1st International Safe and Green Tomorrow Congress
16-18 April 2025 , Kyrenia / Cyprus

Short-Term Wind Speed Prediction Based on Artificial Intelligence

ABDALLAH DWIKAT
ZAFER ASLAN
ŞÜKRAN SİBEL MENTEŞ
AHMET TOKGÖZLÜ

Istanbul Aydin University,
Istanbul Technical University,
Süleyman Demirel University

UNIVERSITY OF KYRENIA

Presentation Plan

- Summary
- Introduction
- Literature Review
- Methodology
- Findings and Discussion
- Conclusion
- Limitations and Recommendations

SUMMARY

The utilization of renewable energy resources is mainly related to the problem of the **prediction precision of wind** speed. Based on wind speed data, the definition of wind energy potential emphasizes ensuring the efficiency and reliability of wind energy systems. The objective of this work is to present a comprehensive review of **artificial intelligence (AI) techniques applied to forecast short-term wind speed.**

The study compares systematically six different AI models, ranging from machine learning models like high-capacity models that comprise Random Forests, Support Vector Machine, autoregressive moving average (ARMA), Linear and Logistic Regression to more complex models like Long Short-Term Memory (LSTM) networks.

The models are trained and validated against extensive historical records of wind speeds. This paper covers some analysis related to the hourly data collected at the west of the Mediterranean Sea (Mugla City at the Latitude: 556335 and Longitude: 4070184 [ED-50 formats]), between 2001 and 2002.

The study highlights the greater significance of new predictive analytics as a move towards ending the issues of incorporating renewable energy sources into traditional power systems. The study also provides useful insight into AI-based forecasting system design and implementation, paving the way for breakthroughs in environmental applications.

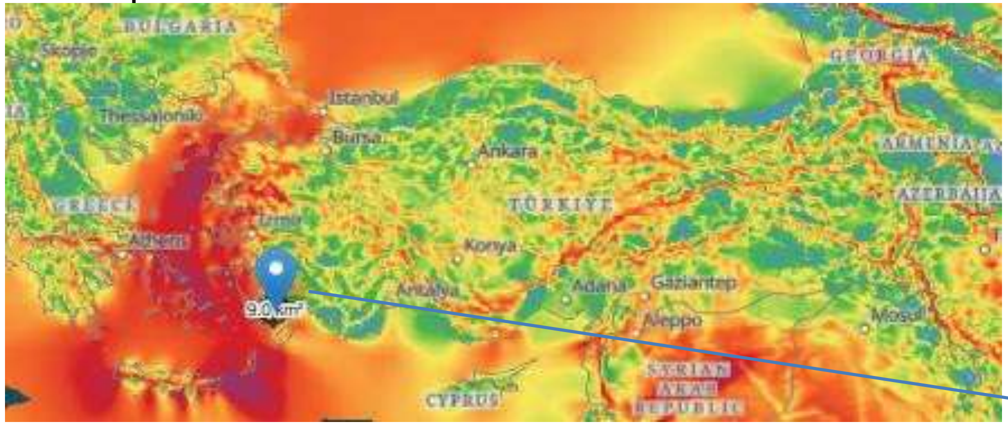
Keywords: Wind energy, sustainability, data driven predictions, machine learning, artificial intelligence.

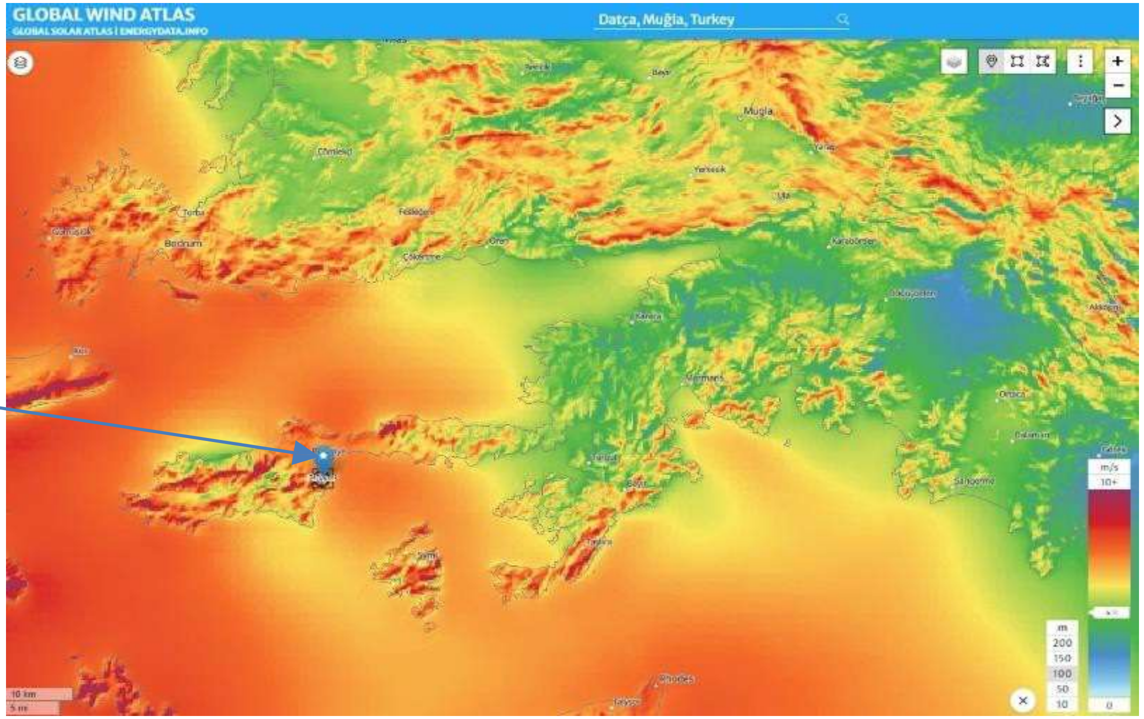
INTRODUCTION

DATA and STUDY AREA

Geographically, Datça is included in the Aegean Region and is located between 27.40-28.00 degrees east meridians and 36.60-36.75 degrees north parallels. Its surface area is 459 km^2. Located in the southwest of Muğla Province, there is the Gulf of Gökova, the wide Hisarönü Gulf, the Marmaris continent, the Aegean and the Mediterranean. Datça is in the Aegean Region as a ministry. It has a mountainous and rugged terrain. The highest points of the Datça Peninsula are Bozdağ (1174m amsl), Kalecik Mountain (881m), Karadağ (786m), Emecik Mountain (704m), Yarım Mountain (615m).

Hourly wind speed data observed 2021 and 2022 are taken into accounted for computer based simulations of wind speed.





METHODOLOGY

Random Forest

The Random Forest algorithm is a method consisting of decision trees. As the name suggests anyway, it creates a forest and does it somehow randomly. The "forest" algorithm builds is a collection of decision trees that are mostly trained by the "bagging" method. In this method, new trees are formed by continually drawing representatives from the sample data set to substitute them, and a community arises from these trees. Based on this, they are referred as ensemble methods. All of the resulting trees are used for modelling, and each one is asked for their opinion.

The steps of this algorithm are as follows:

Step 1: In the first step, the number of decision trees (n) to be created according to the data properties is determined.

Step 2: At each node in the created decision trees, m random variables are selected, and the best branch is determined by calculating with the Gini index.

Step 3: The best branch determined in the previous step is divided into two sub-branches. This process is continued until the Gini index becomes zero, in other words, until there is only one class left in each node.

Step 4: In the last stage, if it is a classification problem, the class with the highest number of votes is chosen as the final decision, if it is a regression problem, the average of individual trees' output is returned as estimation among the predictions made by n decision trees.

