Predicting Minute-Level Power Grid Conditions Using Generative Adversarial Networks

Umme Mahbuba Nabila,[1] Gavin McDuffee,[2] Xingang Zhao,[3,4] William L. Gurecky,[4] Pradeep Ramuhalli,[4] and Majdi I. Radaideh[1,*]

[1]*Department of Nuclear Engineering and Radiological Sciences, University of Michigan, Ann Arbor, MI 48109, USA*

[2]*Department of Computer Science, Tennessee Tech University, Cookeville, TN 38505, USA*

[3]*Department of Nuclear Engineering, University of Tennessee, Knoxville, TN 37996, USA*

[4]*Nuclear Energy and Fuel Cycle Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA*

[leave space for DOI, which will be inserted by ANS]

ABSTRACT

As renewable energy systems are increasingly integrated into grid operations, optimal and flexible use of nuclear energy resources in such integrated setups requires accurate forecasting of grid demand and power produced by renewable energy systems. Precise prediction of contributions from renewable energy sources is vital for planning nuclear-provided generation in a future system where nuclear energy is expected to be operated flexibly. This becomes highly crucial for ensuring power grid stability in the face of increasing power demand, most notably from recent and upcoming high electricity consumption from data centers. This study investigated generative adversarial networks (GANs) to predict minute-level power grid data, focusing on solar power, wind power, and load power in California Independent System Operator (CAISO) Zone 1. Conditional GAN (cGAN), Wasserstein GAN (WGAN), Gated Recurrent Unit (GRU), and sequence-to-sequence (seq2seq) GRU were used to capture the complex temporal dependencies in the data. The study systematically evaluated model performance using various lookback (1 h, 12 h) and lookforward (1 h, 12 h) windows, assessing each model's accuracy using metrics such as the root mean square error and the mean absolute error. The results demonstrate that WGAN achieves higher accuracy than cGAN, GRU, and seq2seq GRU in almost all scenarios. This higher accuracy is attributable to WGAN's use of the Wasserstein distance as the loss function, which enhances stability and mitigates mode collapse. The results in this paper show the ability of generative models to generate accurate time series data that may be used to assist in optimizing nuclear energy generation in future grids.

Keywords: Generative Adversarial Network (GAN), Wasserstein GAN (WGAN), Gated Recurrent Unit (GRU), sequence-to-sequence (seq2seq) GRU, time series forecasting, power grid prediction

1. INTRODUCTION

As global advancement accelerates, the demand for electricity is growing at an unprecedented rate. Integrated energy systems (IESs) are essential in the modern era with high energy demand because they enable efficient, flexible, and sustainable management of diverse energy sources, thereby ensuring a reliable and balanced power supply. Predicting grid load is a critical component of managing IESs, which combine diverse energy sources, such as nuclear, solar, wind, and storage systems, into cohesive, optimized frameworks. In an IES setup, accurate forecasting of power demand is essential for ensuring that energy supply matches real-time demand, thereby minimizing waste and maintaining grid stability. This is particularly important for systems in which nuclear and renewable energy sources work together because these energy sources have fundamentally different operating characteristics [1]. Nuclear power—with its ability to provide steady, continuous output—acts as a baseload energy source, whereas renewable energy sources, such as solar and wind, are inherently intermittent and can vary significantly based on weather conditions. The complementary nature of nuclear and renewable energy sources within an IES highlights the importance of robust forecasting methods [2]. In the US, for example, Idaho National Laboratory is working to advance integrated energy generation, storage, and delivery technologies to support a net-zero future by developing multigeneration systems that seamlessly combine nuclear and renewable energy [3]. This research on predictive modeling for grid loads will enhance these efforts by providing accurate forecasts, thereby enabling better coordination of energy sources within these integrated systems.

Traditional methods for time series prediction, such as Autoregressive Integrated Moving Average (ARIMA), are effective but are limited by assumptions of linearity and stationarity. Recent advances in Machine Learning, particularly deep learning, have enhanced time series prediction by capturing complex, nonlinear patterns. Recurrent neural networks (RNNs) are powerful for sequence data because of their ability to model temporal dependencies, though they struggle with long-term dependencies because of the vanishing gradient problem. Enhanced architectures, such as Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM), address this issue. GRUs offers a simpler, faster structure and comparable performance to LSTMs, making GRUs ideal for time series tasks [4]. GRU was proposed by Cho et al. in 2014 [5]. Its effectiveness and efficiency were showcased by Chung et al. who explored the performance of GRUs against LSTMs across different sequence modeling tasks. They found that GRUs, with their simpler structure and fewer parameters, matched the performance of LSTMs while offering faster and more efficient training [6].

