Group 27 ECE251C Final Report

Hatim Alhazmi Nathan Newbury Zeting Luan
University of California, San Diego
haalhazmi@ucsd.edu nnewbury@ucsd.edu zluan@ucsd.edu

Abstract

Feature extraction from the Discrete Wavelet Transform (DWT) and variants as inputs into machine learning regression and classification models has shown promise for time series data. This technique has mainly seen used in classifying EKG readings, where a combination of Wavelets with certain features can match the EKG. There is some research on this for weather prediction, but often with limitations. This performance review of various wavelet techniques to predict weather information indicates promise in this area.

1. Introduction

We propose using coefficients from the DWT and the Stationary Wavelet Transform (SWT) to predict wind speed with a regression model. Wavelets have variable time and frequency resolution, which could make them a better option than the Short Time-Frequency Transform (STFT) to highlight specific features that predict wind speed. Our data set[4] is comprised of temperature, humidity, cloud cover, wind direction, visibility, and finally wind speed. Data for these measurements were taken every hour for 11 years. Three models utilizing the neural net shown in Figure 1 to predict the wind speed of the next hour were compared. In the first model DWT coefficients for each subband were inputted as features into a single network. The DWT is not translation-invariant. To address this, in our second network we inputted SWT coefficients from each subband as features into a single network. These methods are inspired from [1]. In the final model, we trained a neural network for each SWT subband coefficients to predict the next coefficient. We reconstructed these coefficients back into a time series with an additional hour of data. Time series data inputted directly into Figure 1 predicted the next hour of wind speed with an MSE of 28.1 All models improved on this with the correct wavelet filter choices.

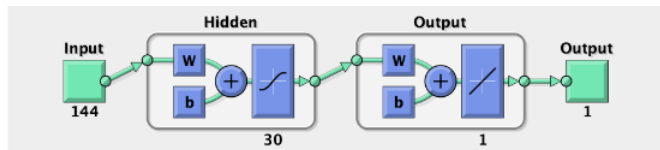

Figure 1. Basic Neural Net Building Block

2. Review of Previous Approaches

2.1. Classification using DWT and random forest

In [2], Emanet proposed a new method for ECG beat classification. He performed the ECG signals feature extraction using DWT and the Random Forest to classify the ECG signals. A total of five types of different ECG beats were classified with a success of 99.8 percent. According to Emanet, for non-stationary signals, a wavelet-based time-frequency representation is the better feature extraction technique. The basic idea of the wavelet transform is representing a function as a superposition of wavelets. There were 265 wavelet coefficients obtained for each ECG segment, which were fed as feature vectors to the Random Forest algorithm.

2.2. Wind speed prediction using Random Forest model

Random forest is an ensemble machine-learning method for many tasks like classification, regression, etc. Specifically for regression tasks, the model returns the mean prediction of the individual trees. It utilized the bagging mechanism, which creates a different training subset from sample training data, and the output is based on majority voting. There are many works regarding weather prediction using random forest algorithms. In [3], Meenal *et al* selected multiple data including maximum temperature, minimum temperature, surface pressure, etc to predict wind speed. By using the mean square error (MSE) for model evaluation, the result beats other machine learning methods like Support Vector Machine (SVM). Similar work includes [1]. Drisya, G. V.*et al* trained a model that was used to predict the remaining data points at an interval of two weeks.

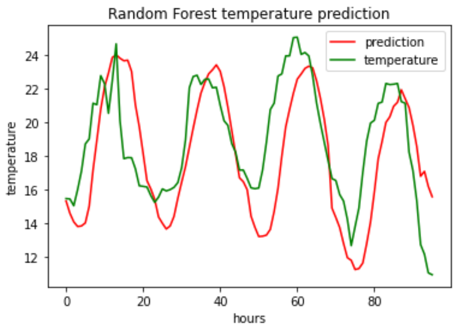

Figure 2. Random forest baseline temperature prediction

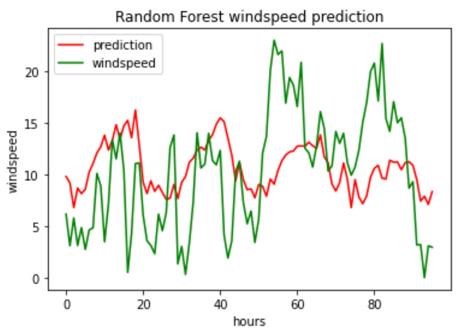

Figure 3. Random forest baseline wind speed prediction

Based on our research on the Random Forest algorithm for weather predicting like [1] and [3], we built a Random Forest model as the baseline. The prediction results were shown in Figure 2 and Figure 3. It produced a decent result in temperature prediction and an underwhelming result in wind speed prediction as they achieved mean squared errors (MSEs) of 9.124 and 34.386 respectively.