Autoregressive Moving Average (ARMA)

ARMA models are applied to non-stationary series but converted to stationary by difference-taking. Models used to non-stationary series but converted to stationary by difference-taking are called nonstationary linear stochastic models. These models are AR, applied to series with a d-degree difference. The general representation of the models is ARMA (p, d, q). Here, p and q are the degrees of the autoregressive (AR) model and the moving average (MA) model, respectively, and d is the degree of difference. ARIMA is denoted as:

$$y'_t = \phi_1 y'_{t-1} + ... + \phi_p y'_{t-p} + \theta_i w_{t-1} + ... + \theta_q w_{t-q}$$

- y'_t is differenced series

Vector Machines (SVM)

Support Vector Machines (SVM) is a supervised learning method. It is used for classification and regression. Points outdoor the tube are support vectors due to the fact they are helping the shape or formation of the tube. These characteristic vectors had been named support vectors. Compared to other traditional learning methods, this method has much better performance and ability to solve nonlinear problems.

In classification problems, the aim are to and the appropriate classifier function that separates the two classes. The goal is to and the best linear classifier plane even for nonexistent data. In the case of SVMs, the maximum distance between two classes is called the margin. There are infinitely many lines to separate these two classes. But there is only one line that maximizes the margin. This line is called the best separating hyperplane. In SVM, the goal is to transform the classification

Vector Machines (SVM)

$$\hat{y} = \begin{cases} 0 \text{ if } \mathbf{w}^T \cdot \mathbf{x} + b < 0 \\ 1 \text{ if } \mathbf{w}^T \cdot \mathbf{x} + b \geq 0 \end{cases}$$

SVM, which has been successfully applied in classification problems, can be applied to

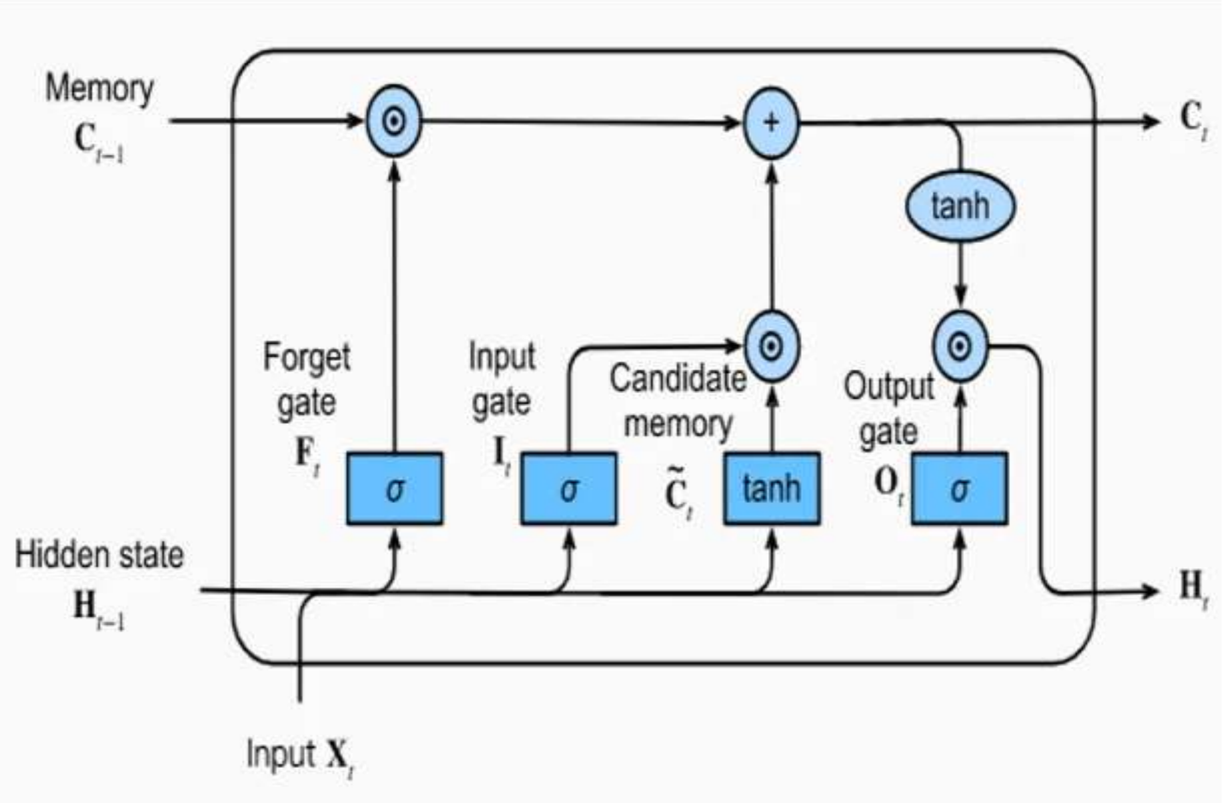
SVMs are preferred for classification processes in many different application areas.

LSTM Architecture

Long Short-Term Memory (LSTM) is a recurrent neural network architecture designed by Sepp Hochreiter and Jürgen Schmidhuber in 1997.

The LSTM architecture consists of one unit, the memory unit (also known as LSTM unit). The LSTM unit is made up of four feedforward neural networks. Each of these neural networks consists of an input layer and an output layer. In each of these neural networks, input neurons are connected to all output neurons. As a result, the LSTM unit has four fully connected layers.

LSTM



The result of the whole process is given as a single output by taking a vote/average opinion by evaluating the opinion of each. With Random Forest, regression and classification analyses can be performed. In order to branch the nodes according to the method, the best random value in the nodes is selected, and certain weights are given to the created decision trees Figure: Architecture of a LSTM Unit (https://d2l.ai/chapter_recurrent-modern/lstm.html)

Wavelet Transforms

- Wavelet
  - A small wave
- Wavelet Transforms
  - Convert a signal into a series of wavelets
  - Provide a way for analyzing waveforms, bounded in both frequency and duration
  - Allow signals to be stored more efficiently than by Fourier transform
  - Be able to better approximate real-world signals
  - Well-suited for approximating data with sharp discontinuities
- "The Forest & the Trees"
  - Notice gross features with a large "window"
  - Notice small features with a small

DEFINITION OF CONTINUOUS WAVELET TRANSFORM

$$CWT_x^{\psi}(\tau, s) = \Psi_x^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \bullet \psi^*\left(\frac{t-\tau}{s}\right) dt$$

Translation

(The location of the window)

