Short Term Load Forecasting Using Wavelet Transform

POUYA IRANI
1937388

Introduction:
Wavelet transform is a mathematical tool that is widely used for time series analysis and forecasting. It has gained popularity due to its ability to capture both time and frequency domain features in a signal. In this report, I will discuss the application of wavelet transform in time series forecasting.

Wavelet Transform:
Wavelet transform is a mathematical tool that decomposes a signal into a set of wavelets with different frequencies and scales. The wavelets are defined as small oscillations with zero mean and a finite energy. The transformation is carried out by convolving the signal with a set of wavelets that are scaled and shifted versions of a mother wavelet. The wavelet coefficients obtained through this process provide information about the time and frequency characteristics of the signal.

Wavelet Transform in Time Series Forecasting:
Time series forecasting involves predicting future values of a signal based on its past behavior. Traditional methods of time series forecasting such as ARIMA (Autoregressive Integrated Moving Average) assume that the signal is stationary and can be modeled using a linear combination of its past values. However, many real-world signals are non-stationary and exhibit non-linear behavior. Wavelet transform can be used to capture the non-stationary and non-linear characteristics of a signal.

Feature Extraction with wavelets:
Wavelet transform can be used for feature extraction from a time series signal. The wavelet coefficients obtained through the transformation provide information about the time and frequency characteristics of the signal. These coefficients can be used as features for a machine learning algorithm to predict future values of the signal.

Applications of Wavelet Transform in Time Series Forecasting:
Wavelet transform has been applied in various fields for time series forecasting. Some of the applications are:

Financial Time Series Forecasting:
Wavelet transform has been used for predicting stock prices, exchange rates, and other financial time series data. It has been shown to improve the accuracy of predictions compared to traditional methods.
Power System Forecasting:
Wavelet transform has been used for predicting power system load demand and wind power generation. It has been shown to capture the non-linear and non-stationary characteristics of the signals, which improves the accuracy of the predictions.
Biomedical Signal Forecasting:
Wavelet transform has been used for predicting epileptic seizures, electrocardiogram (ECG) signals, and other biomedical signals. It has been shown to improve the accuracy of the predictions compared to traditional methods.
Wavelet transform is a powerful mathematical tool that can be used for time series forecasting. It provides a way to capture both time and frequency characteristics of a signal and can be used for denoising and feature extraction. Wavelet transform has been applied in various fields and has shown to improve the accuracy of predictions compared to traditional methods.

Main types of wavelet decomposition methods:
Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT) are two different types of wavelet transforms used in time series analysis and forecasting. While both of these methods

are used to decompose a signal into a set of wavelets, there are some fundamental differences between them.

Continuous vs. Discrete:
The key difference between CWT and DWT is that CWT is a continuous transformation, whereas DWT is a discrete transformation. In CWT, the wavelet function is continuously scaled and shifted to cover the entire frequency spectrum of the signal. In contrast, in DWT, the wavelet function is only scaled and shifted at discrete intervals.

Time-Frequency Resolution:
Another important difference between CWT and DWT is their time-frequency resolution. In CWT, the time and frequency resolution are inversely proportional to each other. This means that if the time resolution is high, the frequency resolution will be low and vice versa. In contrast, DWT provides a better time-frequency resolution because it uses a set of different wavelet filters with different frequency bands.

Boundary Effects:
CWT is known to produce boundary effects when the signal being analyzed is not periodic. This occurs because the wavelet function extends beyond the boundaries of the signal, leading to artificial fluctuations at the boundaries. DWT, on the other hand, does not suffer from boundary effects because it uses a finite-length filter bank.

Computational Complexity:
CWT is computationally more complex than DWT. This is because CWT involves an infinite number of wavelet coefficients, which makes it computationally expensive. DWT, on the other hand, uses a finite-length filter bank, which makes it computationally less expensive.
Application of CWT and DWT in Time Series Forecasting:
CWT and DWT have their own advantages and disadvantages in time series forecasting. CWT is preferred when high-frequency components of a signal are important, such as in speech recognition and music analysis. DWT is preferred when time and frequency resolution are both important, such as in financial time series forecasting and load forecasting.

The DWT uses a discrete dyadic (octave) grid for scale parameter i and shift parameter k, and the equation for forward DWT is as follows:

$$c(i,k) = \sum_{t} f(t)\Psi^{*}_{i,k}(t), \text{ where } \Psi^{*}_{i,k}(t) = 2^{\frac{i}{2}}\Psi\left(2^{i}t - k\right).$$

The mathematical equation for inverse DWT is as follows:

$$f(t) = \sum_{k}\sum_{i} c(i,k)\Psi_{i,k}(t)$$

where $f(\cdot)$ is a function and Ψ the mother wavelet.