Generative AI models, particularly generative adversarial networks (GANs), are being widely explored for time series prediction tasks. GAN, a novel generative model based on minimax game theory, was introduced by Goodfellow in 2014 [7]. GAN has two networks: the generator and the discriminator. The generator creates data as realistic as possible to deceive the discriminator, which then attempts to distinguish fake samples from real ones. The training process continues until the generator succeeds in fooling the discriminator to no longer distinguish the fake data [8]. Many variants of GAN have been proposed, including conditional GAN (cGAN) and Wasserstein GAN (WGAN). cGAN implements class labels or labeled data to effectively execute specific tasks. WGAN addresses the instability of training traditional GANs by using the Wasserstein distance (Earth mover's distance) as a metric for measuring the divergence between the real and generated data distributions [9]. The efficacy of cGAN and WGAN for time series forecasting has been demonstrated in numerous studies, highlighting their potential to generate accurate predictions across various fields [10] [11].

Another powerful approach for time series prediction is the sequence-to-sequence (seq2seq) model, which leverages gated recurrent units (GRUs) for efficient handling of sequential data. Seq2seq models excel in capturing temporal dependencies by encoding input sequences into a fixed-length representation and decoding it to generate output sequences [12]. The GRU-based architecture offers a lightweight yet effective alternative to LSTMs, enabling faster training and reduced computational overhead while maintaining strong performance in time series forecasting tasks. This makes seq2seq GRU models a compelling choice for applications that require both accuracy and efficiency.

This paper aims to address the challenge of accurately forecasting renewable energy contributions to enable the optimal and flexible use of nuclear energy resources in IESs. By leveraging advanced generative models—including cGAN, WGAN, GRU, and seq2seq GRU—this work compares the ability of these models to capture complex temporal dependencies in time series data derived from the power-systems machine learning (PSML) dataset developed by Zheng et al. [13]. Alongside comparing four models, the effect of lookback and lookforward time window size on each model is also investigated.

The structure of the remaining sections of this paper is organized as follows. Section 2 describes the training data collection and preprocessing methods conducted for this study. The models implemented in this paper, along with the assessment cases, are presented in Section 3. Section 4 provides the major findings obtained from the models and the performance assessment of their data needs based on the power grid datasets. Finally, the conclusions drawn from this study are highlighted in Section 5.

2. DATA COLLECTION AND PROCESSING

The PSML dataset used in this analysis is described in this section. This multiscale time series dataset was developed by Zheng et al. by capturing real-world and synthetic time series data across various spatiotemporal scales, including minute-level and millisecond-level measurements from the electric grid [13]. Minute-level data were collected over a period of 3 years, from 2018 to 2020, across 66 regions in the US. California Independent System Operator (CAISO) Zone 1 was used for this study. A visualization of feature trends over the time period selected for this forecasting study is provided in Figure 1. The dataset includes key features such as the following:

- **Load power**: the total electrical power consumed in the load zones.
- **Wind power**: the electrical power generated by wind turbines.
- **Solar power**: the power output from solar systems.
- **Weather-related data**: parameters such as diffuse horizontal irradiance (DHI), direct normal irradiance (DNI), global horizontal irradiance (GHI), dew point, solar zenith angle, wind speed, relative humidity, and temperature.

