5 Results

5.1 Model Architecture

5.1.1 Conditional Variational Autoencoder (CVAE)

The core of our model is based on the Conditional Variational Autoencoder (CVAE) architecture. CVAEs extend the capabilities of traditional Variational Autoencoders (VAEs) by conditioning the encoding and decoding processes on additional information. In our case, this allows the model to generate data under specific conditions, such as time and tariff rates, improving the relevance and accuracy of the generated load profiles.

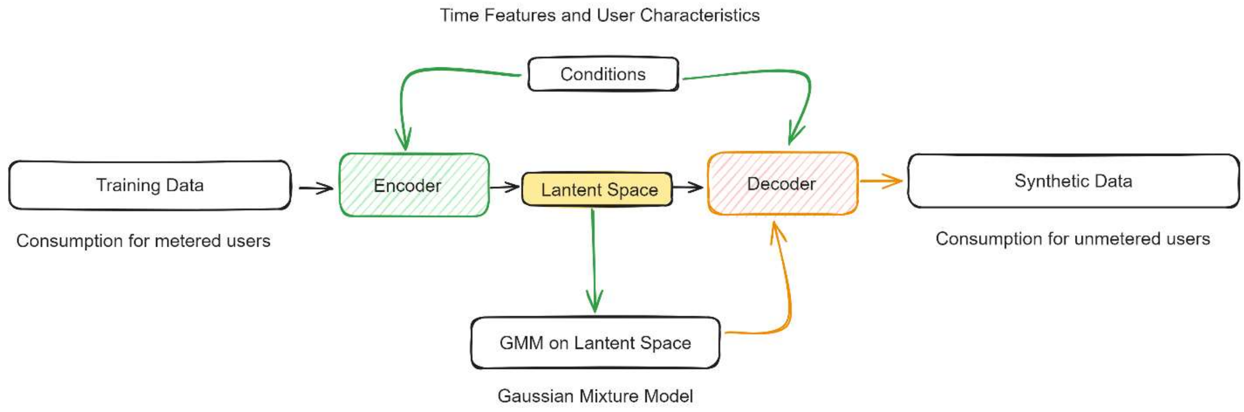


Figure 1 Conditional Variational Autoencoder (CVAE) with GMM fitted on the latent space

The CVAE consists of two main components:

Encoder: Transforms the input data (load profiles) and conditional information into a latent space representation.

Decoder: Reconstructs the load profiles from the latent space representation, conditioned on the same information used in the encoding process.

5.1.2 Modifications to the Base CVAE

To enhance the performance and adaptability of our model, we introduced several modifications to the base CVAE architecture:

Mean Maximum Discrepancy (MMD) Loss: In addition to the traditional Kullback-Leibler (KL) Divergence loss, we incorporated MMD loss to better handle non-normal and skewed data distributions. This helps in capturing the complex patterns present in electricity consumption data.

Gaussian Mixture Model (GMM) Sampling: We modeled the latent space using a Gaussian Mixture Model to better capture its distribution. This allows for more nuanced and realistic sampling from the latent space during the generation process.

5.1.3 Data Source

For this study, we utilized the afore mentioned dataset provided by Alliander NV, containing anonymized historical electricity consumption and generation data from approximately 17,000 industrial and commercial users for the year 2023. The data, recorded at a 15-minute resolution in kilowatts (kW), directly reflects power consumption over each interval.

Data Processing Pipeline

1. **Data Splitting:** We created a dataset of 300 randomly sampled users. An 80/20 split was performed for training and testing.
2. **Data Preprocessing:** and Merging Tariff Data: We loaded merged data, obtained high/low tariff and timestamp information, converted timestamps to match the dataset format, and merged new features with the main data.
3. **Adding Time-Based Features:** Features such as hour, day of the week, month, weekend indicator, time interval, and day of the year were extracted.

4. **Computing and Merging Consumption Statistics:** Monthly total, average, and maximum consumption, as well as yearly maximum and total consumption, were computed and merged back into the main dataset
5. B**inning and Encoding:** The binning function was applied to each group. One-hot encoding was applied to create new columns for binned values and baseload profiles.

5.2 Experiments

Our experimentation process involved an exploration of various model configurations and techniques to optimize performance:

5.2.1 Layer Variations

We experimented with different layer configurations to improve the model's ability to capture complex patterns in the load profiles:

1. **LSTM layers**: Incorporated to model long-term dependencies in the time series data. LSTM (Long Short-Term Memory) layers are particularly useful for time-series data as they can capture temporal dependencies and patterns over long periods.

2. **Attention layers and positional encoding**: Implemented to improve the model's ability to focus on relevant parts of the input sequence and capture position-dependent information. Attention mechanisms allow the model to weigh various parts of the input sequence differently, thereby focusing more on the important time steps. Positional encoding helps the model to understand the order of the input sequence, which is crucial for time-series data.

5.2.2 Normalization Techniques

We explored two main normalization approaches to prepare the input data:

1. **Global Max Normalization**: Normalizing the entire dataset based on the maximum value observed across all users and time periods.
2. **Localized Monthly Max Normalization:** Normalizing the data on a month-by-month basis, allowing for better capturing of seasonal variations.

5.2.3 Sampling Techniques

To generate diverse and realistic load profiles, we investigated several sampling techniques from the latent space:

1. **GMM Simple Sampling**: Direct sampling from the fitted Gaussian Mixture Model. This method samples directly from the learned mixture of Gaussians, ensuring that the generated samples reflect the underlying data distribution.
2. **Conditional Sampling from GMM**: Sampling from the GMM conditioned on specific input features. This technique allows us to generate samples that meet certain conditions, such as specific time periods or tariff rates, making the generated profiles more relevant to specific scenarios.
3. **Importance Sampling from GMM**: Using importance sampling to generate more representative samples from the GMM. Importance sampling is a method that adjusts the sampling process to give more weight to certain samples, improving the representativeness and quality of the generated profiles.

5.3 Training Process

The training process for our CVAE model involved several steps to ensure effective learning and generalization. We utilized the Adam optimizer with a learning rate of 1e-4, and incorporated a mixed precision training approach to leverage GPU capabilities efficiently. The model was trained over 20 epochs with a batch size of 128, allowing it to learn the patterns present in the electricity consumption data.

