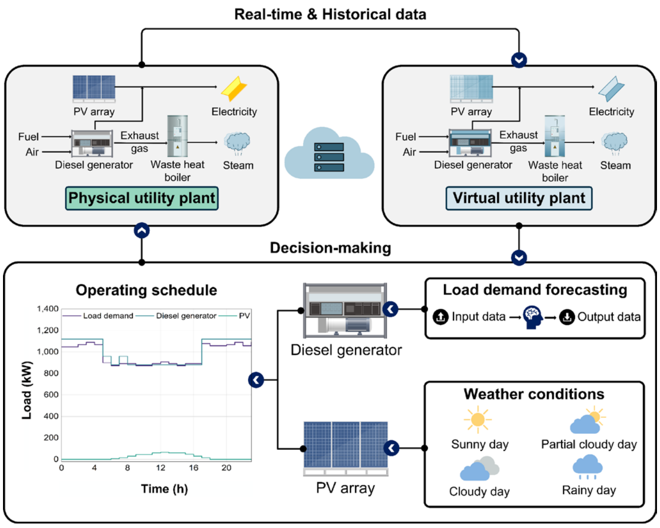

Figure 2. Illustration of digital twin for industrial utility management

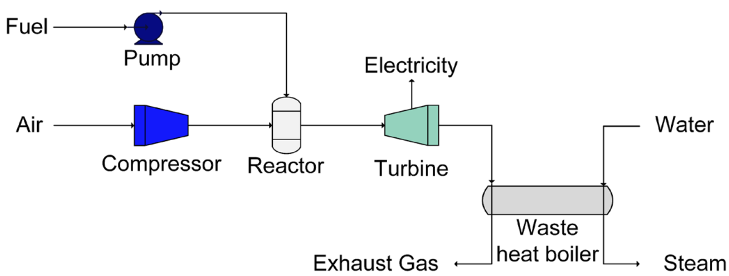

Figure 3. Process flow diagram of utility plant

2.2.2 Artificial intelligence-based load demand forecasting

Load forecasting involves the forecast of future load demand, which is helpful in energy management and improving decision support systems for utility management. Several energy management system decisions, including operation, scheduling, and planning, can depend on accurate load forecasting. Furthermore, weather factors can significantly impact load demand. All factors impacting load demand, including historical load and weather data, must be included as forecast model inputs to produce better forecasting results (Raza and Khosravi, 2015). Several techniques can be applied for load demand forecasting. The current research focused on artificial intelligence techniques, with machine learning being one of the most important applications of artificial intelligence in load demand forecasting, as shown in Figure 4. Machines are developed in such a way that they learn from data and get trained. Once trained based on the training data, the machine learning model can make forecasts on new examples of data that it has never seen before. The hourly load demand is the result of the load demand forecasting process. The estimated energy demand for every hour of the anticipated period is provided by these data. The efficient allocation of energy resources and the optimization of their production depend on accurate load demand forecasting.