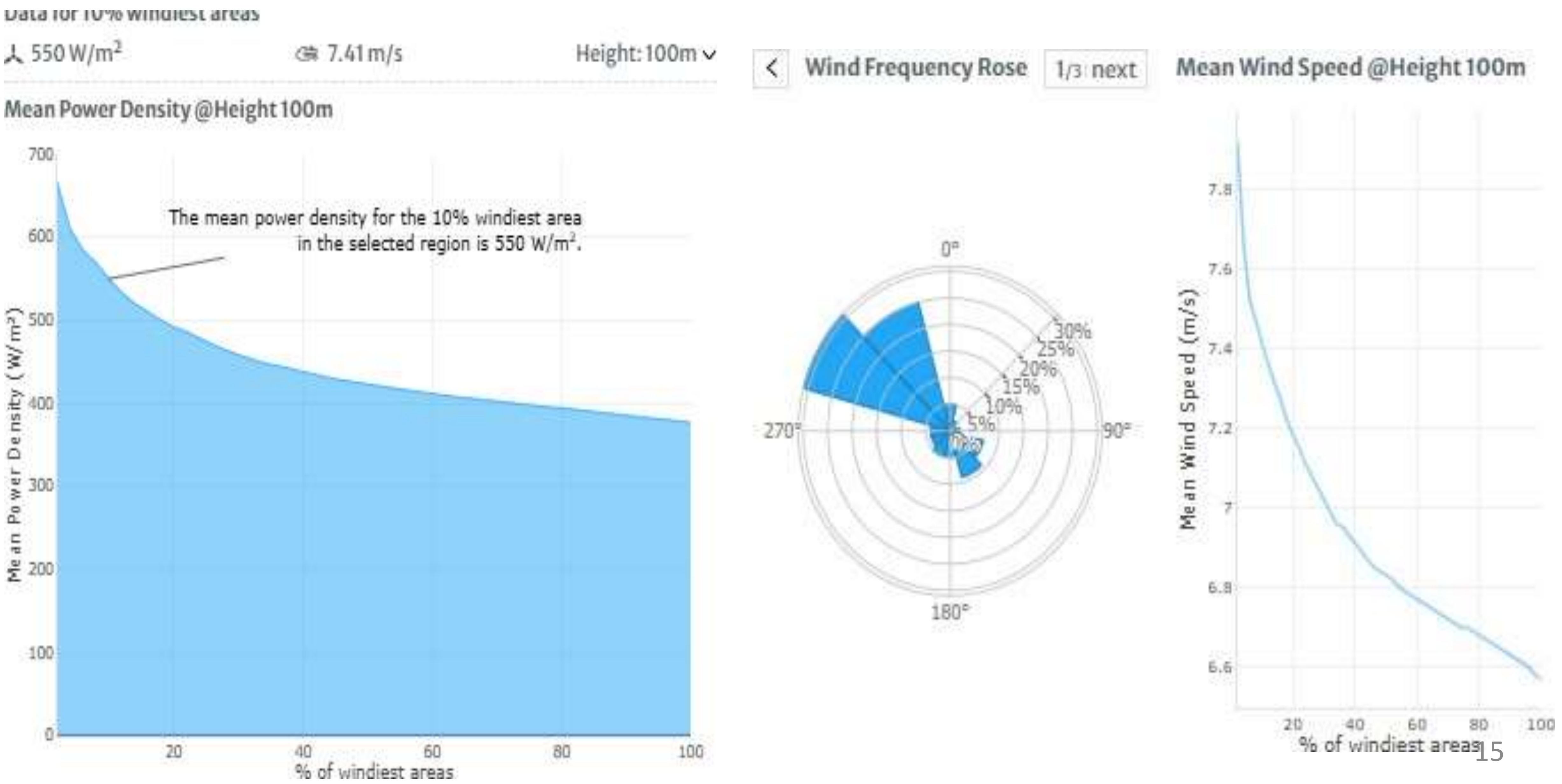
Scale

Mother Wavelet

- Wavelet
  - Small wave
  - Means the window function is of finite length
- Mother Wavelet
  - A prototype for generating the other window functions
  - All the used windows are its dilated or compressed and shifted versions

Findings and Discussion

Statistical Analyses



Mean wind power density and wind frequency and mean wind speed at 100m amsl

Center (Lat, Long): 36.726778°, 27.689667°
Address: Muğla, Aegean Region, Turkey

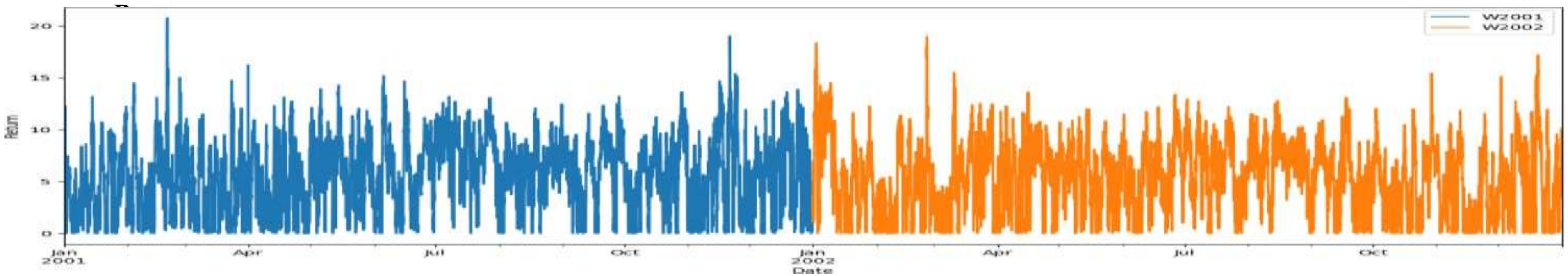


Fig. Hourly wind speed variation in Datça (Muğla) from 2001 to 2002.

In 2001, slightly higher wind speed values were recorded than 2002 at the study area. Times series are very similar with each other.

The figure shows annual wind speed values and their Box Plots. 2001 has higher Wind Speed as Mean and outliers as well than 2002. Both plots show almost bel shape distribution of wind speed at study area.

Fig. Annual comparison of box plots at study area.

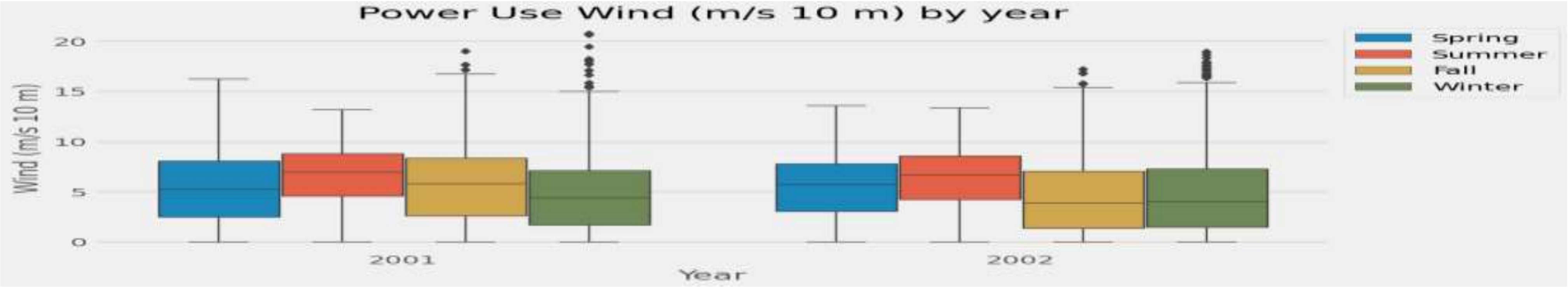


Fig. Seasonal Boxplots shows four season variation: (blue; spring, red; summer, yellow; autumn and green winter. In summer slightly, negative skewness was observed at the both study period. In 2002, fall skewness values are opposite than other seasons, there is a positive skewness.

Wind speed has the highest outliers in fall and winter, however, in summer it has the highest Mean Wind Speed Values of the year.

Monthly Boxplots:

February has the highest outliers of the year however as we see that July and August has the highest mean values for the Wind Speed.

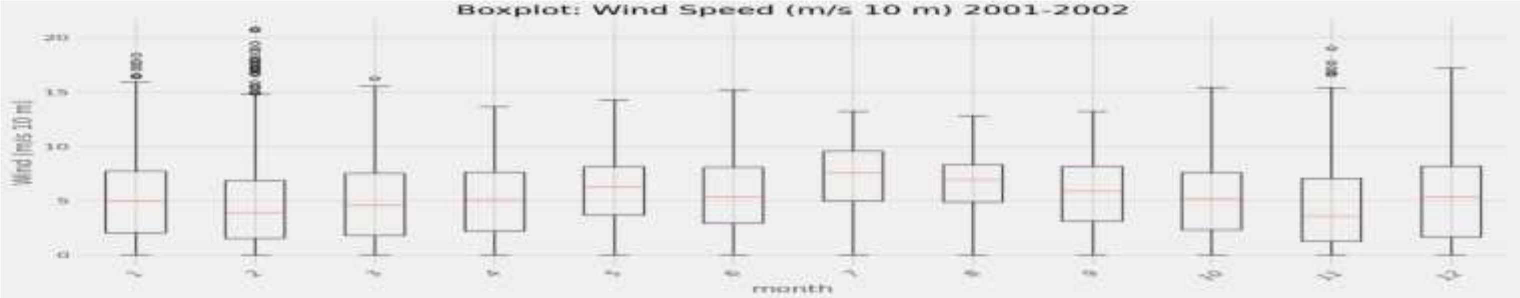


Fig. In February, November and January we have recorded highest outliers of the year, however, wind speed values have highest mean values in July and August.