The training process was carefully monitored to ensure effective learning and model convergence:

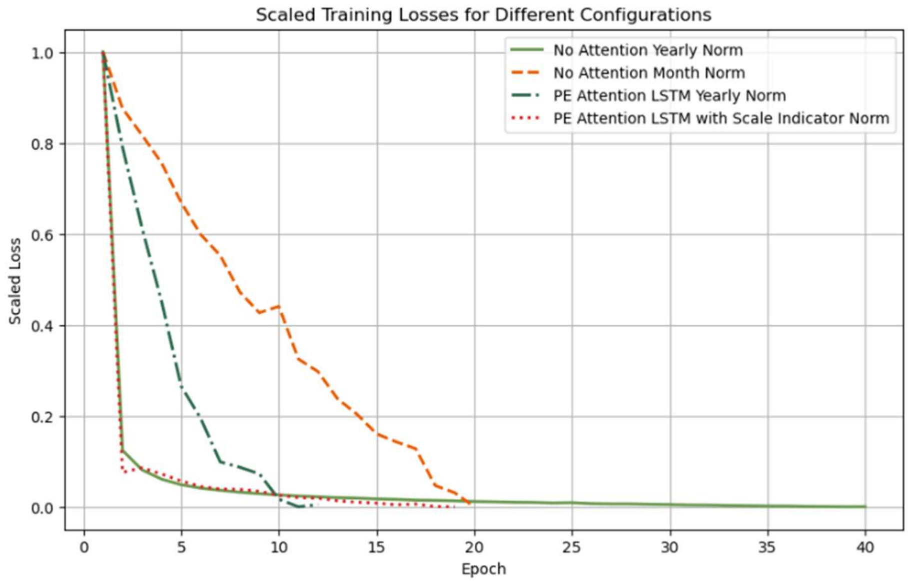


Figure 2 Training Losses - Training losses over epochs, showing steady reduction in both reconstruction and KL divergence losses

Key observations from the training process:

1. **Effective Learning**: The steady reduction in loss values indicates that the model was effectively learning to encode and decode the load profiles.
2. **Convergence**: The gradual reduction in loss values suggests that the model was converging to a stable state, balancing between reconstruction accuracy and latent space regularization.

5.4 Results and Analysis

5.4.1 Sanity Check: Training Reconstruction

Our initial validation involved reconstructing the training data using only the conditional information. This step was crucial to validate that our model could accurately reconstruct known data under specified conditions.

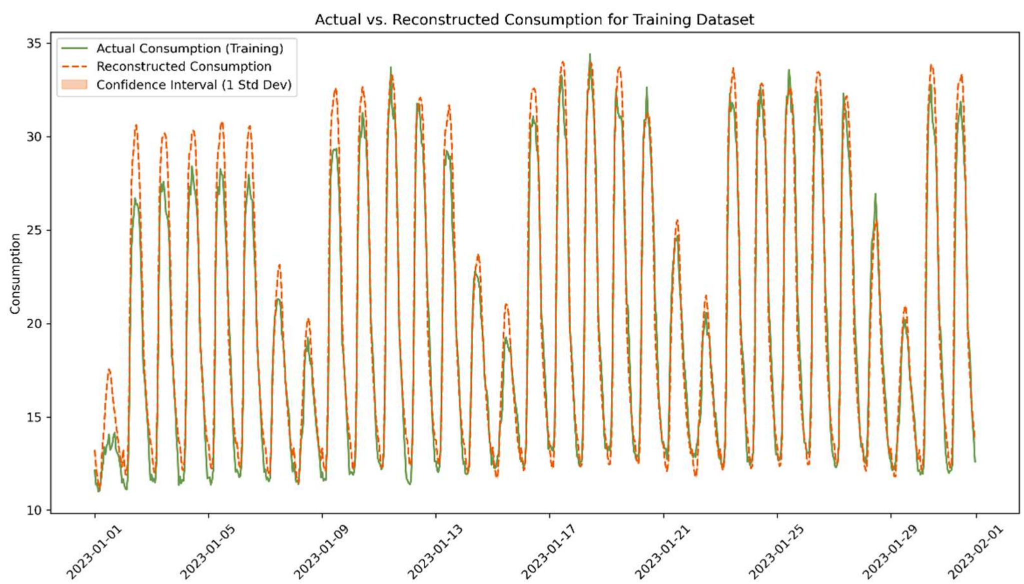


Figure 3 Training Reconstruction - Comparison of original and reconstructed load profiles from the training set.

Figure 2: Training Reconstruction - Comparison of original and reconstructed load profiles from the training set.

Performance metrics:

1. Mean Absolute Error (MAE): 0.99
2. Mean Squared Error (MSE): 1.64

These results validate that the model can perform well on seen conditions, accurately reconstructing the training data.

5.4.2 Reconstruction of Test Data - Global Year Max Normalization Results

We evaluated the model's performance on the test data by comparing the reconstructed profiles with the actual consumption data. We observed that the model captured the general trend of the consumption data quite well but struggled to accurately represent the magnitude of consumption. This discrepancy was particularly notable in the timing and scale of the reconstructed data.

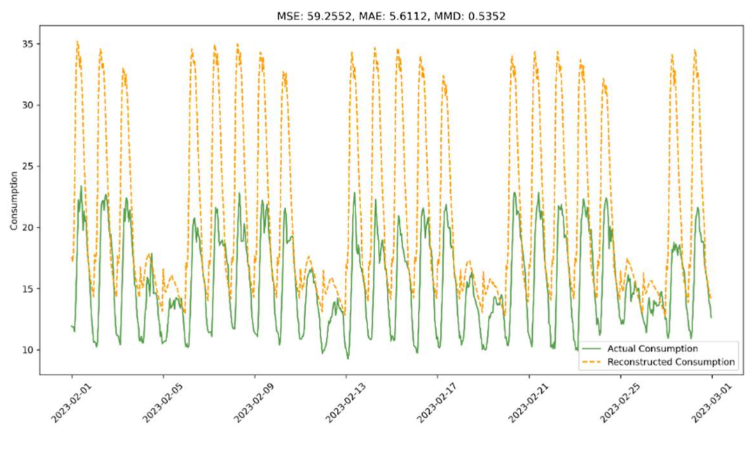


Figure 4 Global Year Max Normalization - Overall - Comparison of test data and generated profiles using global year max normalization.

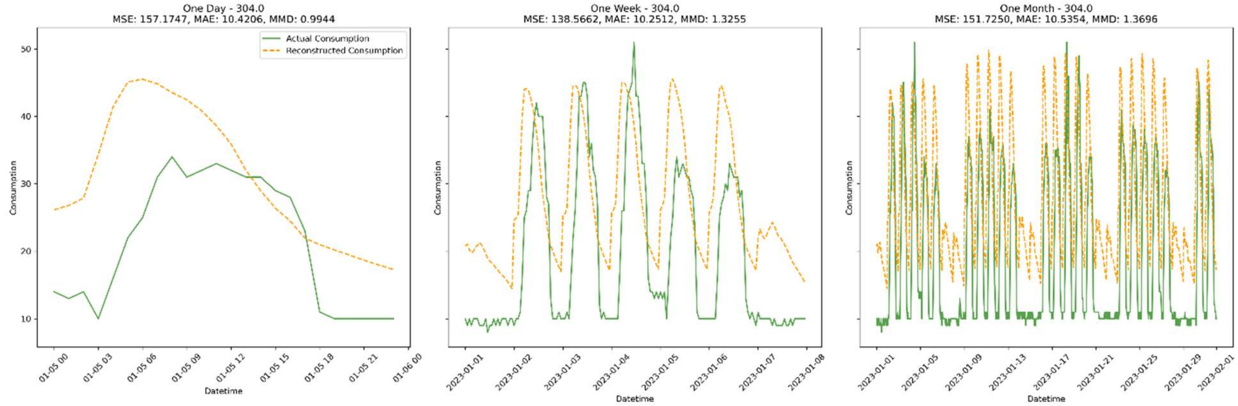


Figure 5 Global Year Max Normalization – Three Levels - Comparison of test data and generated profiles using global year max normalization.

