AN-BEATSfor Short-Term Electricity Load Forecasting with Adjusted Seasonality Blocks and Optimization of Block Order

Anh Tuan Nguyen, Thanh Chau Do Thi, Anh Ngoc Le and Ngoc Anh Nguyen Thi

June 24, 2022

AN-BEATS for short-term electricity load forecasting with adjusted seasonality blocks and optimization of block order

Nguyen Anh Tuan[1], Do Thi Thanh Chau[1], Le Anh Ngoc[2] and Nguyen Thi Ngoc Anh[1*]

[1]Hanoi University of Science and Technology, Hanoi, Vietnam.
[2]Swinburne University of Technology, Hanoi, Vietnam.

*Corresponding author(s). E-mail(s): anh.nguyenthingoc@hust.edu.vn;
Contributing authors: anhtuannguyen280798@gmail.com; thanhchau1704@gmail.com; ngocla2@fe.edu.vn;

Abstract

For the proper operation of electrical systems, accurate electricity load forecasting is essential. This study focuses on solving the problem that is the optimization of the block order to archive better accuracy of the forecasting model. Furthermore, the seasonality blocks of N-BEATS are adjusted in theory by correctly using the Discrete Fourier Transform. Therefore, AN-BEATS-Adjusted Neural Basic Expand - Analysis Time series model is proposed to forecast short-term power loads based on electricity load history. Experiments show that the proposed model works better than the LSTM model and the order of blocks strongly affects the model's prediction results.

Keywords: N-BEATS, Time-series, Forecasting, Electricity load, Seasonal decomposition

1 Introduction

Electricity load forecasting at the national level is becoming more crucial in successful power management and is the foundation for corresponding decision making [1]. Deep learning are now chosen by academics for short-term electricity load forecasting due to their various benefits in empirical data [2].

Yaoyao He et al. [3] presented a model to handle an issue of short-term power load probability density forecasting. The kernel-based support vector quantile regression (KSVQR) and Copula theory are used in this model. The short-term electrical load forecasting is solved in the paper [4] by integrating the extreme learning machine (ELM) with a novel switching delayed PSO algorithm. The proposed model outperforms existing cutting-edge ELMs in terms of performance. Song Li et al. present a novel synthetic technique for short-term electrical load forecasting based on wavelet transform, extreme learning machine (ELM), and partial least squares regression in their publication [5] (PLSR). The wavelet transform was used to decompose the fundamental time series, and ELM was employed for each sub-component resulting from the wavelet decomposition. The findings showed that the suggested approach might improve load forecasting performance. In this paper [6], the authors study the influence of deep-stacked LSTM and Bi-LSTM models on electricity load forecasting. In addition, they tune the hyperparameters of the LSTM model by using the Bayesian optimization method to gain the best performance forecasting. The experiment results show that deep-stacked LSTM layers do not increase prediction accuracy significantly.

In this paper [7], the authors proposed to use a Fourier series form with a detrended linear function what to represent a periodic time series function with a linear trend. The complete Fourier representation is used in the comparative analysis. Diagnostic tests demonstrated that the suggested method outperforms the pure Fourier approach. [8] A novel method forecasts the trend-cycle and seasonal components using fuzzy techniques. The proposed method gets good performance without considering irregular fluctuation for time series with clear trends and seasonal components. [9] Both of the mean price reversion (MRP) and the seasonality component are represented by the Fourier series. In addition, the Fourier series represent the MRP have a stochastic factor. The Kalman Filter is used to estimate the long-term and seasonality components in the representation state space. In another research [10], the authors focus on the ability of Singular Spectral Analysis, Seasonal-Trend decomposition using Loess, and attribute selection pre-processing methodologies combined with neural network models to forecast monthly river streamflows in Turkey's Nallihan stream.

N-BEATS [11] is a deep learning model built only by fully connected layers to solve the univariate time series forecasting problem and performs well with time-series data in many fields of competition. The most important advantage of the model is interpretation. N-BEATS can decompose time series data into three components trend, seasonality, and residual. Another advantage of the model is fast training and inference than another deep learning model.