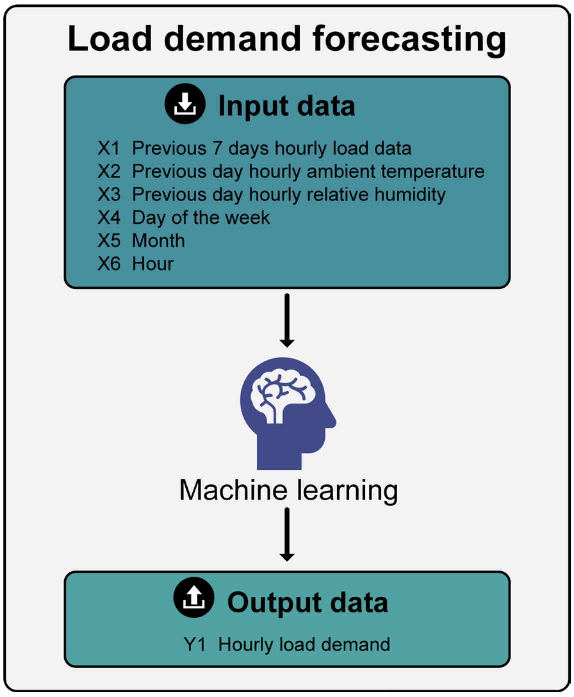

Figure 4. Schematic diagram of load demand forecasting

2.3 Modeling framework for load demand forecasting

Managing industrial utilities requires a framework for modeling load demand forecasts. By accurately forecasting future load demand, energy sources such as a diesel generator and PV arrays can be used efficiently to increase energy efficiency and lower costs (Kotb et al., 2021). In addition, operating schedules can be efficiently developed based on the utility's anticipated future load requirements to meet industry demands. The framework for load demand forecasting is depicted in Figure 5. The mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and R^2 are used as indicators to evaluate the accuracy of the model. Equations (1)–(4) present the formulas for these performance indicators. These indicators help assess how well the model is doing overall and how accurate it is at forecasting future load demand.

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}| \tag{3}$$

$$R^2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2}{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4}$$

where y is the mean value of the initial output value, $\bar{y}$ is the forecasted output value, and y_i is the actual output of the i[th] sample.