Observations:

1. The model captures the overall trend of the test data.
2. However, it struggles to accurately represent the magnitude of consumption.

5.4.2.1 Customer Size Analysis

We analyzed the model's performance across different customer sizes by separating larger users (with average consumption above 40 kW) from smaller users. The results showed that for daily normalization:

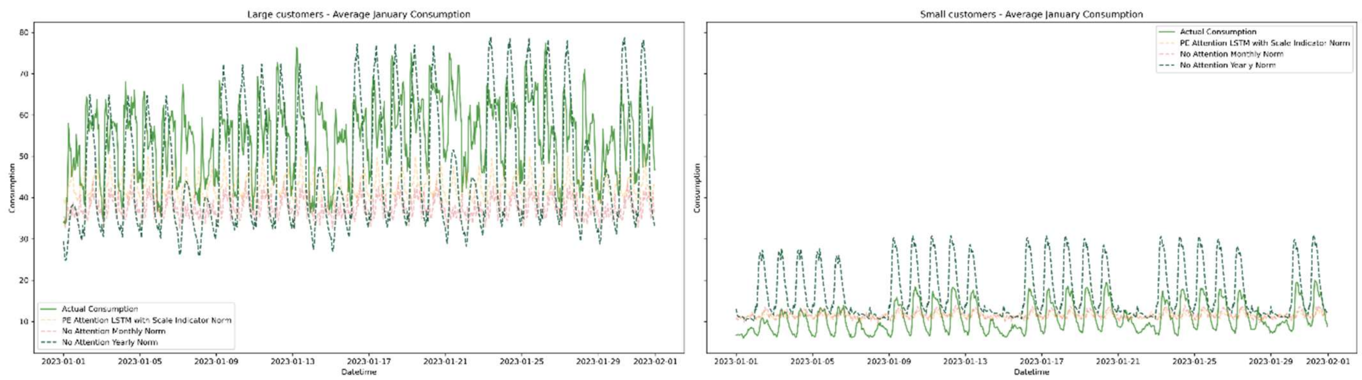


Figure 6 Global Year Max Normalization - Large vs Small Customers - Comparison of generated profiles for large (>40 kW) and small customers.

Observations:

1. Large Customers: The model performed well in capturing the magnitude but struggled with the trend.
2. Small Customers: The model captured the trend accurately but was off in magnitude, suggesting an upward bias in the training dataset.

5.4.3 Monthly Max Normalization and Attention Results

We then examined the results of the models using monthly max normalization and attention mechanisms:

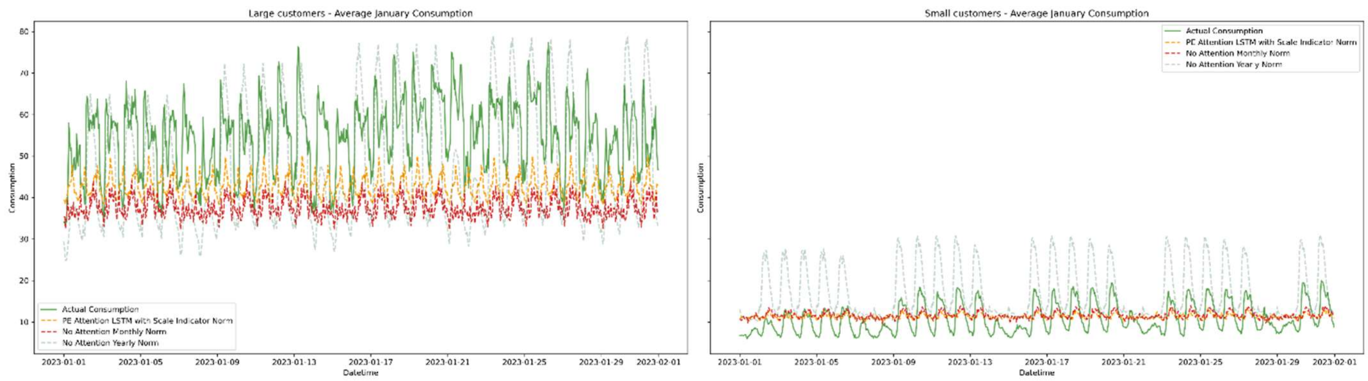


Figure 7 Monthly Max Normalization with Attention - Comparison of test data and generated profiles using monthly max normalization and attention mechanisms.

Observations:

1. Improved performance in capturing both trends but lower performance in the magnitudes of consumption when compared to global normalization.
2. The model both over and under predicted for smaller customers and

5.4.4 Daily Analysis

To gain deeper insights, we performed a daily analysis of the generated profiles:

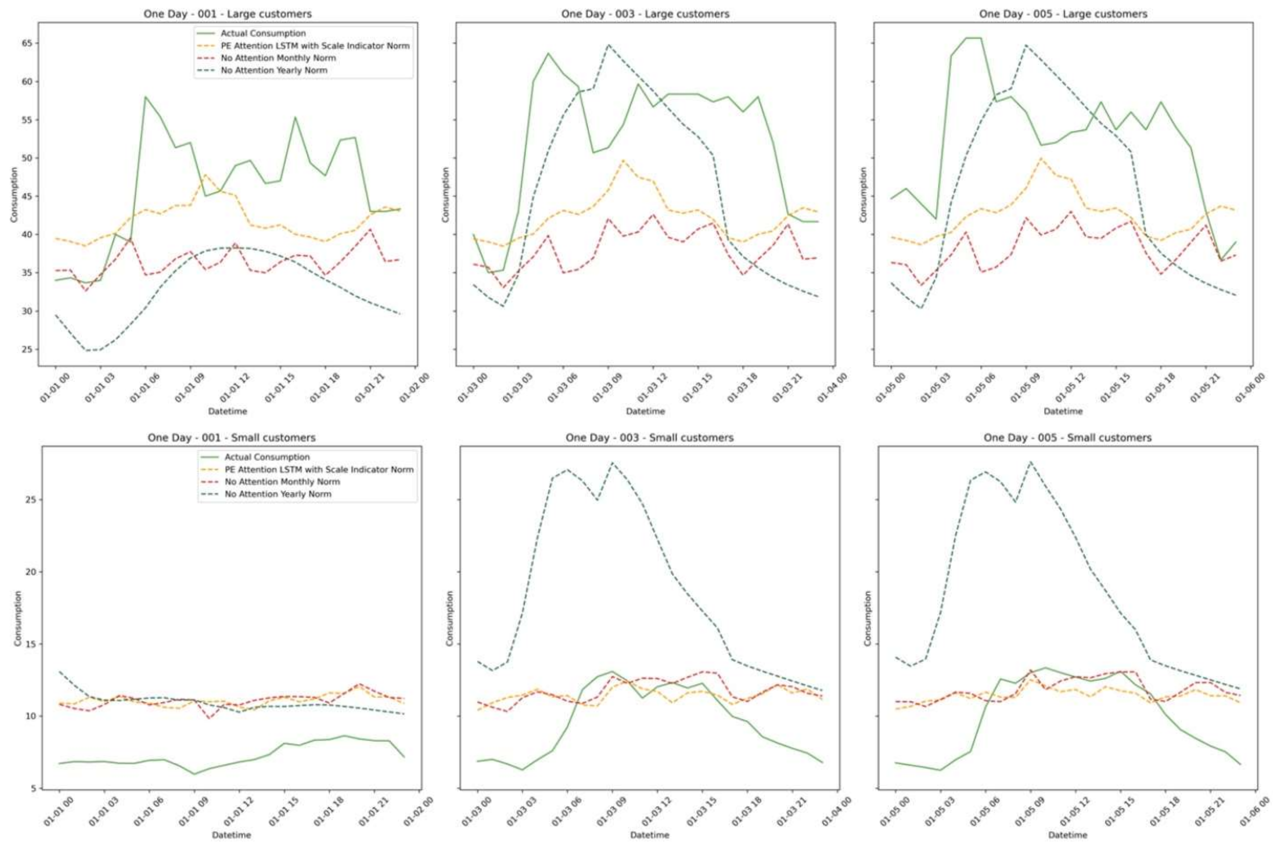


Figure 8 Daily Comparison - Large vs Small Customers - Daily comparison of generated profiles for large and small customers.