Maximal Overlapping Discrete Wavelet Transformations (MODWT):
The traditional Discrete Wavelet Transform (DWT) has some limitations in terms of its ability to capture the time-frequency characteristics of a signal. In particular, it is difficult to capture the highest frequency components of a signal, which can be important for my applications. To address this limitation, the Maximal Overlapping Discrete Wavelet Transform (MODWT) was developed.

The MODWT is similar to the traditional DWT in that it decomposes a signal into a set of wavelet coefficients. However, it differs in the way that the signal is decomposed. In the MODWT, the signal is decomposed using a set of overlapping wavelet filters. This allows the highest frequency components of the signal to be captured more effectively.

Application of MODWT in Time Series Forecasting:

Data Preprocessing:
Wavelet transform can be used for data preprocessing by denoising the load data. The load data often contains noise and other irregularities that can affect the accuracy of the forecasting model. Wavelet transform can be used to remove noise from the load data by filtering out the high-frequency components that are not relevant to the forecasting task.
Feature Extraction:
Wavelet transform can also be used for feature extraction by decomposing the load data into a set of wavelet coefficients. These coefficients can be used as features in a forecasting model. Wavelet transform can also be used to extract the trend and seasonal components of the load data, which can be used as additional features in the model.
Time-Frequency Analysis:
Wavelet transform can be used for time-frequency analysis of the load data. This can be useful in identifying the dominant frequency components of the load data, which can be used to improve the forecasting accuracy. Time-frequency analysis can also help in identifying the periods of high and low load demand.
Forecasting Model:
Wavelet transform can be used as a preprocessing step in forecasting models such as Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs). The wavelet coefficients can be used as inputs to the forecasting model, along with other relevant features such as weather data, holiday data, and historical load data.
In conclusion, wavelet transform is a powerful tool for short-term load forecasting in the energy industry. It can be used for data preprocessing, feature extraction, time-frequency analysis, and as a preprocessing step in forecasting models. The use of wavelet transform can lead to more accurate load forecasts, which can help in ensuring the stability and reliability of the power grid.

My experiment:
My proposed method involves using the MODWT algorithm to preprocess data and obtain more detailed information about the observed time series. By applying the MODWT algorithm to the time series $Y_t(t = 1,2,\dots,N)$, we can obtain J detail coefficients (higher frequency bands) and a smooth (trend component). I illustrate this using a Haar filter in a visual representation. The wavelet levels are chosen based on an integer value that indicates the number of decompositions into smooth details, which we set as $J + 1 = \lfloor \log_e N \rfloor$.
The resulting decomposed series consist of high and low-frequency wavelet (details) and scaling (smooth) coefficients that can accurately track the original series.

$$Y_t = \sum_{j=1}^{J} D_{j,t} + S_{j,t}$$

where $D_{j,t}$ is the j^{th} level details and $S_{j,t}$ is the smooth of the decomposed time series.

Each level of the time series and the smooth component are then passed into several independent Transformers for the prediction task.

The whole procedure is illustrated in the algorithm below.

Algorithm 1 W-Transformers

Input: The original Time Series (training set): Y_t

$(\mathrm{t} = 1,2, \ldots, \mathrm{N})$

Output: The prediction of the testing series: $\hat{Y}_{N+h}$

NF orecast ← Number of data point of testing set (h);

$J + 1$-decomposition ← MODWT (Y_t);

for i in $J + 1$ do

train ← $(J + 1)$-Levels $[i][1\colon N]$;

test ← $(J + 1)$-Levels $[i][N + 1\colon N + h]$;

for each epoch do

Embed In ← Embedding1 1 train);

Embed Out ← Embedding2 (train);

Encode Out ← Encoder (embedIn);

decode $1Out$ ← Decoder1 (embedOut);

decode $2Out$ ← Decoder2 (encodOut, decod $1Out$);

Model ← SoftMax (Linear (decod $2Out$));

end for

transformer Results $[i]$ ← prediction (Model, test);

end for

$\hat{Y}_{N+h}$ ← Inverse MODWT (Transformer Result);

Also the figure below gives us a better understanding of the model which I have used.