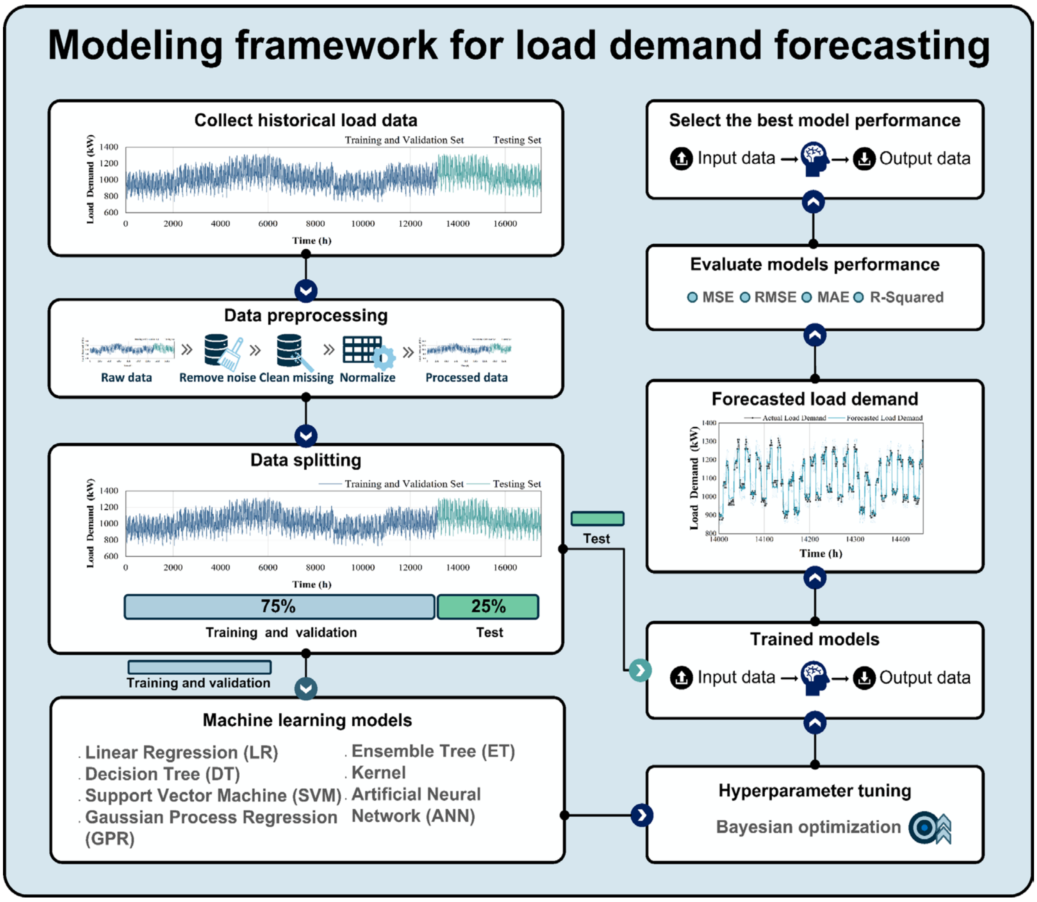

Figure 5. Schematic diagram of modeling framework for load demand forecasting

3. RESULT AND DISCUSSION

3.1 Modeling and forecasting

This study applied several different machine learning regression models for load demand forecasting. Table 1 shows that the ensemble tree (ET) model performed the best in terms of accuracy, with lower MSE, RMSE, and MAE values and a higher R^2 value. The ET model had MSE values of 535.550 and 759.440 for the validation and test datasets, respectively, with RMSE values of 23.164 and 27.558, respectively, and MAE values of 17.682 and 21.148, respectively. Furthermore, as shown in Figure 6, the R^2 values for the validation and test datasets were 0.96 and 0.95, respectively, showing a good fit between the forecast and actual load demands. Table 2 shows that the decision tree (DT) model has the fastest forecasting speed, with a rate of 94,000 obs/sec and a training time of 13.96 sec. Table 3 shows the optimal hyperparameters for each model, which might be helpful for future model development efforts. Although the DT model could forecast more rapidly and used fewer computing resources than the ET model, when selecting a forecasting model, it is critical to consider the tradeoff between forecast accuracy and resource consumption. Accuracy is the most important component in optimizing the use of utilities in an industrial system, which is required for generating an effective operating plan. As a result, choosing the ET model with better accuracy may be better for industrial applications.

Table 1. Model performance for load demand forecasting

Indicator		LR	DT	SVM	GPR	ET	Kernel	ANN
MSE	Validation	1096.200	767.890	707.320	630.230	535.550	1052.200	1258.900
	Test	1312.200	1175.900	1001.200	1024.300	759.440	1302.800	1420.400
RMSE	Validation	33.108	27.711	26.595	25.104	23.164	32.437	35.481
	Test	36.224	34.291	31.642	32.005	27.558	36.094	37.688
MAE	Validation	26.110	20.001	20.384	19.618	17.682	24.390	28.321
	Test	28.967	24.546	24.537	25.219	21.148	26.845	29.592
R^2	Validation	0.93	0.95	0.95	0.96	0.96	0.93	0.95
	Test	0.92	0.92	0.94	0.93	0.95	0.92	0.95

Table 2. Speed indicator for load demand forecasting

Model	Forecasting speed (obs/s)	Training time (s)
LR	25,000	17.94
DT	94,000	13.96
SVM	4,400	261.53
GPR	1,200	2282.30
ET	6,500	303.23
Kernel	25,000	2346.40
ANN	37,600	19.39

Table 3. Optimized hyperparameters of ensemble model

Hyperparameter	Range	Value
Ensemble method	{LSBoost, Bag}	Bag
Number of learners	{1–200}	10
Learning rate	{0.0001–1}	-
Minimum leaf size	{1–10,000}	3
Number of predictions to sample	{1–20}	1

