# Multivariate and Multistep Forecasting of System Marginal Price Using a Modified WaveNet

Jaeyun Jung, Minkyu Lee, Daegun Ko, Jeonghoon Choi, Hyeseung Han, Bumsu Park, Nayeon Park, Kyoungjoo Kim, Hyunsup Kim and Sungkyu Kim



May 21, 2024

# Multivariate and Multistep Forecasting of System Marginal Price Using a Modified WaveNet Architecture

Jaeyun Jung, Minkyu Lee, Daegun Ko, Jeonghoon Choi, Hyeseung Han, Bumsu Park, Nayeon Park, KyoungJoo Kim, Hyunsup Kim, Sungkyu Kim
*Electrified Propulsion Advanced Development Team, Hyundai Motor Company*
*jayjung@hyundai.com*

## Abstract

*This paper explores the application of a WaveNet deep learning model structure for forecasting of System Marginal Price (SMP) in the context of the electrical grid, focusing on the multivariate and multi-step prediction challenges inherent in real-world application. SMP, the real-time cost of balancing for supply and demand for electricity, it's inherently dynamic and volatility nature in response to numerous factors make the SMP forecasting challenging. By applying a WaveNet model renowned for its proficiency in capturing temporal relationships in time series data, with modifications, we aim to navigate the complexities of SMP forecasting, providing insights and methods that could potentially benefit such dynamic times series applications. The effectiveness of the model is evaluated by comparing with three other state-of-the-art deep learning models: Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and long short-term memory (LSTM) models. The results show that the modified WaveNet model is well-suited for the multivariate and multi-step regression-type problems.*

**Keywords:** Multistep forecasting, Multivariate, time series, System Marginal Price (SMP) , WaveNet

## 1. Introduction

The System Marginal Price (SMP) represents the wholesale electricity price in an electrical grid, reflecting the real-time balance between supply and demand. The development of the SMP was driven by the need to address the challenges of grid stability and peak demands in electricity markets. For various stakeholders, particularly for energy aggregators, accurate forecasting of SMP forecasting is crucial since it not only facilitates the optimization of energy trading strategies but also enhances profitability by enabling participation in the market at advantageous times such as peak pricing periods. The forecasting of SMP, however, presents a challenging task due to its inherent volatility to numerous variables, including fuel prices that are used to generate electricity, weather conditions, unexpected consumer demands and renewable energy outputs. Furthermore, real-world operational constraints that 24-hour period SMP values announcement once per day necessitating a capability of predicting multiple steps ahead.

Past studies have been explored and methods such as statistical models, machine learning and deep learning are used to deal with the time-series prediction issue. Statistical models often represent strong generality and interpretability and among the statistical models, implementation of auto regressive integrated moving average (ARIMA) models in the electricity markets of Spain and California can be found in [1]. However, such traditional autoregressive models tend to fail to predict non-linear patterns. Similarly, application of Support Vector Machines (SVMs) based price forecasting is presented in [2]. Machine learning methods provide reasonable results but for such SMP case where high volatility presents, it is difficult to handle non-linearity between different time steps and other influencing factors. Among deep learning methods, three widely used approach for price forecasting are Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) and in [3-5] proves that these are well-suited for regression-type problems. RNN combined with the gated recurrent units and long short-term memory (LSTM) units is elaborated as more sophisticated model, specifically for time-series problems [6]. Alike LSTM, WaveNet is a modified type of CNN, but it is more specifically designed for audio waveforms generation [7]. WaveNet leverages dilated causal convolutions, enabling it to capture long-range dependencies within the audio sequence. Furthermore, by stacking these dilated convolutions, the network can learn relationships between data points across different times [7]. Motivated by CNN and WaveNet structures, [8] introduces CNN in the form of WaveNet architecture which enables multivariate time series forecasting. The modified WaveNet from this paper is inspired from the CNN in the form of WaveNet, while our model concentrates on enhancement of multistep forecasting particularly for dynamic and volatile data.

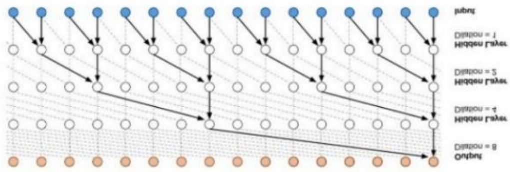

**Figure 1. Visualization of a stack of dilated causal convolutional layers [7]**

The study leverages three years of hourly SMP data (2021-2023) from Jeju Island, South Korea. This region is characterized by its dynamic SMP fluctuations, like shown in figure 2. High reliance on renewable energy resources, accounting for 19% of its energy in 2023, primarily from wind turbines and solar installations makes the region a rigorous testing ground. Through this research, we aim to demonstrate the potential of modified WaveNet model's ability to capture the complex dynamics of SMP which could provide insights that could benefit energy aggregators in maximizing profits through informed decision-making.