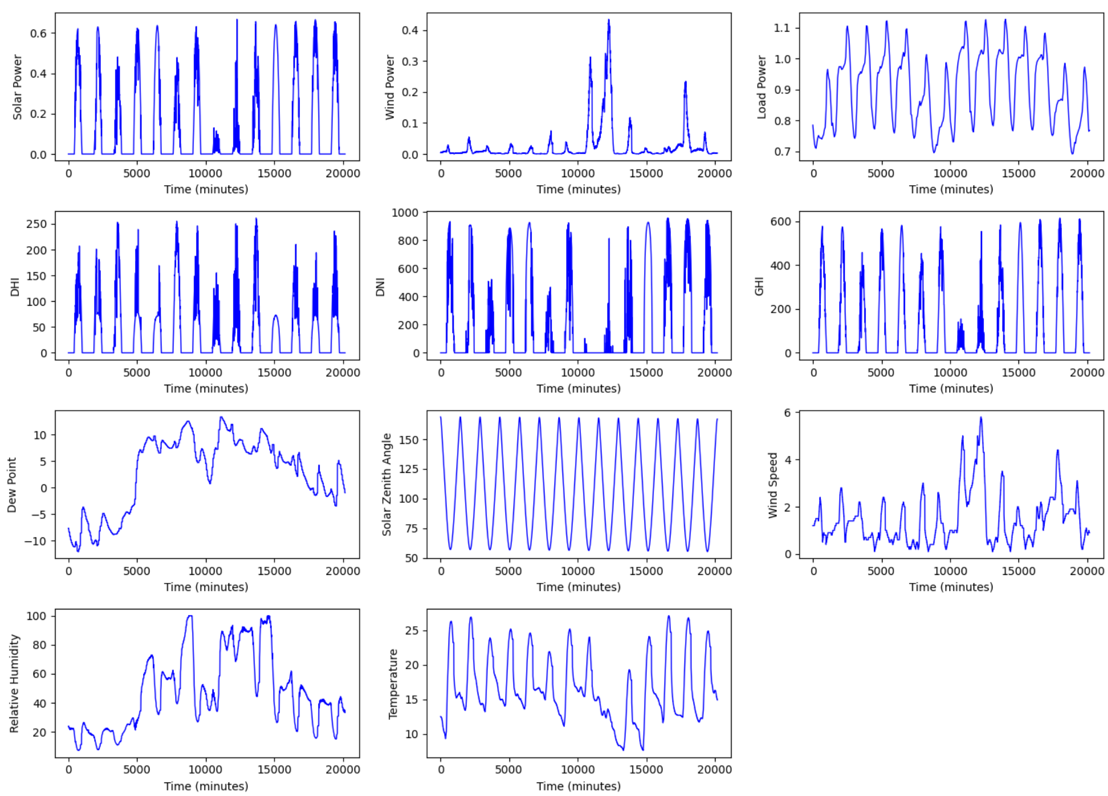

Figure 1: Variation of weather-related and load variables for a selected time period in CAISO Zone 1.

3. METHODOLOGY

3.1. Generative Adversarial Networks

GANs are a class of ML models that consist of two neural networks: the generator $\mathcal{G}$ and the discriminator $\mathcal{D}$. The generator takes a random noise vector z sampled from a prior distribution $p_z(z)$ as input and produces synthetic data samples $\hat{x} = \mathcal{G}(z)$. The discriminator, on the other hand, aims to distinguish between real data samples x from the true data distribution $p_{\text{data}}(x)$ and fake samples $\hat{x}$ produced by the generator [7]. The objective of the generator is to maximize the probability that the discriminator misclassifies its generated samples as real.

$$\min_{\mathcal{G}} \max_{\mathcal{D}} V(\mathcal{D}, \mathcal{G}) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log \mathcal{D}(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - \mathcal{D}(\mathcal{G}(z)))], \tag{1}$$

where $V(\mathcal{D}, \mathcal{G})$ represents the value function or the adversarial loss. Here, the first term in the objective function represents the log probability that the discriminator correctly classifies real data as real, and the second term represents the log probability that the discriminator incorrectly classifies fake data as real.

3.1.1. Conditional Generative Adversarial Network

Unlike traditional GAN, cGANs are not completely unsupervised in their training methods. Instead, cGAN architectures implement class labels or labeled data to effectively execute specific tasks. To create a cGAN structure, a small adjustment is applied to the original GAN architecture by adding a "y-label" to both the discriminator and generator networks. This adjustment transforms the previous probabilities into conditional probabilities [14]. With this modification, the training process guarantees that the generator produces outputs aligned with the specified labels, which are provided as conditions. Similarly, the discriminator scrutinizes the authenticity of the generated output, ensuring that it matches the expected label. Consequently, after training, providing a specific input label yields the desired output from the generative network. Figure 2 illustrates a typical cGAN structure.

$$\min_{\mathcal{G}} \max_{\mathcal{D}} V(\mathcal{D}, \mathcal{G}) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log \mathcal{D}(x|y)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - \mathcal{D}(\mathcal{G}(z|y)))] \tag{2}$$

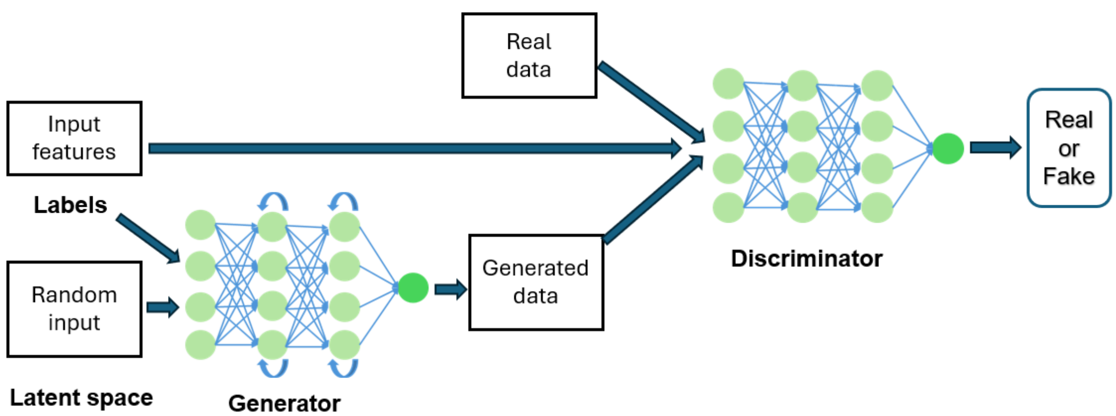

Figure 2: A typical cGAN structure.