Time Series Analysis:

Fig. A decreasing trend for wind speed from 2001 to 2002.

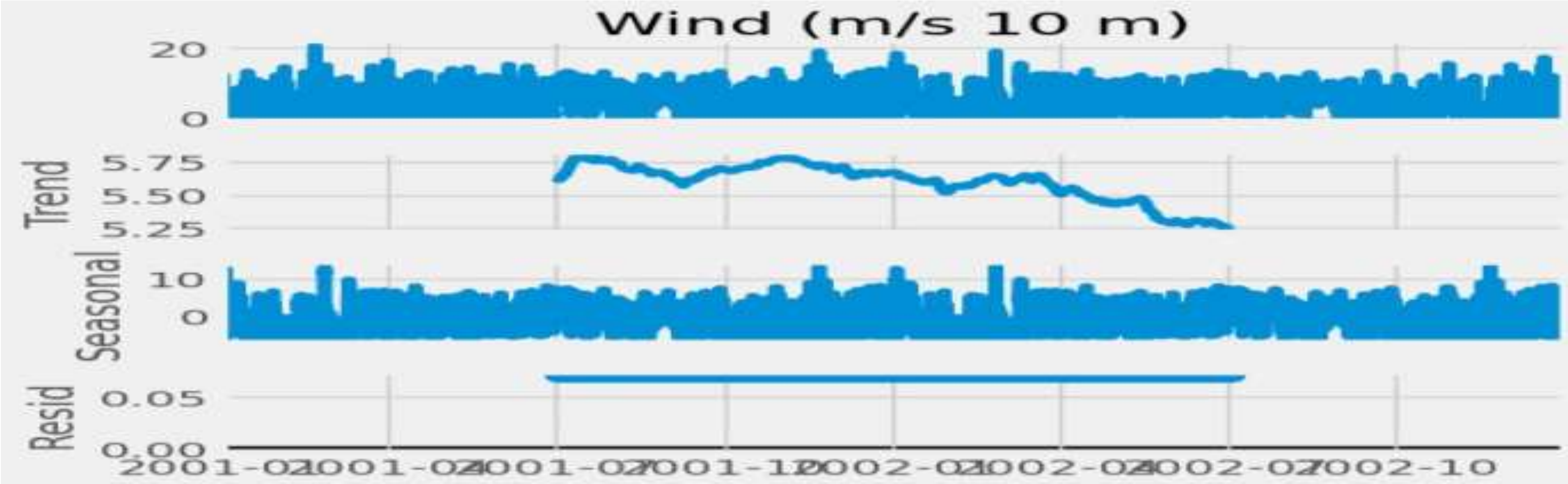


Fig. Time series of two years data (decomposition), hourly wind speed, in Datça.

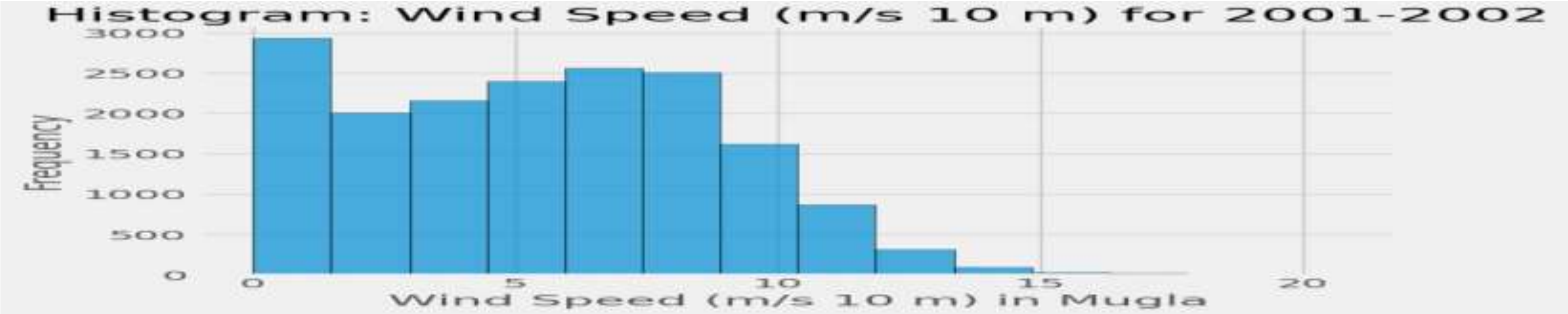


Fig. Frequency histogram of Wind Speed in two years. There is a bi-modal distribution.

Heatmap: It is clear that July 2001 has a very high (unusual) Wind Speed, and for November 2002 it is very low Wind Speed than the usual.