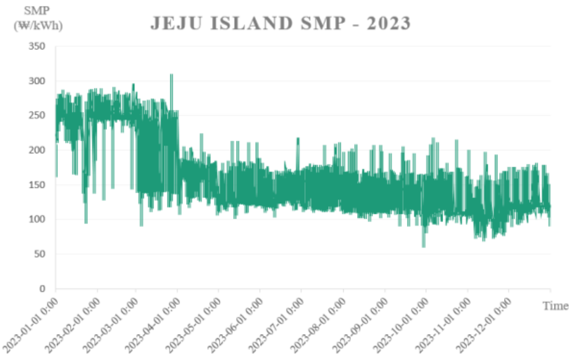


**Figure 2. Jeju Island SMP in 2023**

## 2. Methodology

### A. Datasets and Data Processing

The SMP and Weather data analyzed in this study were acquired from the following organizations: Korea Power Exchange (KPX) and Korea Meteorological Administration (KMA). Recordings spanning from January 1, 2021, to December 31, 2023, at a frequency of every hour, resulting in a total of 26,280 data samples. From KPX, the System Marginal Price (SMP) specific to the Jeju Island region was retrieved as the primary target variable for the forecasting model. Complementarily, the KMA provided a suite of meteorological variables including temperature, humidity, wind speed, wind direction, and both rain and snow precipitation levels. These exogenous variables were deemed essential for understanding the multifaceted influences on SMP fluctuations. These raw time series data underwent several preprocessing steps to prepare it for analytical modeling. First, to enhance the model's ability to learn from historical trends and mitigate short-term volatility, rolling statistics were applied to the Jeju SMP values. Specifically, a 24-hour step was chosen for rolling calculations, aligning with the daily cycle of power consumption and generation. Collected feature set was refined through a feature selection process aimed at reducing model complexity and avoiding overfitting. Variables such as wind speed, wind direction, and precipitation metrics were excluded due to their low correlation with the SMP, as quantified by a Pearson Correlation Coefficient (PCC) with threshold of less than 0.05. The dataset was partitioned into training and validation sets following an 80:20 ratio. This distribution was selected to ensure sufficient data for model training while reserving an adequate subset for validation purposes, thereby facilitating the assessment of the model's generalization capabilities across unseen data. Consequently, a MinMaxScaler which is a prerequisite for efficient training of neural network models was applied to facilitate faster convergence and reduce the likelihood of getting trapped in local optima.

### B. Model Architecture

Once datasets are prepared, WaveNet-based time series model for multivariate and multi-step forecasting is constructed. The model accepts two types of input: exogenous variables such as temperature and the target variable, SMP. The inputs are processed through several layers of causal and dilated convolutions, designed to capture temporal dependencies at varying scales. The core architecture utilizes causal convolutions for initial feature extraction, followed by dilated convolutions to capture long-range dependencies within the sequence. This approach allows for an increase in the receptive field without introducing additional hidden layers, thereby promoting model efficiency. A residual connection after each dilated convolutions is added to implement shortcut connections, which allows information to skip one or more layers and connect directly to the output layer. This method allows the model to dynamically adjust its inputs based on generated values from different layers thus enhancing its ability to model non-linear sequential dependencies. Skip connections and residual blocks are employed to facilitate deeper model architectures without the vanishing gradient problem. Unlike standard Convolutional Neural Networks (CNNs), which capture spatial dependencies, the proposed model is tailored for temporal data, to understand time-series dynamics. Subsequently, the final layer concatenates all intermediate outputs and passes them through a linear activation to produce the forecasted values. In contrast to the gated activation functions typically employed in conventional WaveNet models, we investigate the use of the Scaled Exponential Linear Unit (SELU) activation function. The functions introduce non-linearity while also controlling for the vanishing and exploding gradient problems. While it draws inspiration from the original WaveNet model, significant modifications, such as the incorporation of exogenous inputs and change in residual connections distinguish it. At-a-glance figure that shows the model structure is shown in Figure 3.

These adjustments are designed to enhance the model's applicability to the specific challenges of SMP forecasting. After model architecting, model parameter tuning process was conducted to determine the optimal configuration for the number of filters, kernel size, dilation depth, and regularization strengths. This optimization aimed to achieve a balance between robust learning capacity and efficient computational resource utilization, considering real-world implementation constraints.