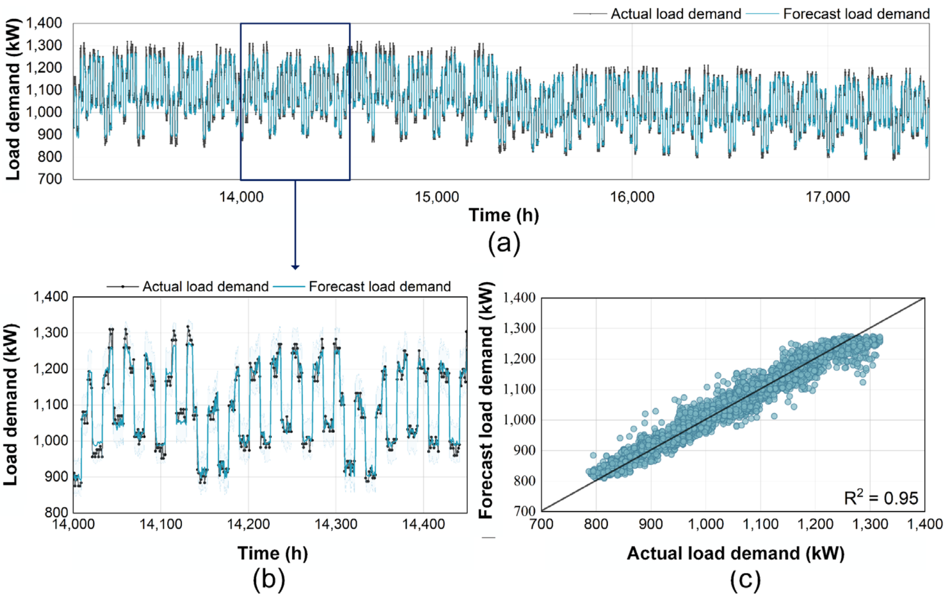

Figure 6. Results of ensemble forecasting model for test dataset: (a) all test data, (b) subset for 14,000–14,500 hours, and (c) R^2 value

3.2 Operating schedule results

The management of energy resources can be substantially influenced by the outcomes of the operating schedule. Accurate load demand forecasts are crucial for optimizing energy source usage, such as the diesel generator and PV arrays in the current study. The diesel generator had a 1,600 kW capacity and a power factor step size of 0.05, making it essential to operate it efficiently by utilizing precise load demand forecasts from the ET model. The schedule results also provided a detailed analysis of energy usage and demand, considering different weather scenarios such as sunny, cloudy, partially cloudy, and rainy days. The load demand forecasting outcomes enable utility management to make informed decisions regarding energy management and optimization. Analyzing these results allows for improved energy resource scheduling and planning, resulting in cost savings and efficient energy utilization.

The forecast load demand was compared between the ET model and the average load demand for the period 745–768 hours. The ET model used six inputs to forecast load demand: hourly load data from the prior seven days, ambient temperature, relative humidity, day of the week, month, and hour. On the other hand, the average load demand was forecast based on hourly load profiles for the same day of the week for four weeks. From Figure 7 (b–c), the ET model provided more accurate load demand forecasting than the average load demand because the former used more data inputs. Table 4 shows that the ET model had a total error of 23.86%, while the average load demand had a total error of 65.29%, as calculated using Equation 5:

$$\%\,Error = \frac{|Forecasted\ load\ demand - Actual\ load\ demand|}{Actual\ load\ demand} x100 \quad (5)$$

Thus, the ET model could be recommended for creating the operating schedule with weather conditions in four scenarios to optimize energy utilization.

Figure 8 shows the summary of the forecast load demand based on the likelihood of better energy management by integrating a diesel generator and PV arrays with hourly loads. The work schedule results are shown for (a) hourly PV generation in four scenarios: (b) a sunny day, (c) a cloudy day, (d) a partially cloudy day, and (e) a rainy day. The results from the operating schedule showed that the ET model improved energy utilization and cost savings by effectively balancing energy supply and demand for the different climatic conditions by aggregating load demand forecasts.