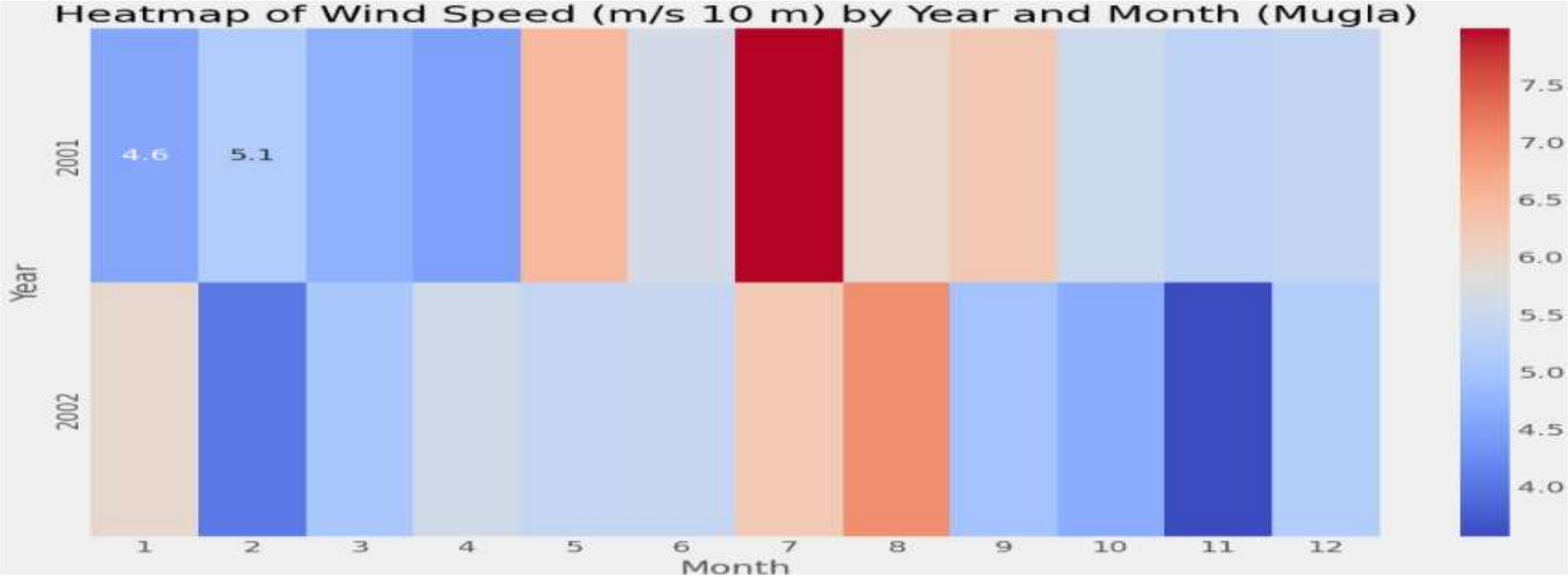
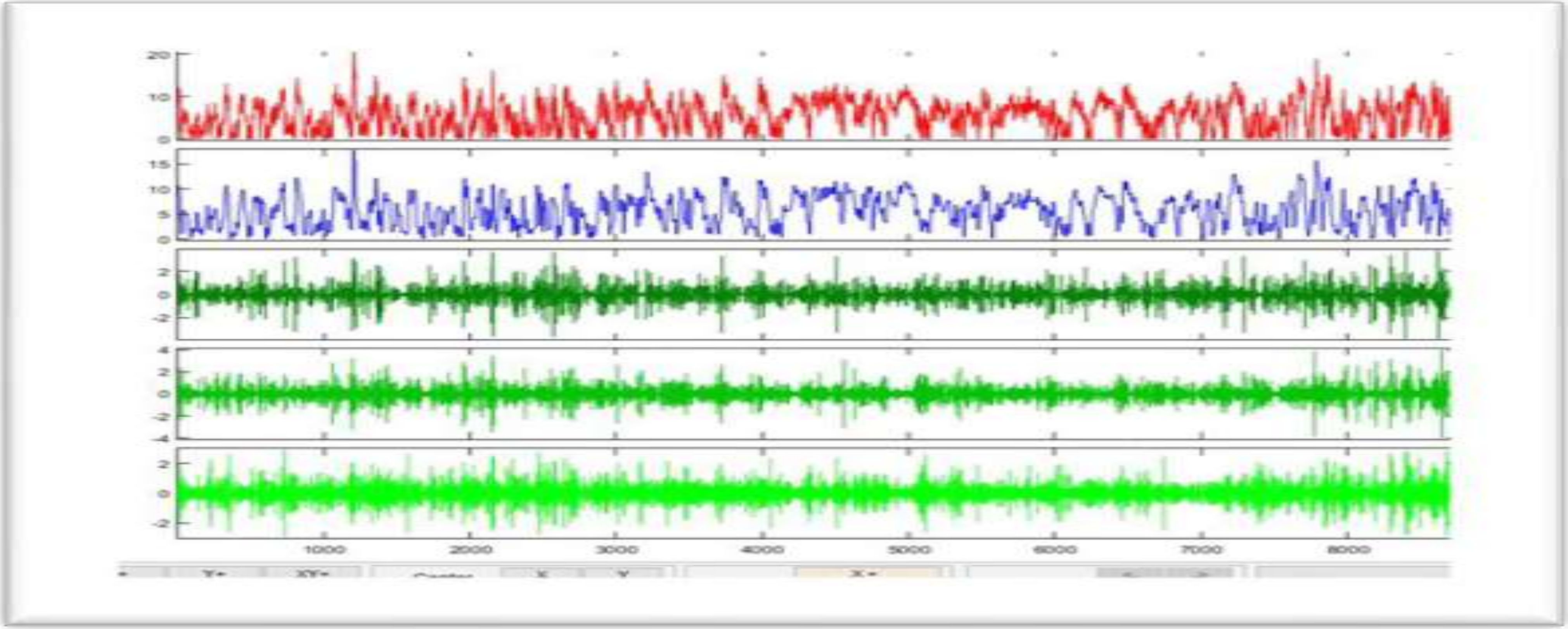
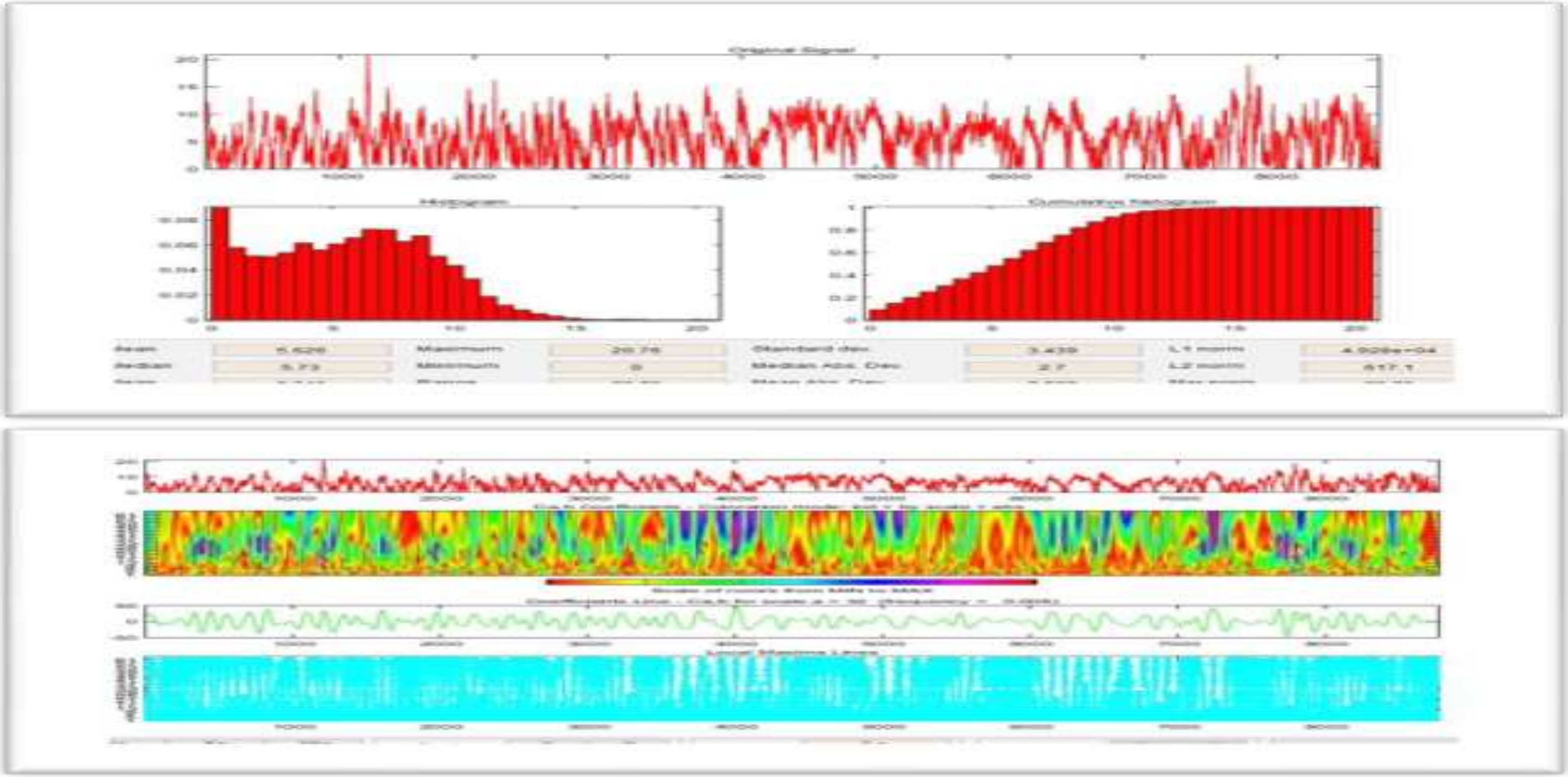


Fig. Heat Map of monthly wind speed in two year.

Wavelet Analyses

1D Wavelet packets and Continuous wavelet (Wavemenu, MATLAB) was applied on hourly wind speed at study area to define role of local, meso and large scale factors on it.





Fig… Wavelet analysis of wind speed in Datça (Muğla), in 2001, (a) 1D wavelet, (b) Statistics and (c) 1D continuous wavelet (db function).