3.1.2. Wasserstein Generative Adversarial Network

Traditional cGANs use the binary cross-entropy (BCE) loss mentioned in the above formulas, Eq. (1) and Eq. (2). The BCE loss is effective for distinguishing between real and fake data samples. However, the BCE loss can cause training instability and mode collapse, particularly when minimal overlap exists between the real and generated data distributions. WGAN addresses the limitations of BCE loss by replacing it with the Wasserstein distance, or Earth mover's distance, as a measure of divergence between the real and generated data distributions [9]. Instead of training the discriminator to classify samples as real or fake, the discriminator (also known as the "critic" in WGAN) is trained to provide a continuous score that estimates how “real” a sample is. The critic network is designed to assign higher scores to real samples and lower scores to generated samples, effectively learning to approximate the Wasserstein distance between real data and the generated data distribution. The generator network is trained to minimize this distance by producing samples that the critic scores similarly to real data, thereby reducing the gap between the two distributions. WGAN reformulates the cGAN loss function to use the expected values of the critic outputs for real samples and generated samples, as shown in Eq. (3):

$$\min_{\mathcal{G}} \max_{\mathcal{D}} V(\mathcal{D}, \mathcal{G}) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\mathcal{D}(x|y)] - \mathbb{E}_{z \sim p_z(z)}[\mathcal{D}(\mathcal{G}(z|y))] \tag{3}$$

where $\mathcal{D}$ is a k-Lipschitz function, which is essential for accurately approximating the Wasserstein distance. WGAN enforces this property by applying a weight-clipping constraint on the critic's parameters within a bounded interval $[-c, c]$, typically with $c = 0.01$. This modification eliminates the need for logarithmic terms in the loss function and thereby enhances training stability.

3.2. Gated Recurrent Unit

GRU is an advanced RNN designed to address the vanishing gradient problem in standard RNNs. It incorporates update and reset gates to selectively retain or discard information, effectively preserving relevant past information across long sequences. The following equations govern the GRU cell:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{4}$$
$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{5}$$
$$\tilde{h}_t = \tanh(W_h x_t + r_t \odot (U_h h_{t-1}) + b_h) \tag{6}$$
$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{7}$$

where: x_t is the input at time step t, h_{t-1} is the hidden state from the previous time step, z_t and r_t are the update and reset gates, respectively, $\tilde{h}_t$ is the candidate hidden state, σ denotes the sigmoid activation function, $W_z, W_r, W_h, U_z, U_r, U_h$ are the weight matrices, and b_z, b_r, b_h are the bias terms.

3.3. Sequence-to-Sequence Gated Recurrent Unit

Consisting of an encoder and a decoder both based on GRUs (as shown in Figure 3), the seq2seq GRU is an extension of GRU for time series forecasting. The encoder processes historical inputs and outputs, encoding temporal dependencies into a fixed-length hidden state vector that serves as a context representation. The decoder uses this context to sequentially generate future predictions. This cohesive approach improves the seq2seq GRU's performance in long-term predictions, making it particularly well suited for tasks such as power grid load forecasting, where capturing complex temporal dependencies and providing reliable bulk predictions are essential.

3.4. Data Assessment Cases

Table 1 outlines the three evaluation cases used to assess the performance of GAN models in predicting power grid demand. These cases strategically explore the effects of lookback and lookforward window sizes on model performance. All three cases are trained on the first 2 weeks of the dataset, and the test dataset is selected as the following 10 days.

The performance of all models is evaluated using mean absolute error (MAE) and root mean square error (RMSE). These metrics provide insights into the accuracy and generalization capability of the models.

- **MAE**: measures the average absolute difference between predicted and actual values.
- **RMSE**: similar to MAE, RMSE measures model prediction error but amplifies the effect of larger errors because of the squaring of the differences between model predictions and actual values.