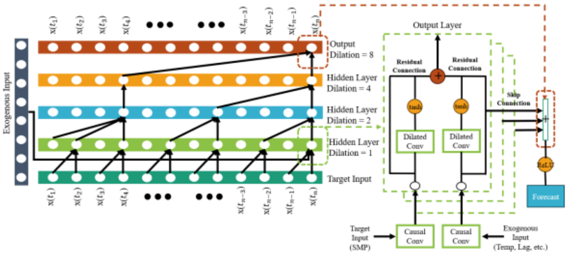


**Figure 3. WaveNet-based time series model**

### C. Model Training and Adjustments

Following the architectural design of the WaveNet-based model, this segment delineates the specifics of the model training strategic selection of hyperparameters aimed at accurate forecasting and optimizing performance. The model was instantiated with the following parameters: sequence length of 72, output length of 24, and 16 exogenous variables. Various sizes of batch size underwent the training process and batch size of 128 was determined to facilitate efficient learning without compromising the model's ability to generalize from the training data. Another critical component of the training procedure was the selection of the learning rate and optimization algorithm. While an initial experiment utilized an Inverse Time Decay learning rate schedule starting at 0.005, subsequent adjustments with aids from learning rate scheduler led to the adoption of a static learning rate of 0.0001. This decision was influenced by the need to ensure stable and convergent training dynamics, particularly given the model's complexity and the depth of temporal dependencies being modeled. The Adam optimizer was chosen for its effectiveness in adapting learning rates. The model underwent training over 100 epochs, with a dynamic adjustment to the batch size with early stopping method to accommodate computational efficiency.

## 3. Experimental Analysis

### A. Indicators for Performance Evaluation

We utilized a set of general metrics which are mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE) specifically chosen to provide a multifaceted view of model performance, allowing for the assessment of both average errors and the distribution of these errors across the dataset.

**Table 1. The overall setting of hyperparameter and architectures for DL models**

| Item | *DNN* | *CNN* | *LSTM* | Mod. *Wave-Net* |
|---|---|---|---|---|
| Input Shape | 72-hour Past Data | | | |
| Output Shape | 24-hour Forecast | | | |
| Hidden Layers | 4 (Dense) | 2 (Conv1D), 2 (Dense) | 2 (LSTM) ,2(Dense) | 4 (Residual Blocks) |
| Channels /Filters | N/A | 64 (Conv1D), 128 (Conv1D) | N/A | 96 (WaveNet Blocks) |
| Dropout Rate | 0.5 | | | N/A |
| Activation Function | ReLU | | | SeLU |
| Kernel Filter | N/A | Size: 2, Stride: 1 | N/A | Size: 2 |
| Optimizer | Adam | | | |
| Regularization | L2: 0.001 | | | |
| Learning Rate | Learning Rate Scheduler (initial: 0.005, steps: 50, rate: 0.01) | | | |
| Batch Size | 128 | | | |

### B. Experimental Settings

To evaluate the effectiveness and precision of the modified WaveNet time series model, comprehensive evaluation framework was established which includes three other benchmarking models: Deep Neural Networks (DNN), Convolutional Neural Networks (CNN) and long short-term memory (LSTM) models. This limited input length of 72 was chosen to achieve a balance between capturing relevant historical information and appropriate computational demanding. Further effort was invested into maintaining consistency across the base models and the modified WaveNet wherever possible. This involved using the same hyperparameters (e.g., learning rate, optimizer, batch size) while only varying the core model architectures to isolate the impact of the structural differences. Table 1 presents a comparative overview of the architectural and hyperparameter settings for the benchmarking models developed in this study. For the proposed WaveNet model, the following parameters were set: network with 3 layers of dilated convolutions, filter size of 2 and 64 number of filters.

### C. Experimental Results and Discussion

A comparative analysis of multistep forecasting performance across four different neural network architectures: Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM), and the proposed WaveNet model are shown in Table 2, in which the best values are highlighted in bold. Figure 4 illustrates the models' forecasting results juxtaposed with actual SMP values for both training and validation datasets.

DNN model presents poor performance in terms of all indicating metrics, which proves a lack of capturing the underlying temporal dynamics. While the CNN and LSTM models capture the overall pattern of the SMP values, the scale of its predictions is compressed compared to the actual data. LSTM's smoother prediction curve suggests a model strength in capturing the underlying trend but possibly at the expense of reacting to short-term fluctuations, which could be critical in a real-world SMP forecasting scenario. The proposed WaveNet model, followed closely by the CNN and LSTM, shows a promising balance between capturing the SMP trends and responding to its volatility, which is crucial for effective real-time forecasting in volatile markets.