Fig shows row data (red), approximation (blue) and wavelet details (d3, d2, d1) at study area, in 2001

Fig.(a), shows wavelet analyses of wind speed variation in Datça in 2001. High wind speed values have been recorded in winter and late autumn, In summer, role of small and large scale factors decreases. Fig. (b) shows frequency histogram, it has a bi-modal distribution. Fig.(c), explains; in winter and early spring, frequency of large scale events were lower than other seasons. The study area is under combined effect of smal, meso and large scale events. In summer and autumn frequency of large scale events are larger than other seasons. Frequency shows a gradualy decreasing trend through winter.

Results of wind speed analyses show a very similar variation to 2001 in 2002 at Datça, (Figs. a-b). Role of large scale events are less than 2001. Descriptive statistics are also lower in 2002. Extremes have been recorded in winter and late summer at the study area.

2002

	min	max	mean
d1	-3,63	3,63	0
d2	-4,6125	4,6125	-0,00013
d3	-4,66	4,66	-0,00024

In general, energy of maximum wind speed is transferred from small scale fluctuations to meso and large-scale fluctuations in 2002. For average wind speed and its energy show inverse process.

MODEL Results, Simulation

Train-Test Split: 80% of data for training and 20% of data was considered for testing.

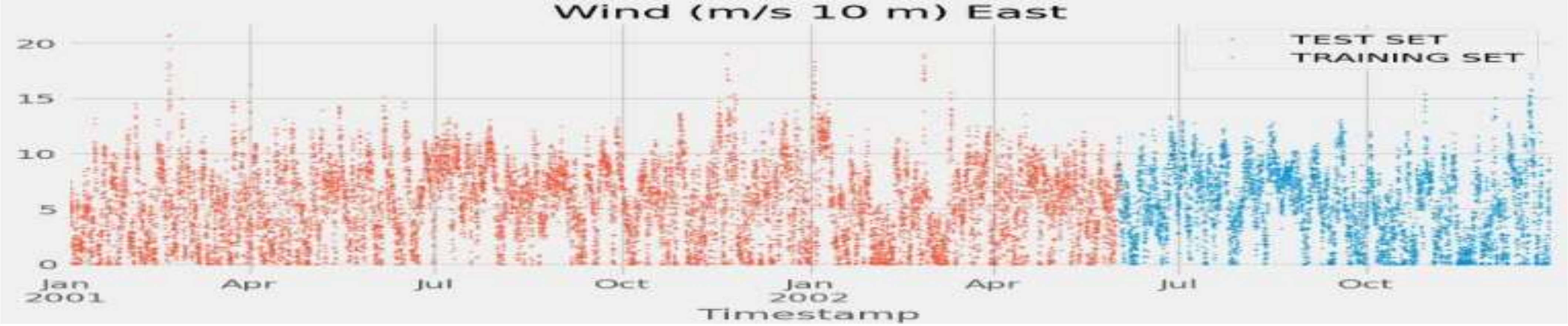


Fig. Train-test split, hourly wind speed in Datça, (2001-2002)

Table. Input data for hybrid modelling

Timestamp	Wind (m/s 10 m)	d1	d2	d3	year	month	hour	day	season	hour_sin	hour_cos	day_sin	day_cos
2001-01-01 00:00:00	9.86	0.635	-0.6925	-0.78500	2001	1	0	1	Winter	0.000000	1.000000	2.012985e-01	0.97953
2001-01-01 01:00:00	8.59	-0.635	-0.6925	-0.78500	2001	1	1	1	Winter	0.258819	0.965926	2.012985e-01	0.97953
2001-01-01 02:00:00	9.93	-0.680	0.6925	-0.78500	2001	1	2	1	Winter	0.500000	0.866025	2.012985e-01	0.97953
2001-01-01 03:00:00	11.29	0.680	0.6925	-0.78500	2001	1	3	1	Winter	0.707107	0.707107	2.012985e-01	0.97953
2001-01-01 04:00:00	10.79	-0.465	-0.2325	0.78500	2001	1	4	1	Winter	0.866025	0.500000	2.012985e-01	0.97953
...	...	...	...	...	...	...	...	...	...	...	...	...	...
2002-12-31 19:00:00	5.01	0.060	0.2475	-1.38125	2002	12	19	31	Winter	-0.965926	0.258819	-2.449294e-16	1.00000
2002-12-31 20:00:00	7.28	0.310	-0.7150	1.38125	2002	12	20	31	Winter	-0.866025	0.500000	-2.449294e-16	1.00000
2002-12-31 21:00:00	8.16	-0.310	-0.7150	1.38125	2002	12	21	31	Winter	-0.707107	0.707107	-2.449294e-16	1.00000
2002-12-31 22:00:00	10.07	-0.740	0.7150	1.38125	2002	12	22	31	Winter	-0.500000	0.866025	-2.449294e-16	1.00000
2002-12-31 23:00:00	10.70	0.740	0.7150	1.38125	2002	12	23	31	Winter	-0.258819	0.965926	-2.449294e-16	1.00000

17520 rows × 13 columns

Optional:

Additional features to data preparation for increasing model performance, (under feature engineering). (Insure the model recognise continuity between the end and start of seasonal cycles).

df['hour_sin'] = np.sin(2 * np.pi * df.index.hour / 24) df['hour_cos'] = np.cos(2 * np.pi * df.index.hour / 24)

df['day_sin'] = np.sin(2 * np.pi * df.index.day / 31) df['day_cos'] = np.cos(2 * np.pi * df.index.day / 31)

LSTM Modelling for Wind Speed in Mugla in 2001-2002

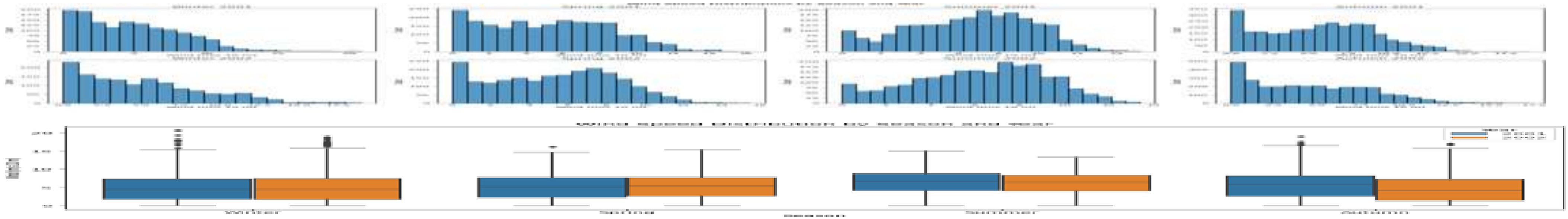


Fig. Frequency histogram and box plot of hourly wind speed for each season, in Datça, in 2001 and 2002.

MODELLING

We used the four machine learning models and we defined errors in the classical machine learning models.

LSTM (Long Short-Term Memory)

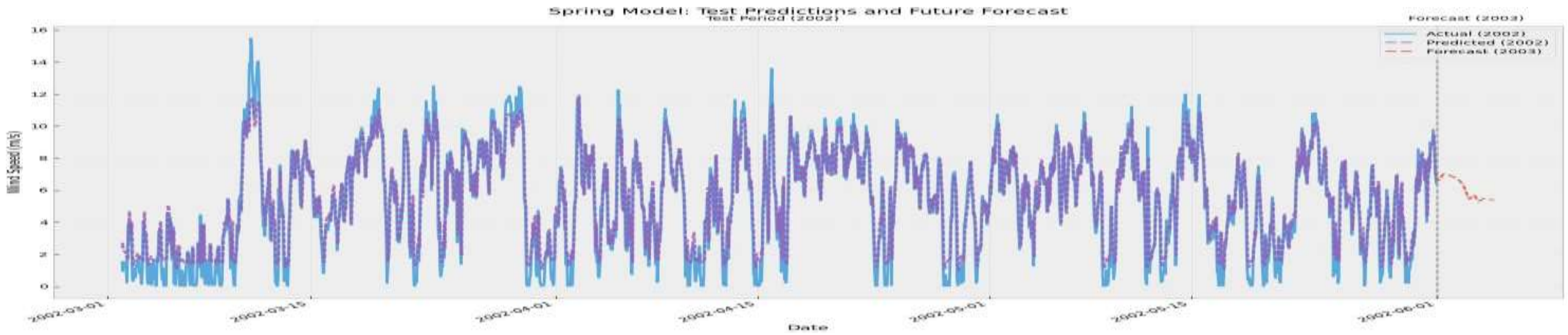
Seasonal LSTM Models for Wind Speed Forecasting

Four Independent Models for Spring, Summer, Autumn, and Winter

We implemented separate LSTM models for each season to forecast wind speeds.