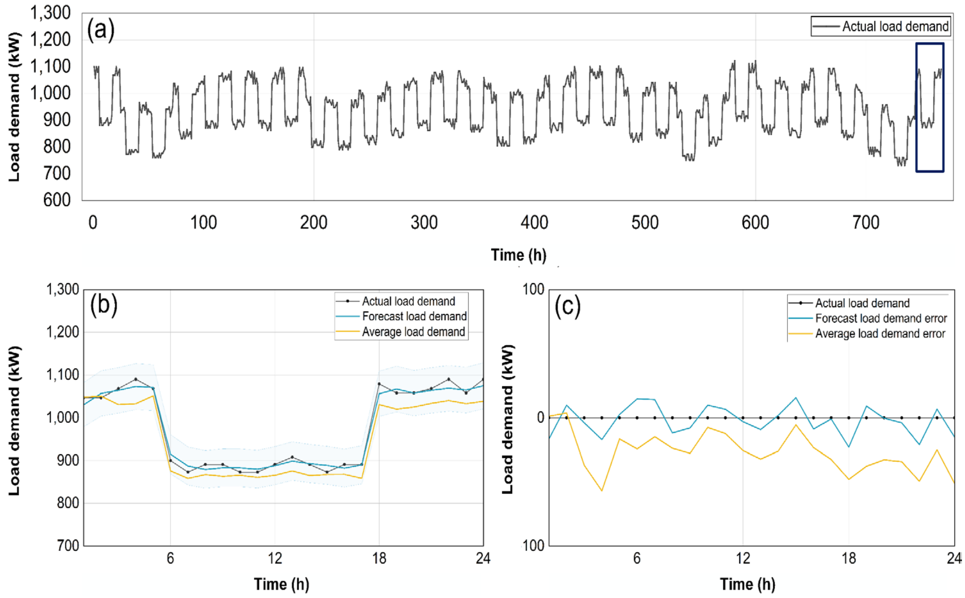

Figure 7. Results of forecast load demand using ET model compared to (a) actual load demand, (b) hourly load profiles, and (c) forecast load demand error