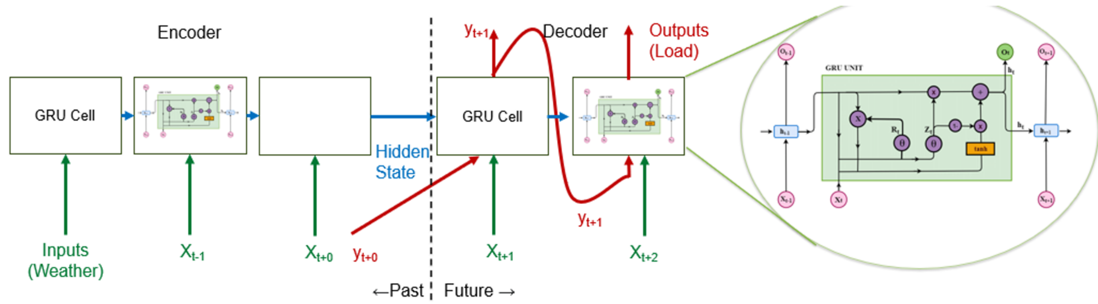

Figure 3: Structure of seq2seq GRU.

Table 1: Description of the test cases used to evaluate model performance in power grid demand

Case	**Description**
Case 1	Lookback is 1 h, and lookforward is 1 h
Case 2	Lookback is 12 h, and lookforward is 1 h
Case 3	Lookback is 12 h, and lookforward is 12 h

4. RESULT AND DISCUSSION

Figure 4 provides a visual comparison of the predictive performance of cGAN, WGAN, GRU and seq2seq GRU for solar, wind, and load power under the condition of 12 h lookback and 12 h lookforward time windows. WGAN performed the best for load power by closely following the power trend with time. All the models performed well for solar power prediction but failed to capture the peaks in the wind power trend. Overall, solar power prediction is the most accurate, which can be explained by the plethora of available input features dedicated to it (e.g, DHI, DNI, GHI, solar zenith angle, and temperature). There are a few input features directly related to wind power—wind speed, relative humidity, and dew point in particular. However, load power lacks any input feature that directly accounts for its transient. This may explain the models' poor performance in predicting the load power.

Table 2: Performance comparison for Case 1 in Table 1

	RMSE				**MAE**			
	cGAN	**WGAN**	**GRU**	**Seq2seq**	**cGAN**	**WGAN**	**GRU**	**Seq2seq**
Wind	0.029	0.029	0.022	0.038	0.018	0.017	0.013	0.029
Solar	0.080	0.074	0.079	0.101	0.049	0.043	0.053	0.069
Load	0.087	0.082	0.096	0.098	0.067	0.064	0.078	0.075

The performance of the models for Case 1 with lookback of 1 h and lookforward of 1 h (as summarized in Table 2) reveals the following key observations: GRU exhibits the best performance for wind power prediction, achieving the lowest RMSE (0.022) and MAE (0.013). Both cGAN and WGAN perform similarly,