This analysis underscores the strengths and limitations of each model type in SMP forecasting. While each model exhibits a degree of proficiency in learning from historical data, the variance in their performance on unseen data highlights the importance of considering both trend capture and volatility responsiveness in model selection and tuning.

**Table 2. Performance comparison among the different forecasting models**

| Metrics | *DNN* | *CNN* | *LSTM* | **Modified *Wave-Net*** |
|---|---|---|---|---|
| MSE | 926.4 | 497.6 | 504.1 | **214.1** |
| MAE | 26.9 | 18.9 | 18.2 | **11.0** |
| RMSE | 30.4 | 22.3 | 22.4 | **14.6** |
| *MAPE* | 18.9 | 14.6 | 14.6 | **8.7** |

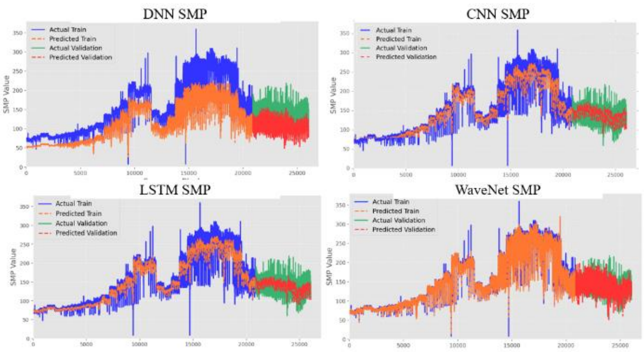


**Figure 4. Actual and predicted values from the different forecasting models**

## 4. Conclusion and Future Work

This paper introduced a novel framework for multistep, multivariate System Marginal Price (SMP) forecasting, leveraging a modified WaveNet architecture. Our approach distinctively combines the WaveNet model's ability to capture complex temporal dynamics with enhancements tailored for the intricacies of SMP forecasting. Experimental results have demonstrated notable improvements across multiple performance metrics, affirming the efficacy of the proposed method in addressing the challenges inherent in SMP forecasting. Future investigations will aim to incorporate Long Short-Term Memory (LSTM) layers by replacing the final layer of the WaveNet model. This hybrid model architecture aims to leverage the LSTM's prowess in capturing long-term dependencies alongside WaveNet model’s effective handling of temporal sequences. Furthermore, the adoption of Mutual Information Coefficient (MIC) and wavelet transform techniques in data preprocessing will be explored. These methods hold the promise of preparing more refined and feature-rich datasets, potentially improving the model's ability to discern and learn from the underlying patterns in SMP data. In conclusion, while this study marks a step forward in SMP forecasting, we aim to further refine our forecasting framework, enhancing its accuracy, robustness, and applicability to real-world energy market operations.

## References

[1]. Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2003). "ARIMA models to predict next-day electricity prices." IEEE Transactions on Power Systems, 18(3), 1014-

[2] Sansom, D. & Downs, T. & Saha, Tapan. (2003). Support Vector Machine Based Electricity Price Forecasting For Electricity Markets utilising Projected Assessment of System Adequacy Data. 2.

[3] A.J Wang, B Ramsay,A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays, Neurocomputing, Volume 23, Issues 1–3, 1998, Pages 47-57, ISSN 0925-2312,

[4] Borovykh, A., Bohte, S., & Oosterlee, C. W. (2017). Conditional Time Series Forecasting with Convolutional Neural Networks. ArXiv. /abs/1703.04691

[5] Zhu, Yongqiong. (2020). Stock price prediction using the RNN model. Journal of Physics: Conference Series. 1650. 032103. 10.1088/1742-6596/1650/3/032103.

[6] K. J, H. E, M. S. Jacob and D. R, "Stock Price Prediction Based on LSTM Deep Learning Model," 2021 International Conference on System, Computation, Automation and Networking (ICSCAN), Puducherry, India, 2021, pp. 1-4, doi: 10.1109/ICSCAN53069.2021.9526491.

[7] Oord, Aaron & Dieleman, Sander & Zen, Heiga & Simonyan, Karen & Vinyals, Oriol & Graves, Alex & Kalchbrenner, Nal & Senior, Andrew & Kavukcuoglu, Koray. (2016). WaveNet: A Generative Model for Raw Audio.

[8] Borovykh, A., Bohte, S., & Oosterlee, C. W. (2017). Conditional Time Series Forecasting with Convolutional Neural Networks.