3. Proposed approach

Our approach is using Discrete/Stationary Wavelet Transform to predict windspeed by first applying DWT/SWT to a windspeed dataset to decompose it into different sub-bands. Then, the decomposition (coefficients) are used as features for an MLP(multi-layer perceptron) machine learning model. This approach has the advantage of exposing specific time-varying and frequency-dependent characteristics of wind data, before inputting features into the model, which hopefully can improve the accuracy of windspeed prediction. We have implemented and tested three different prediction models. Also, we explored the effects of changing the wavelet types and the set of coefficients(sub-bands) used. This section is divided into four parts: Dataset, MLP Model, DWT system and SWT system.

3.1. Dataset

The weather dataset[4] consists of 96453 hours of measurements(Temperature, Humidity, Windspeed...etc) from 2006 to 2017. We have split it into a training set and a test set in a ratio of 80:20. We will use the training set to train our model, and the test set to evaluate its performance. The input to our model is the wind speed of 144 hours, and the output is the prediction for the next hour. We believe that this dataset is well-suited for our machine learning model and will provide us with accurate and reliable predictions. The raw windspeed data is split into 669 vectors of length 144 each and the next data point 145 is the ground truth as shown in Figure 4 below.

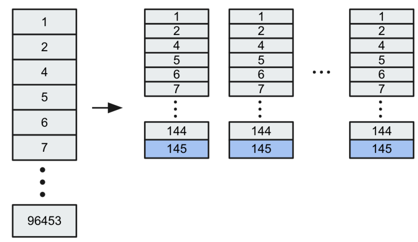

Figure 4. Dataset split.

3.2. Discrete Wavelet Transform Model

In wavelet analysis, discrete wavelet transform (DWT) breaks down a signal into sets of mutually orthogonal wavelet basis functions. The functions are different from sinusoidal basis functions as they are spatially localized. The wavelet functions are dilated, translated, and scaled versions of a common function ϕ, known as the mother wavelet. Due to its invertibility, the original signal can be completely recovered from its DWT representation. Figure 2 shows the block diagram of the DWT model. We implemented a 2-layer Multi-perceptron neural network with 30 neurons in the first layer and an output layer. The cost function used is the Mean Square Error function. The raw windspeed data is split into 669 vectors of length 144 each and the next data point 145 is the ground truth. The machine learning model is trained by the coefficients of the DWT of each vector, and it is compared to the ground truth, which is the windspeed of the next hour. MATLAB Wavelet toolbox and nntraintool toolbox are used.

3.3. Stationary Wavelet Transform Model

Stationary wavelet transform (SWT) is a variant of the wavelet transform. Unlike the discrete wavelet transform (DWT), which downsamples by 2 at each decomposition level, the SWT does not down-sample the signal or subbands at later levels. The SWT however down-samples the mother wavelet, thus altering the low pass and high pass filters, at each level. This allows SWT to better capture the local characteristics of the signal, such as its mean and variance, at different scales. The SWT is additionally transitional invariant. In contrast, DWT is more sensitive to global changes in the signal, and can better capture the overall trend and shape of the signal. Both SWT and DWT have their own strengths and weaknesses, and which one is more suitable for a particular application depends on the specific characteristics of the signal and the goals of the analysis. Figure 5 shows the block diagram of the prediction of windspeed by using SWT coefficients.

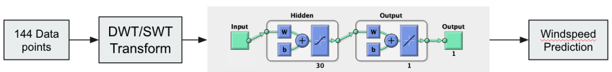

Figure 5. DWT/SWT to Windspeed prediction model.

3.4. Stationary Wavelet Coefficients Prediction Model

In this model, instead of directly predicting wind speed, each subband was trained on its own network to predict the next coefficient. The predictions were added to the end of each subband and then reconstructed back to time series data. This reconstructed time series data now has a prediction tagged on. We use the time-invariant stationary wavelet transform for this network such that the concatenation of a predicted coefficient, when reconstructed, would result in the concatenation of a time series prediction. Attempts to use the discrete wavelet transform failed as the subband array lengths varied in size due to down-sampling. A block diagram of this process is shown in Figure 6.

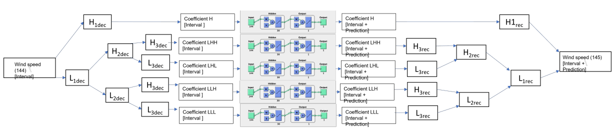

Figure 6. SWT coefficients model for packet 1.