In [1], the authors used the ensemble technique to build N-BEATS models bootstrapped to solve the mid-term electricity load forecasting problem. Compared proposal models with ten other methods (ARIMA, k-NN, MLP,...), the ensemble of N-BEATS has the best performance. In [12], the authors apply a Neural Architecture Search to add Recurrent Neural Network to N-BEATS model called N-BEATS-RNN that help to reduce training time. However, N-BEATS-RNN does not refer to interpretability when compared to N-BEATS. In [13], the authors extend the N-BEATS model in combination with exogenous variables called the N-BEATSx model for the multivariate time series forecasting problem.

This paper aims to improve the accuracy of electricity load forecasting by the AN-BEATS model. We made some changes to the orthonormal basis of the Discrete Fourier Transform in the seasonality blocks of the N-BEATS model. In addition, we also considered the influence of block order on the accuracy of electricity load forecasting. Concretely, we create combinations of five blocks that include three trend blocks and two seasonality blocks. Then, for each order obtained, we train the model and evaluate the results of the model error to find the best model.

The left structure of the paper is as follows: Section 2 explains the methodology of the N-BEATS model, AN-BEATS's Seasonality blocks, and the Proposed model. The experiments and results have been presented in Section 3. The final of this paper is Section 4 that provides the conclusion and discussions.

2 Methodology

2.1 Problem Statement

We consider the problem of univariate time series forecasting in discrete time. Given the observed series history $[y_1, y_2, ..., y_T] \in \mathbb{R}^T$, we need to predict the vector of future values, denoted as $\widehat{\mathbf{y}} = [\widehat{y}_{T+1}, \widehat{y}_{T+2}, ..., \widehat{y}_{T+N}] \in \mathbb{R}^N$. The model can be shown as follow:

$$\widehat{y}_{T+1}, \widehat{y}_{T+2}, ..., \widehat{y}_{T+N} = F(y_1, y_2, ..., y_T) \tag{1}$$

where F is a model we aim to learn.

2.2 N-BEATS

The N-BEATS's structure proposed by the authors in this paper [11]. N-BEATS model consists of stacked blocks which to form stacks. The basic building block includes two-part. The first part is a multi-layer fully connected model with a Relu activation function. Then, we obtain two outputs forward θ^f and backward θ^b. The second part consists of the backward g_l^b and the forward g_l^f basis layers that accept the respective θ_l^f and θ_l^b expansion coefficient. The idea of double residual stacking is used to arrange the blocks into stacks.

The structure of the model is shown in the following figure 1:

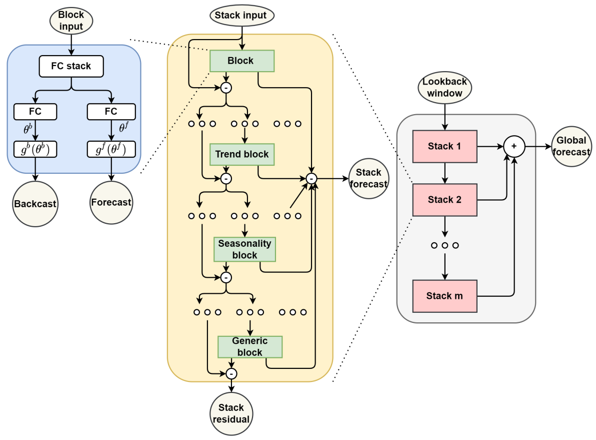

Fig. 1: N-BEATS architecture

2.3 AN-BEATS's Seasonality blocks

Adjusted Neural Basic Expand - Analysis Time series model called AN-BEATS is proposed by adjusting seasonality blocks in this section. To model seasonality, $g^b_{s,l}$ and $g^f_{s,l}$ is periodic functions. Given a length-N forecast horizon, according to the [11], Fourier representation is expressed in the following formula:

$$\widehat{\mathbf{y}}_{s,l} = \sum_{i=0}^{\lfloor N/2-1 \rfloor} \theta^f_{s,l,i} \cos(2\pi f_i t) + \theta^f_{s,l,i+\lfloor N/2 \rfloor} \sin(2\pi f_i t) \tag{2}$$