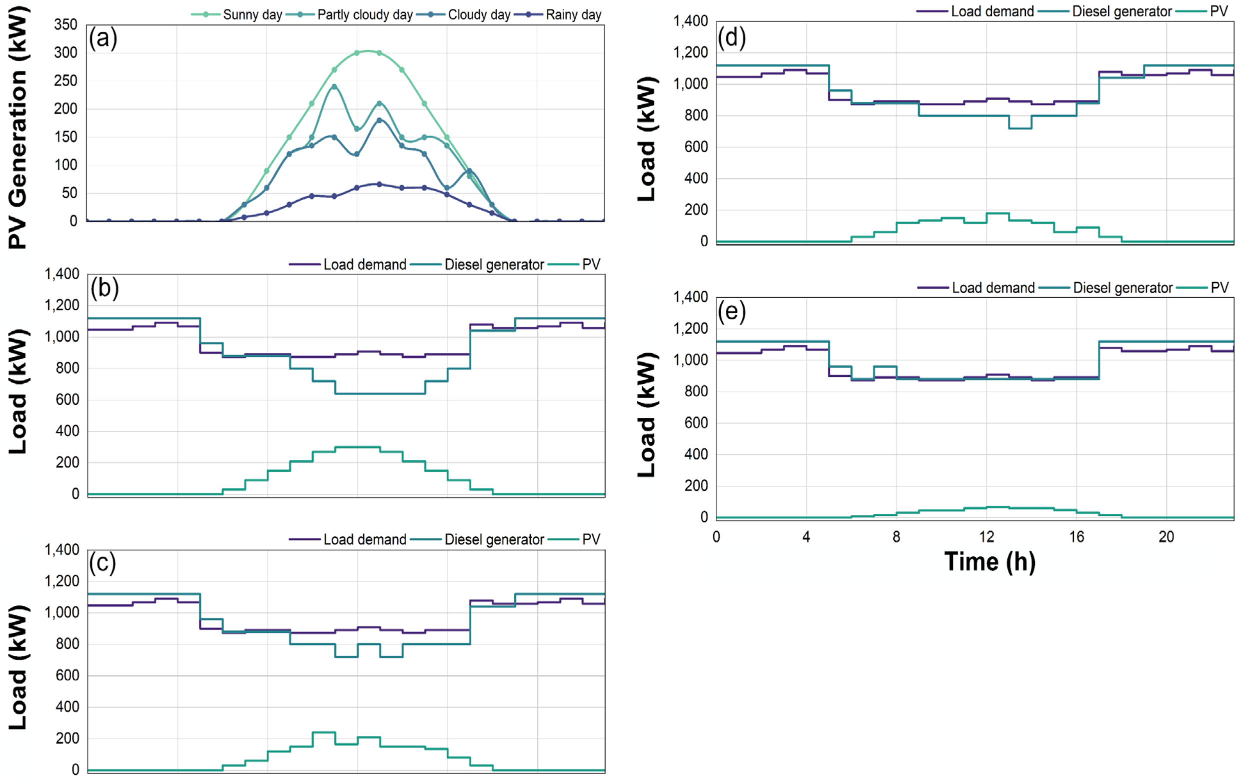

Figure 8. Results of operating schedule results presented for (a) hourly PV generation (four scenarios), (b) a sunny day, (c) a cloudy day, (d) a partially cloudy and (e) a rainy day

Table 4. Comparison between forecast load demand and average load demand

Time (h)	Load demand (kW)			% Error	
	Actual	Forecast	Average	Forecast	Average
1	1,047	1,030.493	1,048.250	1.577	0.119
2	1,047	1,056.865	1,050.750	0.942	0.358
3	1,068	1,064.096	1,031.000	0.366	3.464
4	1,090	1,073.020	1,033.000	1.558	5.229
5	1,068	1,070.619	1,051.500	0.245	1.545
6	900	914.771	875.750	1.641	2.694
7	873	887.213	858.250	1.628	1.690
8	891	879.205	867.250	1.324	2.666
9	891	883.136	863.250	0.883	3.114
10	873	882.879	865.500	1.132	0.859
11	873	879.719	860.750	0.770	1.403
12	891	887.958	865.500	0.341	2.862
13	908	898.803	875.750	1.013	3.552
14	891	892.915	865.000	0.215	2.918
15	873	888.763	867.500	1.806	0.630
16	891	882.257	867.750	0.981	2.609
17	891	890.086	858.500	0.103	3.648
18	1,079	1,056.096	1,031.000	2.123	4.449
19	1,058	1,067.168	1,020.250	0.867	3.568
20	1,058	1,057.683	1,025.250	0.030	3.095
21	1,068	1,064.106	1,033.750	0.365	3.207
22	1,090	1,068.981	1,040.500	1.928	4.541
23	1,058	1,064.740	1,033.000	0.637	2.363
24	1,090	1,074.854	1,038.750	1.390	4.702
Total	23,467	23,416.425	22,827.750	23.86	65.29

4. CONCLUSION

A methodology was proposed that utilizes digital twin technology and artificial intelligence-based load demand forecasting techniques for optimizing industrial systems. The primary contributions of this research were the use of a digital twin to simulate and analyze system behaviors and the development of an operating schedule that predicts future utility demand using load demand forecasting for the prior seven days, ambient temperature, relative humidity, days of the week, month, and hour. The ET model provided the highest accuracy based on a comparison with several machine learning regression models. Additionally, the digital twin and load demand forecasting approaches could be valuable tools to support effective decision-making for industrial utility management when creating the operating schedule. The ET model had a total error of 23.86%, while the average load demand approach had a total error of 65.29%. Therefore, the ET model with weather conditions in four scenarios is recommended for use in developing the operating schedule to maximize energy efficiency.

ACKNOWLEDGMENT

The Department of Chemical Engineering, Faculty of Engineering, Kasetsart University, Bangkok, Thailand provided funding support. Honeywell and GRD Tech Co., Ltd. provided the UniSim Design Suite simulation software.