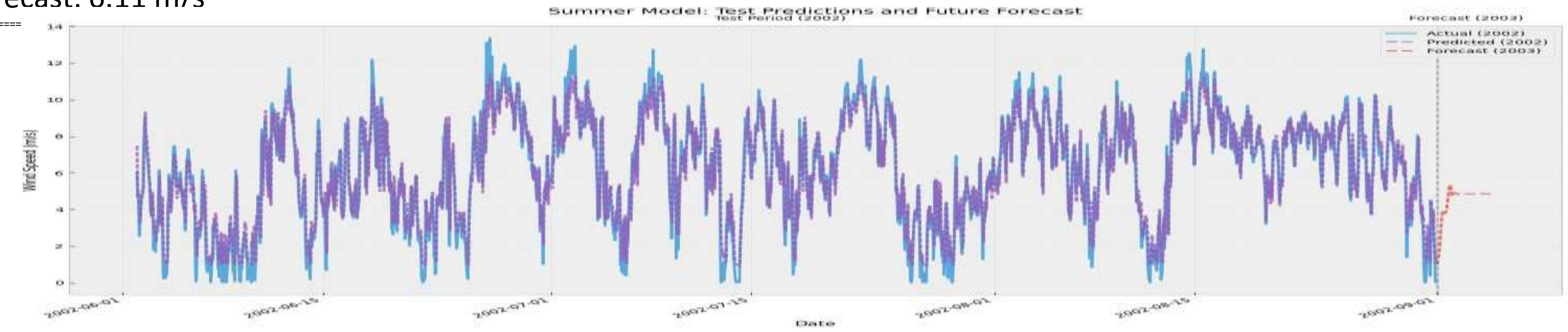
Dataset: Hourly wind speeds from 2001-2002

Goal: Forecast the first 16 hours (short term) of each season in 2003



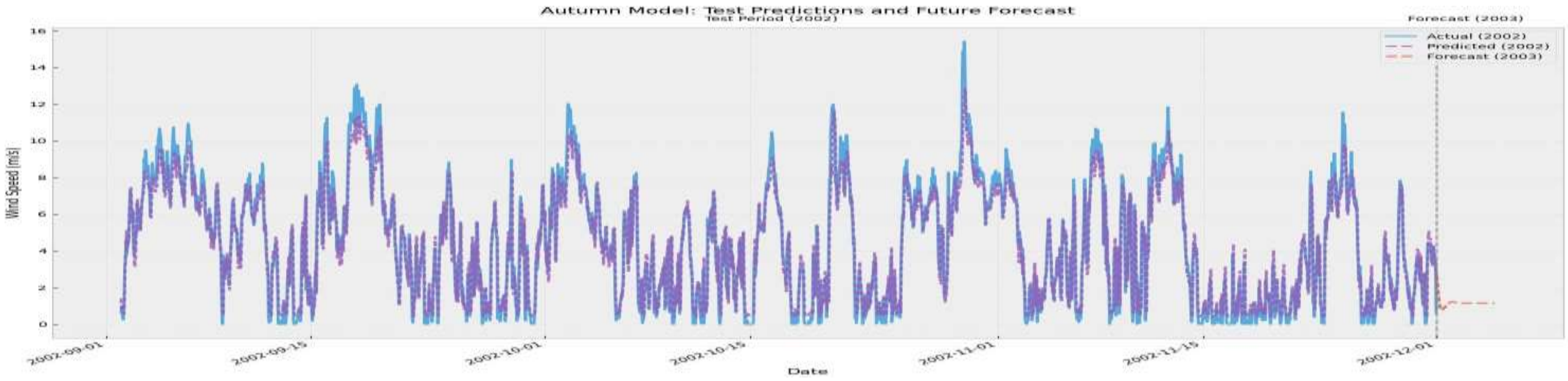
Mean Wind Speed Values:
2002 Actual: 5.38 m/s
2002 Predicted: 5.56 m/s
2003 Forecast: 6.11 m/s

Fig. LSTM-Wavelet in Spring, (2001 and 2002)



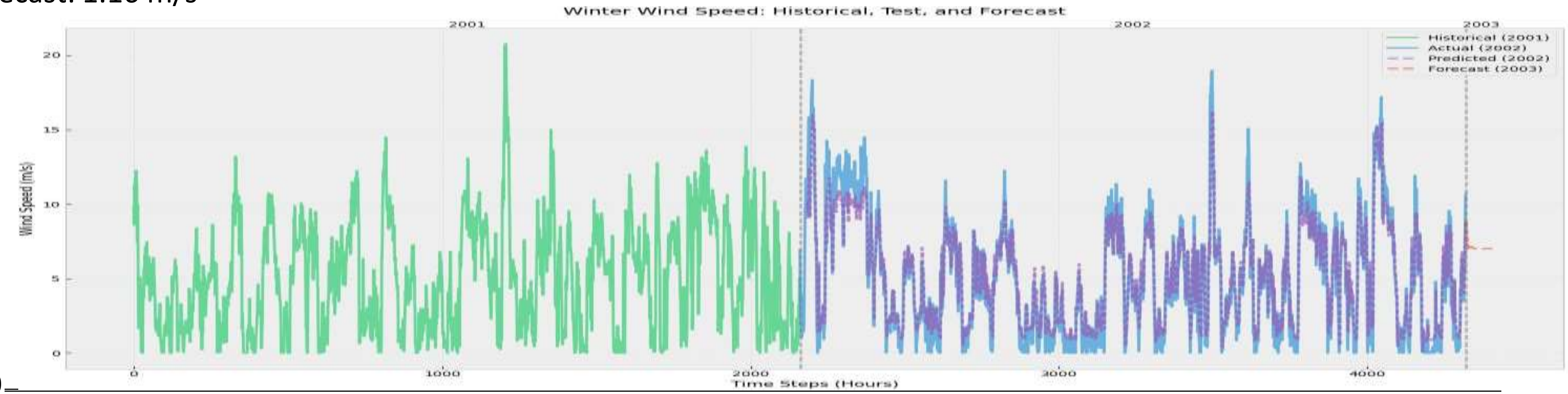
Mean Wind Speed Values:
2002 Actual: 6.20 m/s
2002 Predicted: 6.25 m/s
2003 Forecast: 4.53 m/

Fig. LSTM-Wavelet in Summer, (2001 and 2002)



Mean Wind Speed Values:
2002 Actual: 4.42 m/s
2002 Predicted: 4.31 m/s
2003 Forecast: 1.16 m/s

Fig. LSTM-Wavelet in Autumn, (2001 and 2002)



Wind Speed Statistics:
2001 Mean: 5.03 m/s
2002 Mean: 5.09 m/s
2003 Forecast Mean: 7.10 m/s

Fig. LSTM-Wavelet in Winter, (2001 and 2002)

LSTM-WAVELET	MAE	RMSE	LSTM	MAE	RMSE
Spring	0.93	1.21	Spring	**0.88**	**1.16**
Summer	**0.80**	**1.04**	Summer	0.81	1.06
Autumn	**0.79**	**1.08**	Autumn	0.83	1.12
Winter	**1.05**	**1.38**	Winter	1.32	1.72

In spring, LSTM shows slightly better performance than LSTM-WAVELET after MAE (m/s) and RMSE (m/s). For all other seasons, hybrid model performs better than single LSTM for simulation of short term (16 hr. wind speed).