Then, they came up with the following seasonal forecast form:

$$\widehat{\mathbf{y}}^{seasonal}_{s,l} = \mathbf{S}\theta^f_{s,l} \tag{3}$$

Where $\theta^f_{s,l} = [\theta^f_{s,l,0}, \theta^f_{s,l,1}, \ldots, \theta^f_{s,l,N-1}]^T$ are predicted by a FC network of layer l of stack s. $\mathbf{S} = [1, \cos\frac{2\pi}{N}t, ..., \cos\lfloor\frac{N}{2}-1\rfloor\frac{2\pi}{N}t, \sin\frac{2\pi}{N}t, ..., \sin\lfloor\frac{N}{2}-1\rfloor\frac{2\pi}{N}t]$ is the matrix of sinusoidal waveforms, and $f_i = \frac{i}{N}$ is the i^{th} harmonic of the fundamental frequency $\frac{1}{n}$.

We can easily see that if N is an odd number, the equation 2 has only $N-1$ terms and we are missing $\cos(\frac{N-1}{2}.\frac{2\pi}{N}t) = \cos(\frac{N-1}{N}\pi t)$ component. In addition, although the equation 2 has enough N terms, we are still missing $\cos(\frac{N}{2}.\frac{2\pi}{N}t) = \cos\pi t$ component. That leads to the accuracy of the forecasting model can decrease. So, we have made some changes to get the original normal basis of the Discrete Fourier Transform as follows:

If N is old, we represent $\widehat{y}_{s,l}$ in the following formula:

$$\widehat{y}_{s,l} = \theta_{s,l,0} + \sum_{i=1}^{(N-1)/2} \theta^f_{s,l,i} \cos(2\pi f_i t) + \theta^f_{s,l,i+(N+1)/2} \sin(2\pi f_i t) \qquad (4)$$

If N is even, we represent $\widehat{y}_{s,l}$ in the following formula:

$$\widehat{y}_{s,l} = \sum_{i=0}^{N/2} \theta^f_{s,l,i} \cos(2\pi f_i t) + \sum_{i=1}^{N/2-1} \theta^f_{s,l,i+N/2} \sin(2\pi f_i t) \qquad (5)$$

Thus, the matrix **S** in the equation 3 is also changed.

2.4 Proposed method for load forecasting

To solve the electricity load problem with the AN-BEATS model, we have proposed a model as follows (see Figure 2):

- Step 1: The time-series data is pre-processed and divided into training, validation, and test sets.
- Step 2: Train the AN-BEATS model on the training set.
- Step 3: Evaluate the model on the validation set to choose the most optimal parameters set.

  We have all $M = \frac{(a+b)!}{a!b!}$ combinations of the order of blocks (with a, b are the number of trend blocks and the number of seasonality blocks, respectively). We train the models with different block orders in turn from $1^{st}, 2^{nd}, ..., i^{th}$ combination, until we run out of M combinations. Thus, after evaluating step 3, we continue to turn the order of blocks and perform step 2 again.
- Step 4: After training M models, we choose the model with the best performance. The final model with optimized blocks order is used to forecast the testing set.

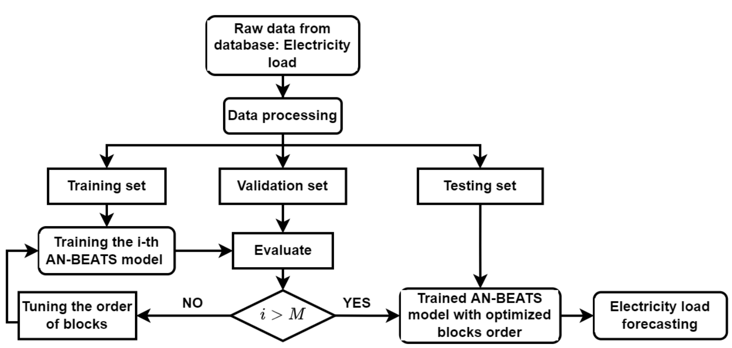

Fig. 2: Proposed model

3 Experiments and results

3.1 Description of data set

The data used for the experiment is the Vietnamese National electricity load data set. The time series start from 01/01/2015 to 05/09/2020. It includes 58560 data points (one-hour resolution). Table 1 describes how to divide the data set for training and testing. In that experiment, we predicted the load value for two days ($N = 48$) from the previous ten days ($T = 240$).