REFERENCES

Ahmad, T., Chen, H., and Shah, W. A. (2019). Effective bulk energy consumption control and management for power utilities using artificial intelligence techniques under conventional and renewable energy resources. *International Journal of Electrical Power & Energy Systems,* 109, 242–258.

Aziz, A. S., Tajuddin, M. F. N., Zidane, T. E. K., Su, C.-L., Alrubaie, A. J. K., and Alwazzan, M. J. (2022). Techno-economic and environmental evaluation of PV/diesel/battery hybrid energy system using improved dispatch strategy. *Energy Reports,* 8, 6794–6814.

Chauhan, A., Upadhyay, S., Khan, M. T., Hussain, S. M. S., and Ustun, T. S. (2021). Performance investigation of a solar photovoltaic/diesel generator based hybrid system with cycle charging strategy using BBO algorithm. *Sustainability,* 13(14), 8048.

Chen, Y., Yang, O., Sampat, C., Bhalode, P., Ramachandran, R., and Ierapetritou, M. (2020). Digital twins in pharmaceutical and biopharmaceutical manufacturing: A literature review. *Processes,* 8(9), 1088.

Hasan, A., Hu, Z., Haghshenas, A., Karlsen, A., Alaliyat, S., and Cali, U. (2023). An interactive digital twin platform for offshore wind farms development. In *Digital Twin Driven Intelligent Systems and Emerging Metaverse* (Karaarslan, E., Aydin, Ö., Cali, Ü., and Challenger, M., Eds.), pp. 269–281. Gateway East, Singapore: Springer Nature.

Kotb, K. M., Elkadeem, M. R., Khalil, A., Imam, S. M., Hamada, M. A., Sharshir, S. W., and Dán, A. (2021). A fuzzy decision-making model for optimal design of solar, wind, diesel-based RO desalination integrating flow-battery and pumped-hydro storage: Case study in Baltim, Egypt. *Energy Conversion and Management,* 235, 113962.

Lan, H., Zhang, C., Hong, Y. Y., He, Y., and Wen, S. (2019). Day-ahead spatiotemporal solar irradiation forecasting using frequency-based hybrid principal component analysis and neural network. *Applied Energy,* 247, 389–402.

Nam, K., Hwangbo, S., and Yoo, C. (2020). A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. *Renewable and Sustainable Energy Reviews,* 122, 109725.

Panjapornpon, C., Bardeeniz, S., and Hussain, M. A. (2023a). Improving energy efficiency prediction under aberrant measurement using deep compensation networks: A case study of petrochemical process. *Energy,* 263(C), 125837.

Panjapornpon, C., Bardeeniz, S., Hussain, M. A., Vongvirat, K., and Chuay-ock, C. (2023b). Energy efficiency and savings analysis with multirate sampling for petrochemical process using convolutional neural network-based transfer learning. *Energy and AI,* 14, 100258.

Rahbar, K., Mahmoud, S., Al-Dadah, R. K., Moazami, N., and Mirhadizadeh, S. A. (2017). Review of organic Rankine cycle for small-scale applications. *Energy Conversion and Management,* 134, 135–155.

Raza, M. Q., and Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews,* 50, 1352–1372.

Roslan, M. F., Hannan, M. A., Ker, P. J., Begum, R. A., Mahlia, T. I., and Dong, Z. Y. (2021). Scheduling controller for microgrids energy management system using optimization algorithm in achieving cost saving and emission reduction. *Applied Energy,* 292, 116883.

Sleiti, A. K., Kapat, J. S., and Vesely, L. (2022). Digital twin in energy industry: Proposed robust digital twin for power plant and other complex capital-intensive large engineering systems. *Energy Reports,* 8, 3704–3726.

Yu, W., Patros, P., Young, B., Klinac, E., and Walmsley, T. G. (2022). Energy digital twin technology for industrial energy management: Classification, challenges and future. *Renewable and Sustainable Energy Reviews,* 161, 112407.