Observations:

1. Large customers: The model captured the magnitude better but struggled with the trend.
2. Small customers: The model captured the trend well but not the magnitude.

This analysis also supports earlier observations that yearly normalization without attention provided the most visually acceptable results.

5.4.5 Error Analysis

5.4.5.1 Comparison of Architectures

We conducted a comprehensive error analysis to understand the model's performance. When comparing different architectures, the results showed that the no-attention model with yearly normalization performed worse in terms of magnitudes, which was evident from the error metrics. This was primarily due to the model's inability to capture the correct magnitude, leading to higher overall errors. This

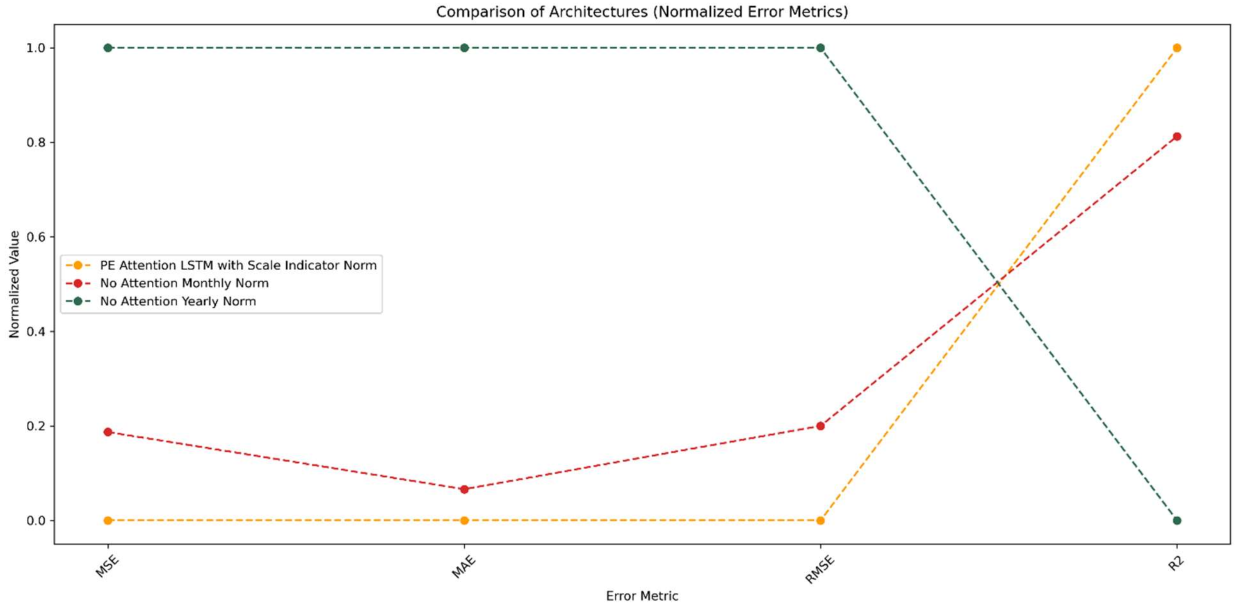


Figure 9 Error Analysis - Error metrics for different model configurations.

5.4.5.2 Comparison of Sampling Techniques

We compared various sampling techniques, including normal sampling, importance sampling, and conditional sampling. Conditional sampling consistently outperformed the others, suggesting that conditioning the sampling process on specific input features helps the model generate more accurate and representative profiles.

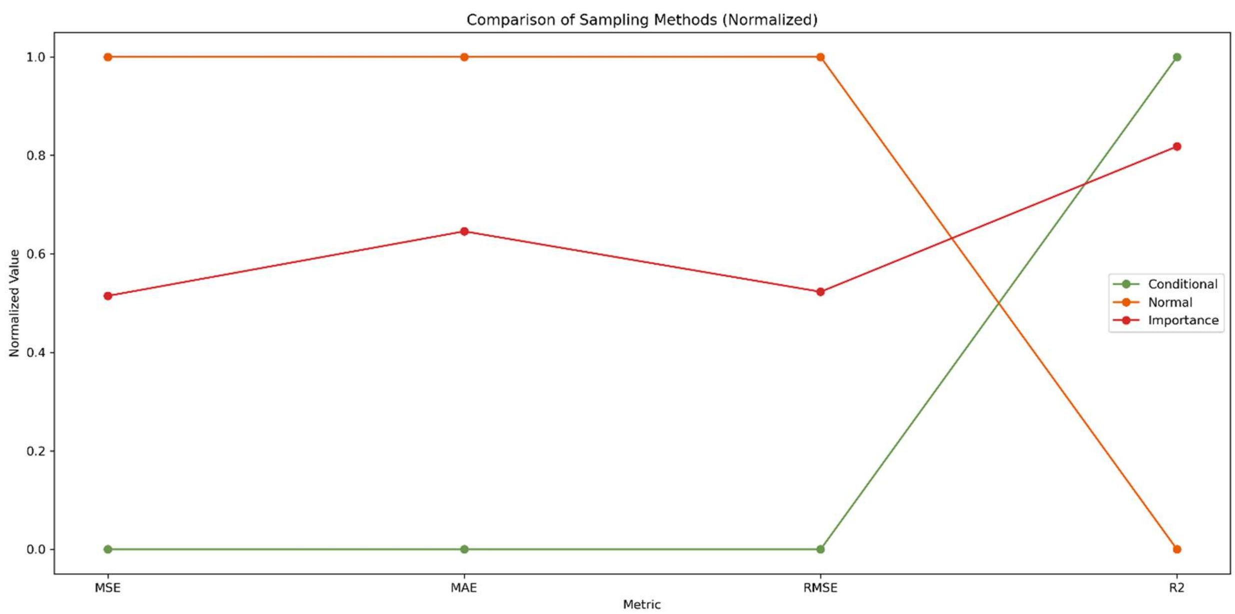


Figure 10 Error Analysis - Error metrics for different latent sampling techniques.