Table 1: Experiment data set

Data Set	From	To	Number of data points
Train	01/01/2015	24/11/2019	52560
Validation	25/11/2019	07/04/2020	2400
Test	08/04/2020	05/09/2020	3600

3.2 Evaluation criteria

Root mean square error (RMSE) and mean absolute percentage error (MAPE) are used to evaluate the performance of the models. These formulas of criteria are as follows:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y(t) - \widehat{y}(t)|}{y(t)} \qquad RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y(t) - \widehat{y}(t))^2} \tag{6}$$

3.3 Experimental Scenarios

Three experimental scenarios will be used for comparison:

1. Scenario 1: LSTM model with different hyper-parameters is used for forecasting.
2. Scenario 2: N-BEATS consists of three trend blocks ($a = 3$), two seasonality blocks ($b = 2$), and one generic block with the different trend and seasonality blocks order.
3. Scenario 3: The proposed model is the AN-BEATS model with the number of blocks and changes order similar to N-BEATS in scenario 2 with the new seasonality blocks.

The models in scenarios trained using Adam optimizer, RMSE for the loss function, were trained on 100 epochs and batch size 512. We also applied early stopping to stop training if the validation loss doesn't decrease for 15 consecutive epochs. The best parameters set for each model are selected through RMSE on the validation set with the most optimal value. Finally, the model is evaluated on the test set. We train and inference with NVIDIA GTX 1080Ti 12GB.

3.4 Results

In scenario 1, we fine-tune hyper-parameters: number of layers (Num layers), number of neurons (NN), and learning rate (LR). In scenario 2 and scenario 3, we change the order of trend blocks (T), seasonality blocks (S) and keep the last block as a generic block (G). There are $\frac{(a+b)!}{a!b!} = \frac{5!}{3!2!} = 10$ models in each scenario 2 and scenario 3. Details of the block orders are depicted in Table 3. In scenarios 2 and 3, we train models with a learning rate is 0.001. For each model in all scenarios, we train and evaluate it 30 times. We then average the results that include the mean and standard deviation of MAPE, RMSE and show them in Tables 2 and 3.

From Tables 2 and 3, we can see that N-BEATS and AN-BEATS model gave superior results when compared with the LSTM model in scenario 1. The LSTM models in scenario 1 also have a longer training time. In addition, the MAPE is above the threshold of 7%, and the RMSE is above the threshold of 6000 (MW). Besides, we see that the order of blocks also affects the performance of the model N-BEATS and AN-BEATS for the load forecasting problem. With the original order blocks (T, T, T, S, S, G), RMSE and MAPE of the N-BEATS model are 3274 (MW), and 3.79% and RMSE and MAPE of AN-BEATS are 3306 (MW) and 3.71 respectively. Meanwhile, both N-BEATS and AN-BEATS got the best results with the order blocks (T, S, T, S, T, G). With the order of (T, S, T, S, T, G), the N-BEATS model has RMSE of 3258 and MAPE of 3.63. Besides, RMSE and MAPE of the AN-BEATS model are 3179 and 3.54, respectively. We see that with some orders, the performance of N-BEATS and AN-BEATS decreases. From Table 3, we figure out that for some orders of blocks, AN-BEATS achieves better results than N-BEATS while for others is worse results. The main reason that interprets AN-BEATS does not improve much compared to N-BEATS is that the load series is pretty clean. Furthermore, the N-BEATS model with the old seasonality block can still capture the seasonality of the data even though it is the missing basis in the space like formula 2.