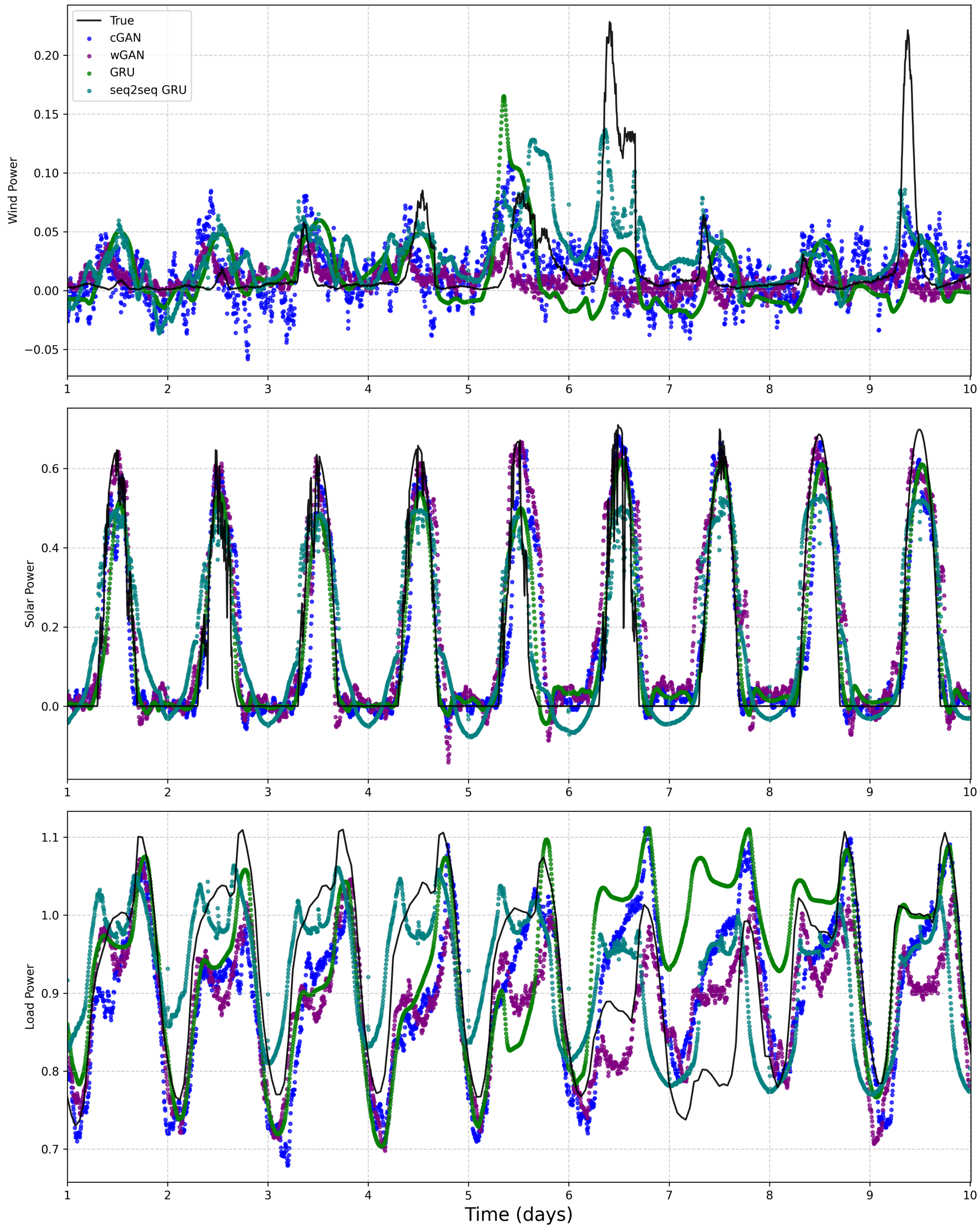

Figure 4: Comparison of power grid prediction using cGAN, WGAN, GRU, and seq2seq GRU.

with identical RMSE values (0.029). However, WGAN achieves slightly better MAE (0.017) compared with that of cGAN (0.018). The seq2seq model has the highest RMSE (0.038) and MAE (0.029) among the four models, indicating lower precision in predicting wind power. For solar power prediction, WGAN demonstrates the best overall performance, achieving the lowest RMSE (0.074) and MAE (0.043). Once again, WGAN outperforms the other models, achieving the lowest RMSE (0.082) and MAE (0.064) for load power. Figure 5 visually illustrates the RMSE and MAE results for Case 1.

Table 3 presents the performance comparison for Case 2 with lookback of 12 h and lookforward of 1 h. GRU achieves the best performance for both solar and wind power predictions, whereas GANs perform the

Table 3: Performance comparison for Case 2 in Table 1

	RMSE				MAE			
	cGAN	**WGAN**	**GRU**	**Seq2seq**	**cGAN**	**WGAN**	**GRU**	**Seq2seq**
Wind	0.030	0.031	0.022	0.045	0.019	0.022	0.012	0.028
Solar	0.074	0.068	0.061	0.127	0.047	0.046	0.041	0.108
Load	0.071	0.077	0.085	0.095	0.058	0.063	0.073	0.078

Table 4: Performance comparison for Case 3 in Table 1

	RMSE				MAE			
	cGAN	**WGAN**	**GRU**	**Seq2seq**	**cGAN**	**WGAN**	**GRU**	**Seq2seq**
Wind	0.035	0.032	0.032	0.035	0.023	0.019	0.022	0.023
Solar	0.080	0.071	0.070	0.067	0.049	0.042	0.045	0.045
Load	0.082	0.070	0.084	0.116	0.065	0.054	0.064	0.085

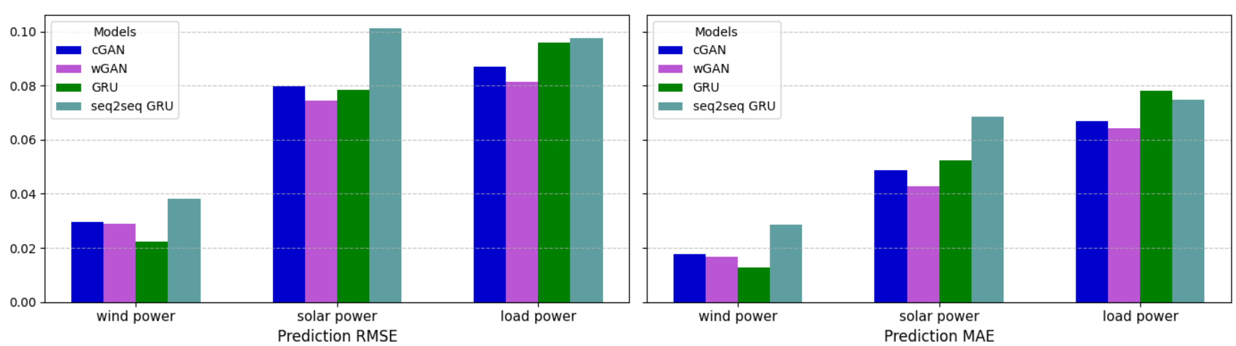

Figure 5: Comparison of RMSE and MAE for Case 1 in Table 1.