5.4.5.3 Smaller Customers vs. Larger Customers

To understand the model's behavior, we analyzed the training data, looking at smaller and larger customers separately:

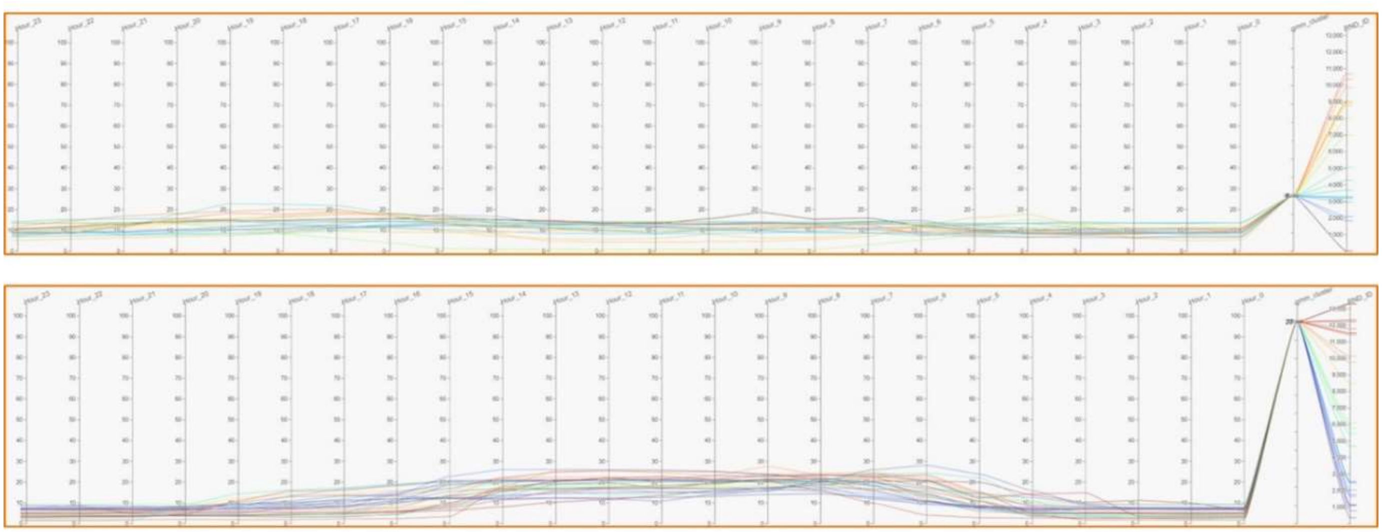


Figure 11Smaller customer segments

Smaller customers: Due to the higher number of smaller customers in the dataset and their varied consumption patterns, the model could capture the different distributions better.

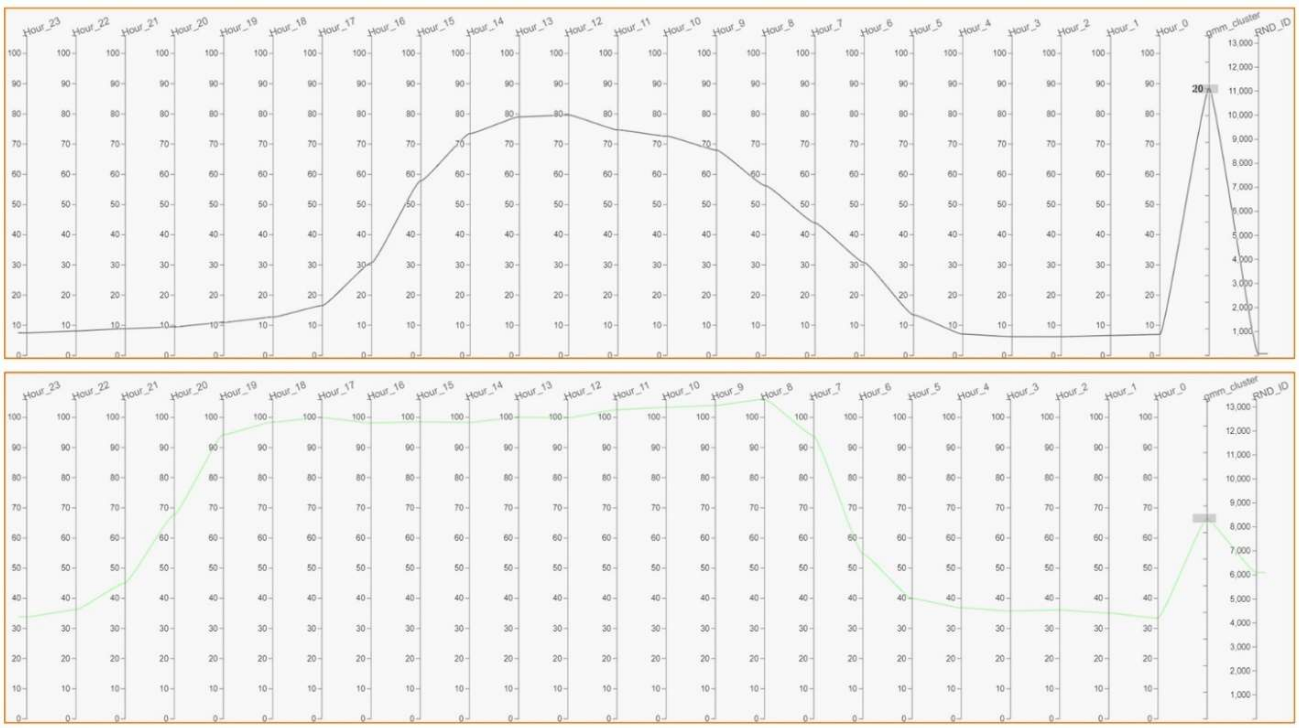


Figure 12 Larger customer segments (the bottom is in light green)

Larger customers: The smaller number of larger customers led to less variation in their patterns, which, combined with the global normalization, biased the model upwards and reduced its ability to capture trends accurately.

This analysis indicates that a balanced dataset with equal representation of different customer sizes could improve the model's performance by allowing it to learn the nuances of each customer segment more effectively.

6 Conclusion

Based on our extensive experimentation and analysis, we have drawn several key conclusions regarding the performance and potential of our CVAE model in generating high-fidelity synthetic load profiles:

Model Performance and Data Fidelity:

The CVAE model demonstrates significant promise in generating realistic synthetic load profiles, particularly in reconstructing training data. Our results indicate that the reconstructed training data closely follows actual consumption trends, validating the model's capability. This aligns with findings from Wang et al. (2022), where contextual load profiles were effectively generated using a combination of operational autoencoders.