Table 2: Results of scenario 1: LSTM with different hyper-parameters

Num layers	**NN**	**LR**	**MAPE**	**RMSE (MW)**	**Training time (s)**
1	128	0.001	7.71 ± 0.13	6474 ± 76	2500
1	256	0.001	7.82 ± 0.11	6490 ± 57	2610
1	128	0.002	7.85 ± 0.15	6531 ± 80	2451
1	256	0.002	7.94 ± 0.13	6598 ± 110	2605
2	128	0.001	7.98 ± 0.29	6605 ± 197	2596
2	256	0.001	8.15 ± 0.14	6722 ± 42	2919
2	**128**	**0.002**	$\mathbf{7.62 \pm 0.03}$	$\mathbf{6409 \pm 120}$	2561
2	256	0.002	7.69 ± 0.13	6453 ± 78	2925

Table 3: Results of scenario 2 and scenario 3

Order of blocks						**Scenario 2: N-BEATS**			**Scenario 3: AN-BEATS**		
1	**2**	**3**	**4**	**5**	**6**	**MAPE**	**RMSE**	**Training time**	**MAPE**	**RMSE**	**Training time**
T	T	T	S	S	G	3.71 ± 0.12	3306 ± 43	1130	3.79 ± 0.01	3274 ± 23	1029
T	T	S	S	T	G	3.78 ± 0.04	3345 ± 21	1085	3.98 ± 0.06	3401 ± 62	1077
T	T	S	T	S	G	3.85 ± 0.02	3354 ± 34	1072	3.91 ± 0.02	3323 ± 40	1068
T	S	S	T	T	G	3.89 ± 0.01	3402 ± 17	1059	4.07 ± 0.04	3434 ± 19	1021
T	**S**	**T**	**S**	**T**	**G**	$\mathbf{3.63 \pm 0.03}$	$\mathbf{3258 \pm 30}$	1038	$\mathbf{3.54 \pm 0.02}$	$\mathbf{3179 \pm 23}$	1027
T	S	T	T	S	G	3.77 ± 0.06	3343 ± 49	1048	3.66 ± 0.03	3254 ± 10	1009
S	S	T	T	T	G	3.80 ± 0.05	3377 ± 46	1065	3.75 ± 0.03	3277 ± 32	1014
S	T	S	T	T	G	3.71 ± 0.07	3314 ± 43	1071	3.86 ± 0.02	3381 ± 11	1040
S	T	T	S	T	G	3.65 ± 0.05	3272 ± 25	1059	3.60 ± 0.01	3238 ± 18	1028
S	T	T	T	S	G	3.69 ± 0.04	3289 ± 36	1068	3.59 ± 0.03	3268 ± 47	1051

Figure 3 shows an example of hourly load prediction of three models LSTM, N-BEATS, and AN-BEATS (with best parameters from Table 2 and Table 3). The models will observe the load values from 240 hours to predict the next 48-hours load values. In general, the AN-BEATS model will give the best results in most of the examples in the test set. Figure 4 shows an example of the interpretable AN-BEATS model (with best parameters in Table 3) based on trend and seasonality components. The left chart in the figure shows the original load series and the prediction results of the AN-BEATS model. The center chart shows the trend component, and the right chart shows the seasonality component. The results display polynomial trend components in the data and evident seasonal effects within the days of the load series.

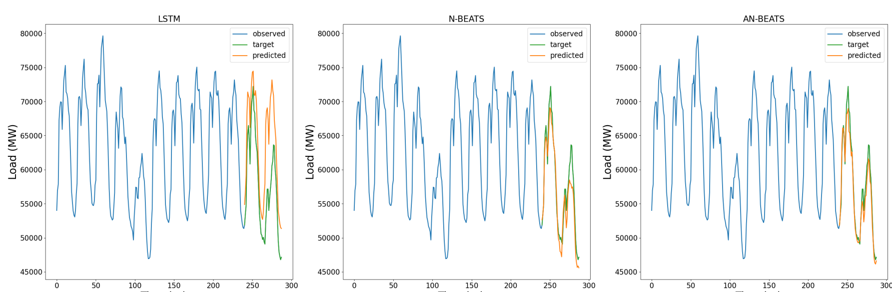

Fig. 3: Results of LSTM (left), N-BEATS (center) and AN-BEATS (right)

4 Conclusion and discussion

If load forecasting is precise and reliable, power system operations can be safer and more cost-effective. This work introduces AN-BEATS, a new model for