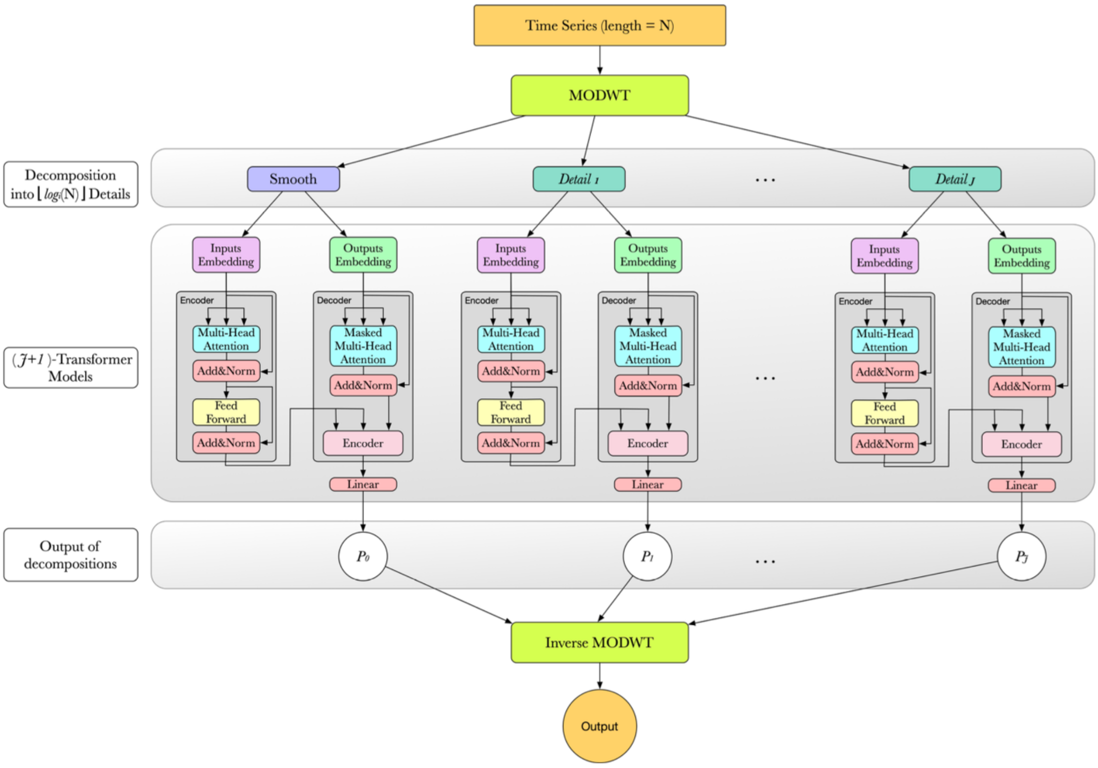

figure 1: the Diagram representing the proposed W-Transformers architecture

Results:

I have developed a MODWT wavelet transformation on “PJM_Load_hourly” dataset which is the energy consumption of the east coast of the United States between 1998 to 2001 in MW.
For the sake of simplicity and since I have limited computational power, I used only first 5000 samples.

After applying the wavelet transform the original signal would decompose to 9 signals with different frequencies.