3.5. Wavelet Packets

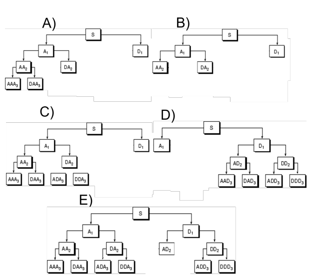

Figure 7. A) The level 3 wavelet decomposition B) The level 2 wavelet decomposition C) Packet 1 D) Packet 2 E) Packet 3

The splitting decisions of an ”optimal” tree were made with the Shannon entropy cost function: $-\sum_{n=i}^{n} s_i^2 log(s_i^2)$ This is related to the logarithm of energy entropy $\sum_{n=i}^{n} log(s_i^2)$. An entropy ”cost” was calculated for each sub-band. A sub-band node was ”split” if the sum of the entropy for its split sub-bands was greater than the node. Packet 1, along with the level 3 decomposition both made proper splitting decisions depending on the interval and mother wavelet. Packet 3 was also the best packet for the bi-orthogonal filter. Packet 2 was sub-optimal by this criteria and was introduced as a control. The table below shows the optimal packet by either criteria.

Wavelet \ Entropy Cost Function	Shannon	Log Energy
Sym4	Level 3	Level 3
Coif4	Level 3	Level 3
Db6	Packet 1	Level 3
Haar	Level 3	Level 3
Bior2.2	Packet 3	Level 3

Table 1.

4. Experiment result

4.1. DWT Coefficients to Windspeed Prediction

Set of coefficients \ Wavelet	sym4	Coif4	Db6	Haar	Bior2.2
Wavelet Packet 1	31.7	27.5	26.8	22.9	25.8
Level 2 Decomposition	25.9	18.9	23.3	26.1	24.5
Level 3 Decomposition	26.2	28.1	27.4	26	33.9

Table 2.

Table 2 above shows the MSE for the testing set when predicting wind speed using DWT coefficients as the input to the model. Although this model did not perform well, the Coif4 wavelet along with the level 2 decomposition provides the best result.The model out performs the random forest and basic MLP model with MSEs of 28.1 and 34.4 respectively. A plot of the predicted wind speed vs. the actual wind speed of the best result is shown in Figure 8.

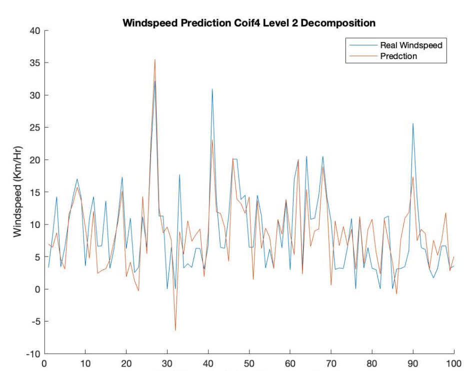

Figure 8. Windspeed prediction using DWT coefficients.

4.2. SWT Coefficients to Windspeed Prediction

Set of coefficients \ Wavelet	Sym4	Coif4	Db6	Haar	Bior2.2
Wavelet Packet 1	30.2	30.4	27	32.4	63.1
Level 2 Decomposition	21.2	20.4	24.4	23.6	23.9
Level 3 Decomposition	27.1	29	30	25.5	23.4

Table 3.

Table 3 above shows the MSE for the testing set of our model using SWT coefficients. In this case, using SWT instead did not improve the results. Similar to the previous model, using level 2 decomposition performs better than level 3 decomposition. In both cases, we notice that using the Coif4 wavelet yields the best results. The plot of using SWT with Coif4 and level 2 decomposition is shown in Figure 9.

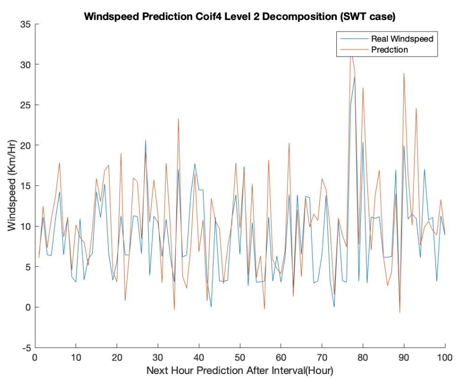

Figure 9. Windspeed prediction using SWT coefficients.

4.3. SWT Coefficients to SWT Coefficients Prediction

Set of coefficients \ Wavelet	Sym4	Coif4	Db6	Haar	Bior2.2
Wavelet Packet 1	.7	7.0	6.8	13	13
Wavelet Packet 2	2	7.8	7.7	3.6	1.5
Wavelet Packet 3	.25	8.3	8.3	.3	1.2
Level 2 Decomposition	1.6	2.7	4.2	4.3	1
Level 3 Decomposition	1.8	2.5	3.4	2.6	0.36

Table 4.