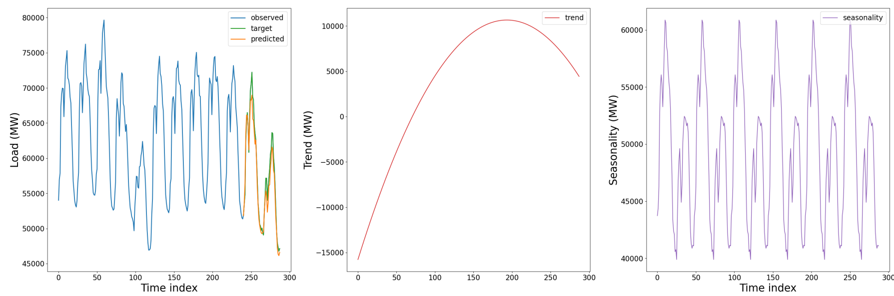

Fig. 4: Example of the interpretable AN-BEATS model

short-term load forecasting that focuses on adjusting seasonality blocks and optimization block order. The following are some of the paper's findings:

- Adjusting seasonality blocks to archive AN-BEATS.
- Successfully proposing a new model AN-BEATS with optimization block order.
- Using AN-BEATS for short-term forecasting.
- Applying the proposed approach for forecasting short-term electricity load in the National Load Dispatch Centre of Vietnam Electricity.

In the future, we intend to solve the challenge to optimize the number of blocks in the model, the number of trend blocks, and the order blocks of the AN-BEATS models. Furthermore, because the new AN-BEATS model can only predict with a single time series, expanding AN-BEATS to multiple time series will be investigated.

Acknowledgments

This work was supported by the Hanoi University of Science and Technology and the National Load Dispatch Centre of Vietnam Electricity.

References

[1] Oreshkin, B.N., Dudek, G., Pełka, P., Turkina, E.: N-beats neural network for mid-term electricity load forecasting. Applied Energy **293**, 116918 (2021)

[2] Serrano-Guerrero, X., Briceño-León, M., Clairand, J.-M., Escrivá-Escrivá, G.: A new interval prediction methodology for short-term electric load forecasting based on pattern recognition. Applied Energy **297**, 117173 (2021)

[3] He, Y., Liu, R., Li, H., Wang, S., Lu, X.: Short-term power load probability density forecasting method using kernel-based support vector quantile regression and copula theory. Applied Energy **185**, 254–266 (2017)

[4] Zeng, N., Zhang, H., Liu, W., Liang, J., Alsaadi, F.E.: A switching delayed pso optimized extreme learning machine for short-term load forecasting. Neurocomputing (2017)

[5] Song, L., Lalit, G., Peng, W.: An ensemble approach for short-term load forecasting by extreme learning machine. Applied Energy **170**, 22–29 (2016)

[6] Atef, S., Eltawil, A.B.: Assessment of stacked unidirectional and bidirectional long short-term memory networks for electricity load forecasting. Electric Power Systems Research **187**, 106489 (2020)

[7] Iwok, I.: Seasonal modelling of fourier series with linear trend. International Journal of Statistics and Probability **5**, 65 (2016)

[8] Nguyen, L., Novák, V.: Forecasting seasonal time series based on fuzzy techniques. Fuzzy Sets and Systems **361**, 114–129 (2019). Theme: Clustering and Rule-based Models

[9] Moreno, M., Novales, A., Platania, F.: Long-term swings and seasonality in energy markets. European Journal of Operational Research **279**(3), 1011–1023 (2019)

[10] Apaydin, H., Sattari, M.T., Falsafian, K., Prasad, R.: Artificial intelligence modelling integrated with singular spectral analysis and seasonal-trend decomposition using loess approaches for streamflow predictions. Journal of Hydrology **600**, 126506 (2021)

[11] Oreshkin, B.N., Carpov, D., Chapados, N., Bengio, Y.: N-beats: Neural basis expansion analysis for interpretable time series forecasting. In: International Conference on Learning Representations (2020)

[12] Sbrana, A., Rossi, A.L.D., Naldi, M.C.: N-beats-rnn: deep learning for time series forecasting. In: 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA) (2020)

[13] Olivares, K.G., Challu, C., Marcjasz, G., Weron, R., Dubrawski, A.: Neural basis expansion analysis with exogenous variables: Forecasting electricity prices with nbeatsx. International Journal of Forecasting, submitted **Working Paper version available at arXiv:2104.05522** (2021)