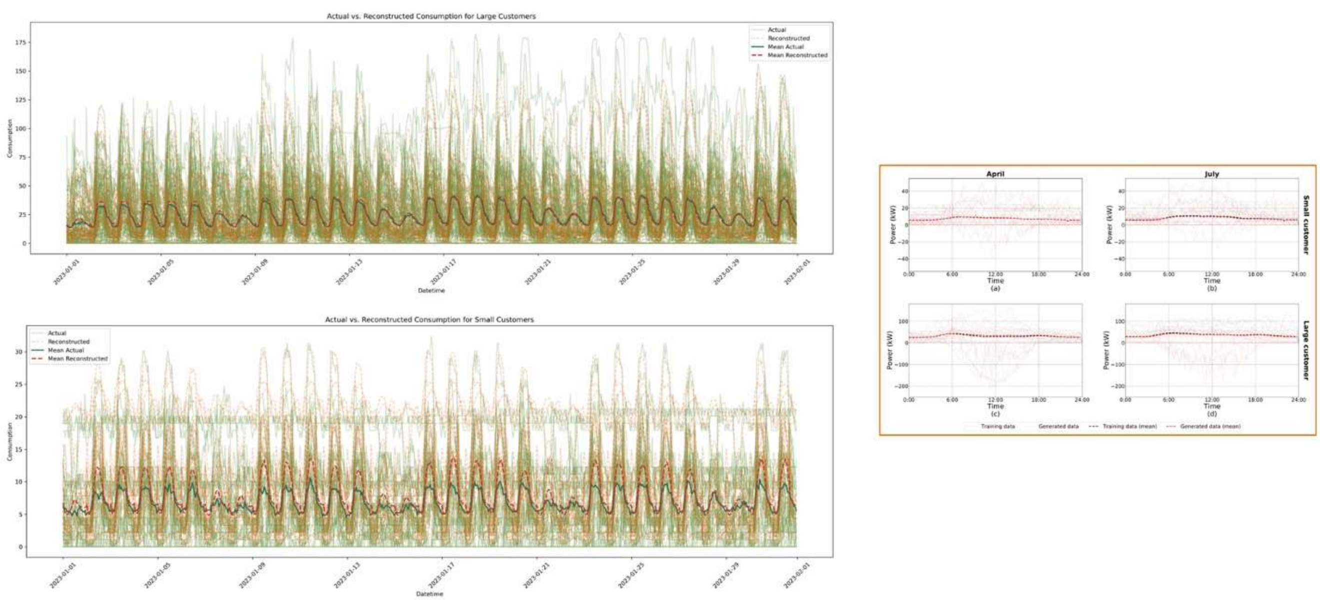


Figure 13 Promise in generating realistic synthetic load profiles

In the context of test data reconstruction, while the model captures the general trend, it struggles with accurately representing the magnitude of consumption, particularly for smaller customers. However, when compared to the Faraday paper (Doe et al., 2023), our model performed reasonably well (albeit worse than their results) despite utilizing significantly less data and not incorporating household characteristics.

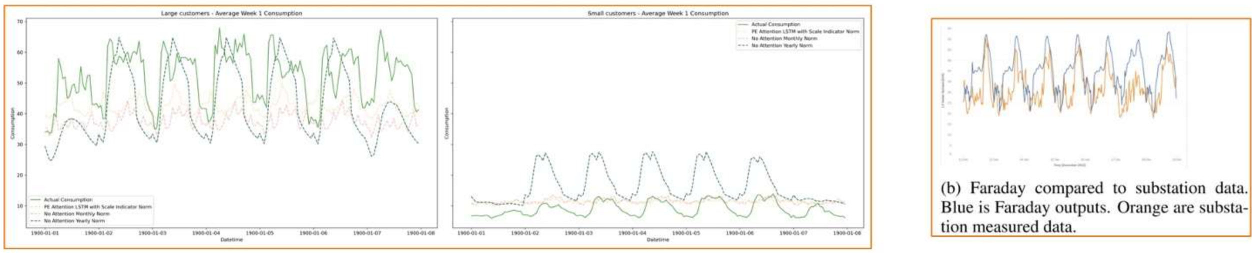


Figure 14 Comparison to Faraday

Normalization Techniques:

Monthly Max normalization generally succeeded in capturing trends effectively. However, Global Year Max normalization produced more visually realistic load profiles on average. This suggests that while both techniques have merits, Global Year Max normalization might be better suited for our specific use case.

Customer Size Performance Variability:

There was a noticeable difference in performance between larger and smaller customers under Global Year Max normalization. Larger customers exhibited better magnitude capture, whereas smaller customers showed better trend capture. This indicates a need for more specialized models, potentially fine-tuned for specific customer segments.

Attention Mechanisms:

The attention mechanism did not significantly enhance the model's ability to capture sequence dependencies. Given the complexity and extended training time required, we did not fully explore its potential. Although it performed best in error metrics, the visual examination did not support its effectiveness, suggesting that further investigation is needed.

Sampling Techniques:

Conditional sampling using Gaussian Mixture Models (GMM) yielded the best results among the tested sampling techniques. This highlights the importance of sophisticated sampling strategies in improving model performance.

7 Recommendation

To build on our findings and enhance the performance of our CVAE model, we propose the following recommendations:

Expand Data Size and Diversity:

Training the model architecture provided on a more balanced and larger dataset would likely improve the model's ability to capture nuanced differences between larger and smaller customers. This includes ensuring an equal representation of different customer sizes to avoid biases in normalization.

Explore Different Normalization Techniques:

While Global Year Max normalization showed promise, other techniques might yield better results. Future studies should explore a variety of normalization methods to identify the most effective approach for different datasets and use cases.

Develop Specialized Models:

Creating separate models for small and large users could address their unique consumption characteristics more effectively. This specialized approach would allow for tailored training and optimization, improving overall performance.

Investigate Advanced Sampling Techniques:

Further research into advanced sampling methods, such as important sampling and other conditional techniques, could enhance the model's ability to generate realistic load profiles. These techniques can help ensure that the training data is representative of all customer segments, leading to better generalization.

8 References

1. Wang, C., Tindemans, S. H., & Palensky, P. (2022). Generating Contextual Load Profiles Using a Conditional Variational Autoencoder. Retrieved from https://arxiv.org/abs/2209.04056v2

2. Chai, S., & Chadney, G. (2024). FARADAY: Synthetic Smart Meter Generator for the Smart Grid. Tackling Climate Change with Machine Learning, ICLR 2024. Retrieved from https://arxiv.org/abs/2404.04314v1

3. Bossmann, T. (2015). Unravelling load patterns of residential end-uses from smart meter data. Retrieved from https://www.researchgate.net/publication/324359967_Unravelling_load_patterns_of_residential_end-uses_from_smart_meter_data

4. Chun Fu, Hussain Kazmi, Matias Quintana, Clayton Miller. (2024). Creating synthetic energy meter data using conditional diffusion and building metadata. Retrieved from https://arxiv.org/abs/2404.00525v1

9 Annexure

9.1 New larger and more balanced dataset

To address the recommendation of expanding data size and diversity, we increased the dataset size from 300 to 500 users. The new dataset includes an equal representation of larger and smaller customers, ensuring a balanced distribution across four quantiles of average consumption. This balanced dataset allows for more robust and generalized model training and evaluation.