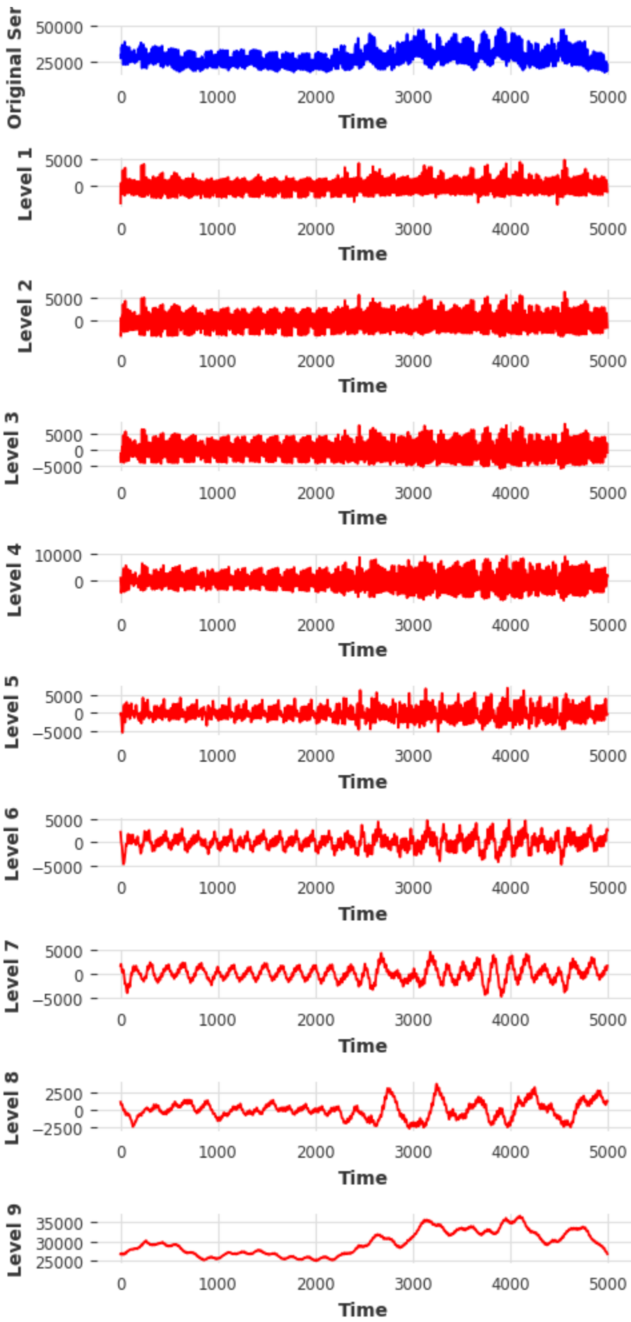

figure 2 wavelet decomposition

since now the original signal decomposed to 9 different signal we need to train them separately and the take inverse of MODWT transform to get the original signal.
I have used the standard transformer model from darts library as I would like to see only the effect of MODWT on the accuracy of the model and make a comparison with previously done projects.
I need to mention that once more that I only used 5000 samples while the original dataset has more than 87000 sample.

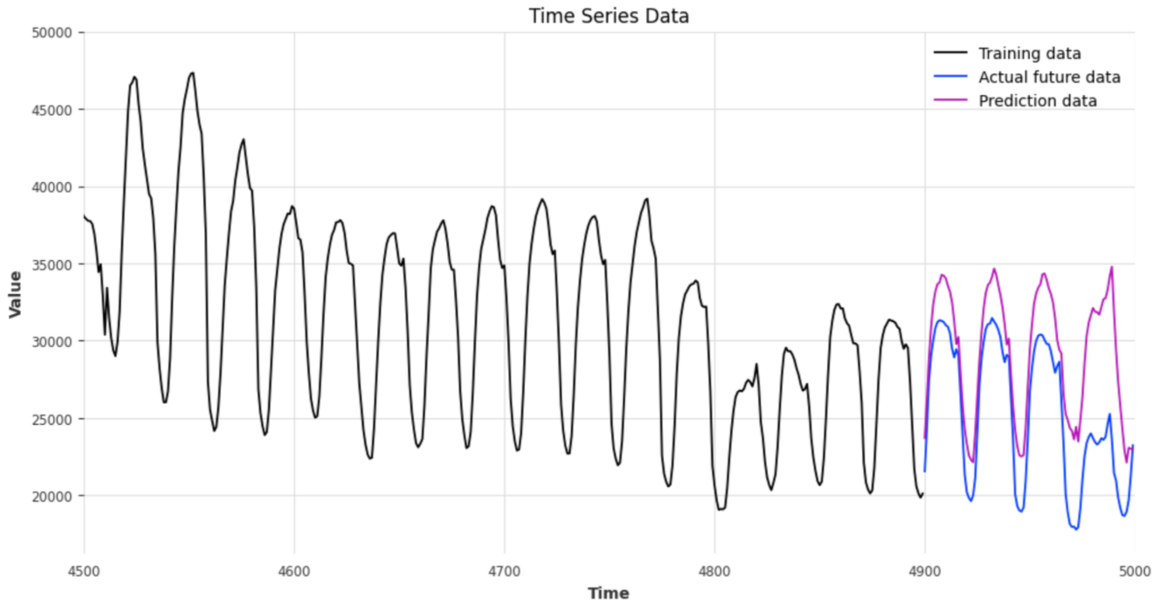

figure 3 test results

in this experiment I set the last 100 samples as the testing samples to illustrate the accuracy of the model and the result is encouraging.

Future steps:
In the next steps I am going to first design a more specific and more accurate model and make a comparison between traditional ARIMA statistical models and also compare with deep learning models such as RNN and LSTM.