The table above shows the mean squared error for a combination of mother wavelets and wavelet packets. The performance of each packet was dependent the mother wavelet. Our best performance was with packet 3 on the Symlet4 wavelet. This performance seems unreal but was recreated multiple times. Even excluding the first and final 8 coefficients of each subband when training and testing resulted in an MSE of around 1 for this configuration. All wavelet and packet combinations were trained and tested using identical neural nets and training data. Packet 3 also performed well with bio-orthogonal filters and the Haar wavelet. This is notable because Haar and bio-orthogonal filters have linear phases and the Sym4 has an approximate linear phase. The linear phase filters performed better in every category except with wavelet packet 1. Using Shannon entropy as our criteria to split the packet, we would expect these results to mirror Table 2. Packet 1 performed the best out of all three packets for our Debauchies 6 wavelet. Packet 3 performed the best out of all three packets for our bi-orthogonal wavelet. However, Haar and Symlet's amazing performance using wavelet packet 3 was not reflected with this criteria. This suggests the criteria for selecting proper wavelet packets is flawed. The level 3 decomposition performed well across the board and better than the level 2 decomposition. Level 3 decomposition is a good trade-off between the number of networks trained and accurate predictions. Packet 3 involves training seven separate networks, level 3 decomposition involves training train four.

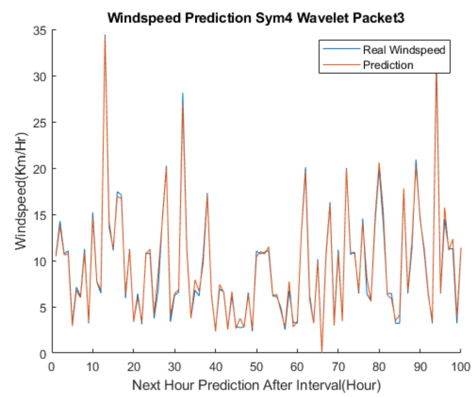

Figure 10. Windspeed prediction by using SWT coefficients to SWT coefficients prediction.

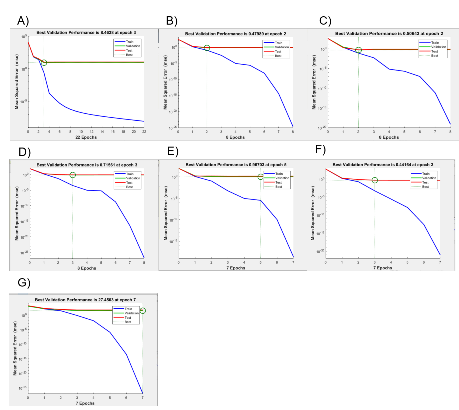

Figure 11. Validation, training and testing loss for packet 3 coefficients. A)AAA3, B)DAA3, C)ADA3, D)DDA3, E) AD2, F)ADD3 G)DDD3

5. Contributions

Nathan - Implemented full SWT coefficients to SWT coefficients model, helped with DWT/SWT to windspeed model, wavelet packet analysis, implemented Random Forest model, trained networks, co-wrote document.

Hatim - Implemented DWT/SWT to wind speed model and basic MLP model, wavelet packet analysis, trained networks, auto correlation analysis, co-wrote document.

Zeting - Implemented Arima model to compare, implemented basic mlp model, reviewed previous work, co-wrote document.

6. Conclusion

This review demonstrates the value in using discrete wavelet transform (DWT) and stationary wavelet transform (SWT) methods to predict wind speed. Each of the three wavelet approaches varied in performance depending on the wavelet type and the set of wavelet coefficients used. Training a model with DWT coefficients as features achieves a similar performance to training the same model with SWT coefficients. This performance is notably better than training the model with time series wind speed data. Training a model with SWT coefficients to predict the next coefficient in the sub-band yielded the best results. Wavelets with linear, or approximately linear phase performed the best. Wavelet packet 3 produced the signal most accurate wind speed prediction, however the basic level 3 decomposition also gave accurate predictions. This is notable because we train nearly half of the neural nets using the level 3 decomposition as using packet 3. This approach can be utilized and improved to predict the wind speed for the next day or week. Overall, our results suggest that SWT can be a valuable tool for wind speed prediction, and further research can be done to identify the best approach for specific applications.

References

[1] G. V. Drisya, Valsaraj P., K. Asokan, and K. Satheesh Kumar. Wind speed forecast using random forest learning method. 2022.

[2] Nahit Emanet. Ecg beat classification by using discrete wavelet transform and random forest algorithm. In *2009 Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control*, pages 1–4, 2009.

[3] Rajasekaran Meenal, Prawin Michael, D. Pamela, and Ekambaram Rajasekaran. Weather prediction using random forest machine learning model. *Indonesian Journal of Electrical Engineering and Computer Science*, 22:1208, 05 2021.

[4] MUTHUKUMAR.J. Weather dataset. `https://www.kaggle.com/datasets/muthuj7/weather-dataset`, 2022. retreived December 5, 2022.