best for load power prediction. For better comprehension, the comparison is also demonstrated in Figure 6.

Table 4 summarizes findings for Case 3 with lookback of 12 h and lookforward of 12 h. A clear visual comparison of model performances is shown in Figure 7. For wind power prediction, WGAN achieves the best performance with the lowest RMSE (0.032) and lowest MAE (0.019), closely followed by GRU. cGAN and seq2seq GRU show a slightly higher RMSE (0.035), indicating comparable but slightly less accurate predictions than those of WGAN and GRU. For solar power, the seq2seq GRU model achieves the lowest RMSE (0.067), highlighting its performance improvements with higher lookback and lookforward. Finally, for load power, WGAN achieves the best performance with the lowest RMSE (0.070) and MAE (0.054), showcasing its robustness and stability in predicting load power.

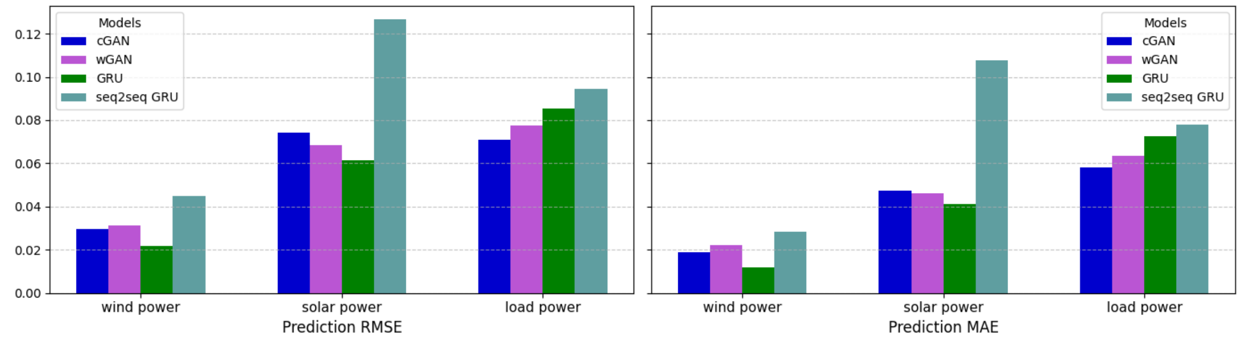

Figure 6: Comparison of RMSE and MAE for Case 2 in Table 1.

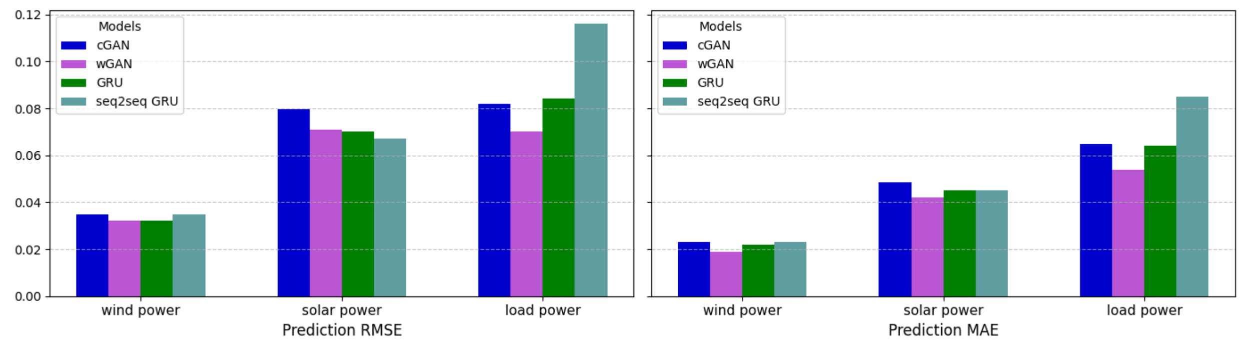

Figure 7: Comparison of RMSE and MAE for Case 3 in Table 1.