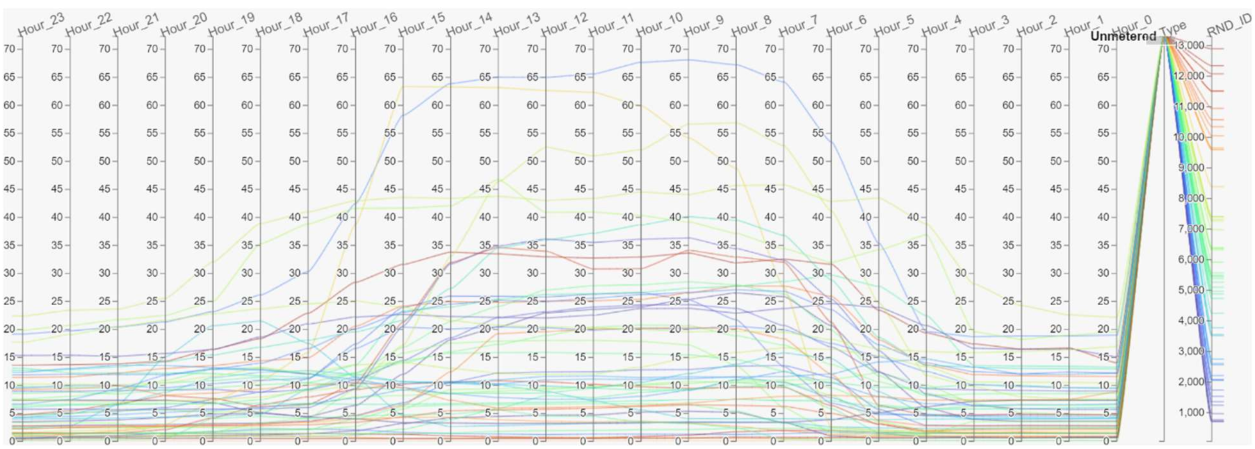


Figure 15 Unmetered

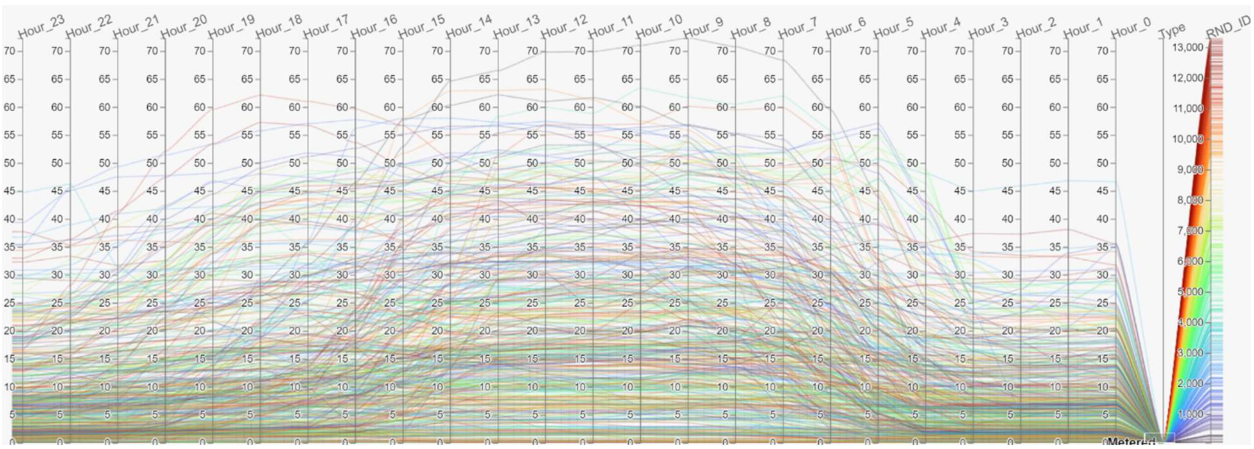


Figure 16 Metered

9.2 Reconstruction of Test Data - Global Year Max Normalization Results

The first set of results demonstrates the comparison of generated profiles for larger and smaller customers. The new balanced dataset shows a significant improvement in the reconstruction quality compared to the previous, smaller dataset. This improvement proves the recommendation of dataset size and balance in training our architecture will improve results significantly.