CONCLUSION

As a conclusion, the Hybrid model (LSTM-Wavelet) shows slightly higher reliability for prediction and estimation of short-term hourly wind speed data at the study area.

LIMITATION AND RECOMMENDATION

The future work will be deal with comparison of different models with longer wind speed data. Longer data will increase precission and accuracy of models.

LIMITATION AND RECOMMENDATION

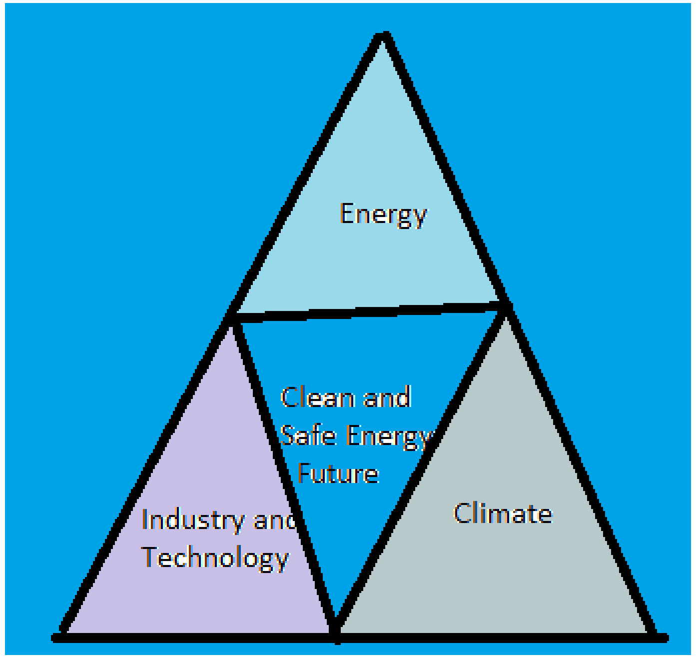


Table. Total wind turbines in Turkey in 2025 (March)

Source	No of Central	Installed Power (MW)
Motorin	1	1
Nafta	1	5
Wind	**376**	13,098
Hard coal	4	841

Türkiye's target of investing 3,500 megawatts of solar energy and 1,500 megawatts of wind energy every year until 2035.

Offshore wind energy system for 2025, 5GW system has been planned.

SDGs, No: 7 Clean Enery

SUMMARY

The utilization of renewable energy resources is mainly related to the problem of the prediction precision of wind speed. Based on wind speed data, the definition of wind energy potential emphasizes ensuring the efficiency and reliability of wind energy systems. The objective of this work is to present a comprehensive review of artificial intelligence (AI) techniques applied to forecast short-term wind speed.

The study compares systematically six different AI models, ranging from machine learning models like high-capacity models that comprise Random Forests, Support Vector Machine, autoregressive moving average (ARMA), Linear and Logistic Regression to more complex models like Long Short-Term Memory (LSTM) networks.

The models are trained and validated against extensive historical records of wind speeds. This paper covers some analysis related to the hourly data collected at the west of the Mediterranean Sea (Mugla City at the Latitude: 556335 and Longitude: 4070184 [ED-50 formats]), between 2001 and 2002.

The study highlights the greater significance of new predictive analytics as a move towards ending the issues of incorporating renewable energy sources into traditional power systems. The study also provides useful insight into AI-based forecasting system design and implementation, paving the way for breakthroughs in environmental applications.

Keywords: Wind energy, sustainability, data driven predictions, machine learning, artificial intelligence

APPENDIX

Pseudo codes: https://github.com/apodwikat/

Acknowledgments:

This study is supported by NATO SPS Multi-Year Project number G5970 named Cube4EnvSec: “Big Earth Data Cube Analytics for Transnational Security and Environment Protection”. The authors also thank to Turkish State Meteorological Service for the data used in the study. The author(s) would like to thank Prof. Dr. Peter BAUMANN, Construction University and Dr. Rene HEISE, NATO HQ: Innovation Hybrid Cyber Division for their support in this paper.

REFERENCES

M AlShafeey, C Csaki, " Adaptive Machine Learning For Forecasting İn Wind Energy: A Dynamic, Multi-Algorithmic Approach For Short And Long-Term Predictions", Heliyon, Volume 10, Issue 15 e34807 August 15, 2024

J Kim, HJ Shin, K Lee, J Hong, " Enhancement Of ANN-Based Wind Power Forecasting By Modification Of Surface Roughness Parameterization Over Complex Terrain", Journal Of Environmental Management, Volume 362, June 2024,- Elsevier

Y Wang, "Deep Learning Techniques in Renewable Energy Forecasting: Solving Intermittency and Instability in Solar and Wind Energy", 4th International Conference on Energy, 2024 - ieeexplore.ieee.org

E. Tan, Ş. S. Menteş, E. Unal, Y. Unal, B. Efe, B. Barutcu, B. Onol, H. S. Topcu, S., İncecik, "Short Term Wind Energy Resource Prediction Using WRF Model For A Location İn Western Part Of ,Turkey", J. Of. Renewable and Sustainable Energy 13, 013303, 09 February, 2021,

U.G.B., Gorgun and Ş.S. Menteş, "Analyzing Wintertime ExtremeWinds over Turkiye and Their Relationships with Synoptic Patterns Using Cluster Analysis" , Atmosphere, Atmosphere 2024, 15(2), 196; PP. 1-26,

REFERENCES

E. Kaya, B.Barutçu, Ş. S. Menteş, “Improving Performance in Deterministic Prediction Methods of Wind Energy with Using Heuristic Methods”, International Energy Raw Materials and Energy Summit (INERMA), Istanbul- Turkiye, 1-3 October, 2015,

Abinet T., Zhang, J. H., Zheng D. H., and Dereje, S., “Short-term wind power forecasting using artificial neural networks for resource scheduling in microgrids,” International Journal of Science and Engineering Applications 5(3), 144–151 (2016).https://doi.org/10.7753/IJSEA0503.1005

Grassi, G. and Vecchio, P., “Energy prediction using a two-hidden layer neural network,” Commun. Nonlinear Sci. Numer. Simul. 15(9), 2262–2266 (2010).

Saleh, A. E., Moustafa, M. S., Abo-Al-Ez, K. M., and Abdullah, A. A., “A hybrid neuro-fuzzy power prediction system for wind energy generation,” Int. J. Electr. Power Energy Syst. 74, 384–395 (2016).https://doi.org/10.1016/j.ijepes.2015.07.039.

https://iicec.sabanciuniv.edu/sites/iicec.sabanciuniv.edu/files/inline-files/IICEC_TEEO_25.pdf • Türkiye Enerji Verimliliği Görünümü Yönetici Özeti (Türkçe ve İngilizce) için: https://iicec.sabanciuniv.edu/sites/iicec.sabanciuniv.edu/files/inline-files/IICEC_TEEO_25_ES.pdf • Türkiye Enerji Verimliliği Görünümü lasnman sunumu için: https://iicec.sabanciuniv.edu/sites/iicec.sabanciuniv.edu/files/inlinefiles/IICEC%20_TEVG_2025_Lansman%20Sunumu.pdf
TÜREB: https://www.tureb.com.tr/

Email:
abdallahdwikat@stu.aydin.edu.tr
, zaferaslan@aydin.edu.tr

Sibel.mentes@.gmail.com

tokgozlu68@gmail.com

LinkedIn:
https://www.linkedin.com/in/abdallah-dwikat-04095927b/

Thank you