5. CONCLUSIONS

This study evaluated the performance of cGAN, WGAN, GRU, and seq2seq GRU models across three cases with varying lookback and lookforward intervals for predicting wind power, solar power, and load power. The results reveal distinct strengths and limitations of each model, offering insights into each model's suitability for different prediction tasks and temporal resolutions.

For Case 1 (lookback of 1 h, lookforward of 1 h), the GRU model emerged as the best performer for wind power prediction, achieving the lowest RMSE (0.022) and MAE (0.013). By contrast, WGAN demonstrated superior performance for solar power (RMSE: 0.074, MAE: 0.043) and load power (RMSE: 0.082, MAE: 0.064), showcasing its robustness with a shorter window. Although cGAN performed similarly to WGAN for wind power, seq2seq GRU exhibited the highest error metrics, highlighting its limited efficacy for short-term prediction tasks. In Case 2 (lookback of 12 h, lookforward of 1 h), the GRU model again excelled in wind and solar power predictions, reflecting its strong ability to capture sequential dependencies. However, GAN-based models, particularly WGAN, outperformed GRU for load power prediction, achieving the lowest RMSE and MAE. In Case 3 (lookback of 12 h, lookforward of 12 h), WGAN achieved the best performance for wind power prediction (RMSE: 0.032, MAE: 0.019).

Overall, the findings indicate that GRU is highly effective for short-term wind and solar power predictions. WGAN consistently excels in load power predictions and performs competitively across all tasks, showcasing its robustness and stability. Seq2seq GRU shows notable improvements with longer lookback and lookforward intervals, particularly for solar power prediction. cGAN is moderately effective but generally lags behind WGAN.

ACKNOWLEDGMENTS

This work was supported by the AI Initiative as part of the Laboratory Directed Research and Development (LDRD) program of Oak Ridge National Laboratory, which is managed by UT-Battelle LLC for the US Department of Energy under contract DE-AC05-00OR22725.

REFERENCES

[1] S. M. Bragg-Sitton, R. Boardman, C. Rabiti, and J. O'Brien. "Reimagining future energy systems: Overview of the US program to maximize energy utilization via integrated nuclear-renewable energy systems." *International Journal of Energy Research*, **volume 44**(10), pp. 8156–8169 (2020).

[2] A. I. Arvanitidis, V. Agarwal, and M. Alamaniotis. "Nuclear-Driven Integrated Energy Systems: A State-of-the-Art Review." *Energies*, **volume 16**(11), p. 4293 (2023).

[3] S. M. Bragg-Sitton. "Summary of INL Integrated Energy Systems Research for the Global National Laboratories Consortium on IES." Technical report, Idaho National Laboratory (INL), Idaho Falls, ID (United States) (2023).

[4] R. Cahuantzi, X. Chen, and S. Güttel. "A comparison of LSTM and GRU networks for learning symbolic sequences." In *Science and Information Conference*, pp. 771–785. Springer (2023).

[5] K. Cho. "On the Properties of Neural Machine Translation: Encoder-decoder Approaches." *arXiv preprint arXiv:14091259* (2014).

[6] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio. "Empirical evaluation of gated recurrent neural networks on sequence modeling." *arXiv preprint arXiv:14123555* (2014).

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. "Generative adversarial nets." *Advances in neural information processing systems*, **volume 27** (2014).

[8] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye. "A review on generative adversarial networks: Algorithms, theory, and applications." *IEEE transactions on knowledge and data engineering*, **volume 35**(4), pp. 3313–3332 (2021).

[9] M. Arjovsky, S. Chintala, and L. Bottou. "Wasserstein generative adversarial networks." In *International conference on machine learning*, pp. 214–223. PMLR (2017).

[10] X. Gu, K. See, Y. Liu, B. Arshad, L. Zhao, and Y. Wang. "A time-series Wasserstein GAN method for state-of-charge estimation of lithium-ion batteries." *Journal of Power Sources*, **volume 581**, p. 233472 (2023).

[11] S. Festag, J. Denzler, and C. Spreckelsen. "Generative adversarial networks for biomedical time series forecasting and imputation." *Journal of Biomedical Informatics*, **volume 129**, p. 104058 (2022).

[12] G. Gong, X. An, N. K. Mahato, S. Sun, S. Chen, and Y. Wen. "Research on short-term load prediction based on Seq2seq model." *Energies*, **volume 12**(16), p. 3199 (2019).

[13] X. Zheng, N. Xu, L. Trinh, D. Wu, T. Huang, S. Sivaranjani, Y. Liu, and L. Xie. "A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids." *Scientific Data*, **volume 9**(1), p. 359 (2022).

[14] M. Mirza. "Conditional generative adversarial nets." *arXiv preprint arXiv:14111784* (2014).