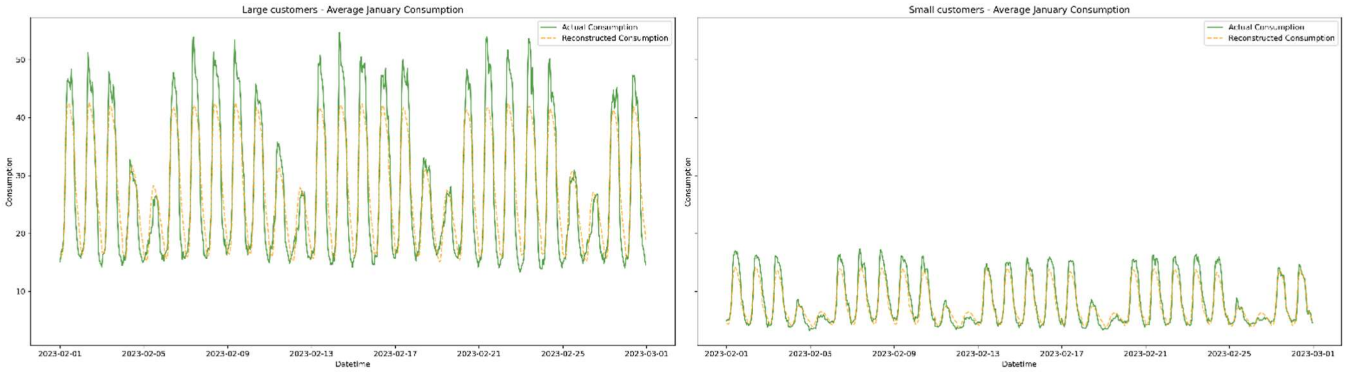
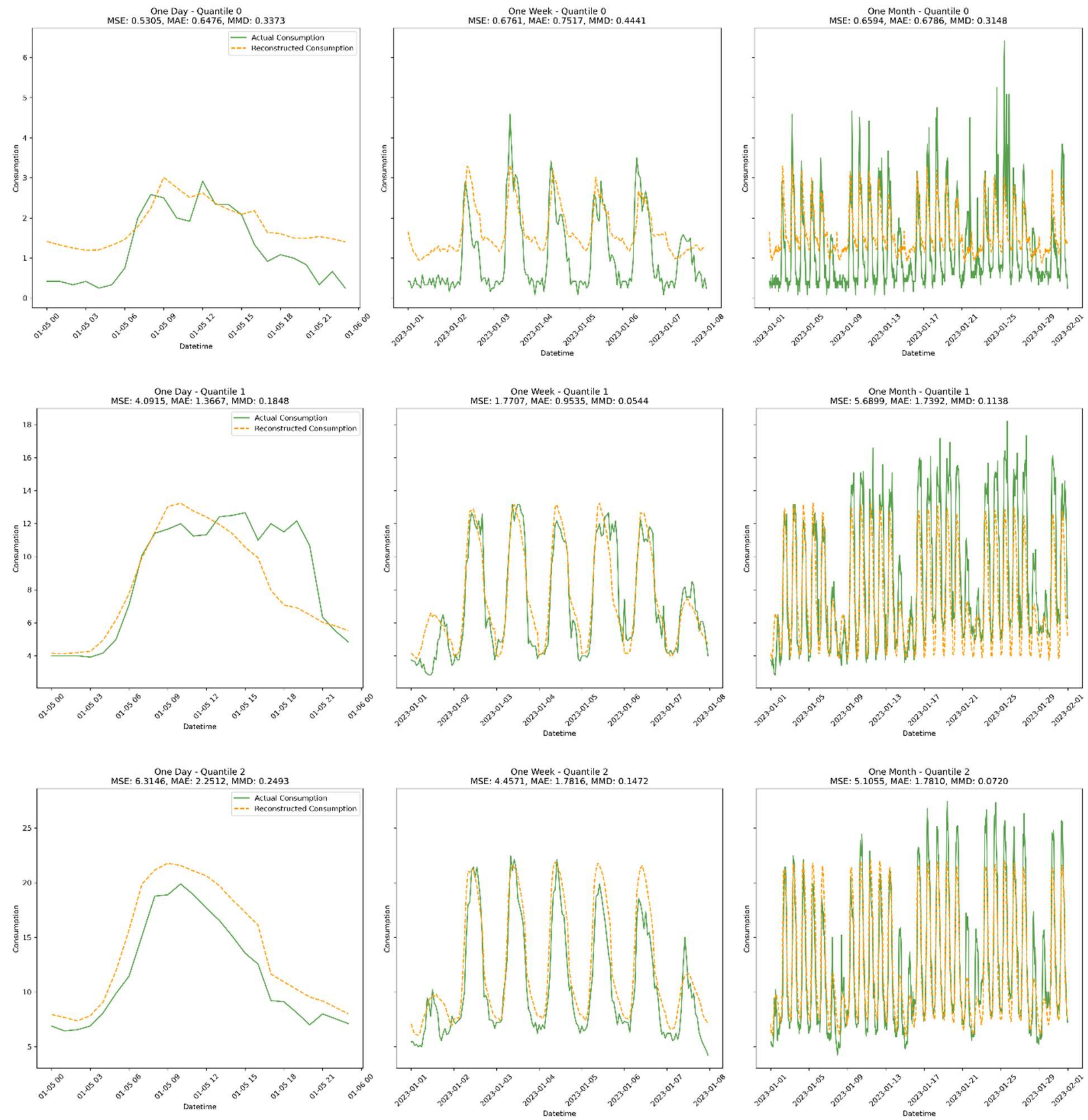
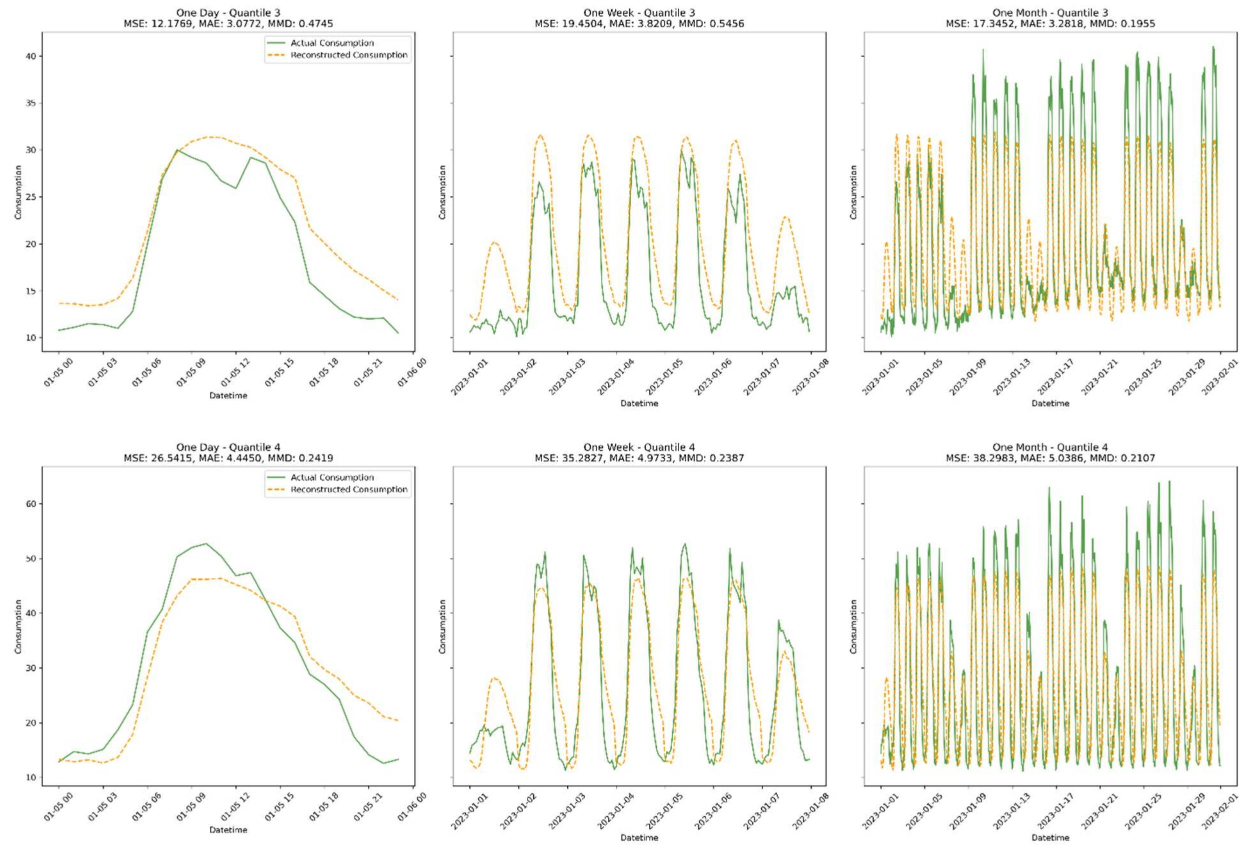


Figure 17 Global Year Max Normalization - Large vs Small Customers - Comparison of generated profiles for large and small customers.

Further, we analyzed the performance across different quantiles to assess the model's ability to capture both the trend and magnitude of consumption. The results, shown in *Figures 18*, indicate that while the model shows promising improvements, particularly in capturing the magnitude, it still struggles somewhat with accurately representing the exact trends. These figures represent the reconstruction results for four quantiles after training for 13 epochs. While the current results are promising, further training with larger datasets and more epochs could yield even better performance, highlighting the potential for continual improvement in future work.



One Day - Quantile 0
MSE: 0.5305, MAE: 0.6476, MMD: 0.3373
One Week - Quantile 0
MSE: 0.6761, MAE: 0.7517, MMD: 0.4441
One Month - Quantile 0
MSE: 0.6594, MAE: 0.6786, MMD: 0.3148
One Day - Quantile 1
MSE: 4.0915, MAE: 1.3667, MMD: 0.1848
One Week - Quantile 1
MSE: 1.7707, MAE: 0.9535, MMD: 0.0544
One Month - Quantile 1
MSE: 5.6899, MAE: 1.7392, MMD: 0.1138
One Day - Quantile 2
MSE: 6.3146, MAE: 2.2512, MMD: 0.2493
One Week - Quantile 2
MSE: 4.4571, MAE: 1.7816, MMD: 0.1472
One Month - Quantile 2
MSE: 5.1055, MAE: 1.7810, MMD: 0.0720
Actual Consumption
Reconstructed Consumption
Consumption
Datetime



Figures 18 Global Year Max Normalization – Three Levels - Comparison of test data and generated profiles using global year max normalization for different quantiles.