Artificial Intelligence Review (2023) 56:S1149–S1198
https://doi.org/10.1007/s10462-023-10554-9

A survey of long short term memory and its associated models in sustainable wind energy predictive analytics

Sherry Garg[1] · **Rajalakshmi Krishnamurthi**[1]

Accepted: 8 July 2023 / Published online: 19 July 2023

Abstract

Sustainable energy is the new normal towards saving the environment, thus resources generating sustainable green energy have gained global attention. Out of all the predominant sustainable energy genres, wind energy is one of the promising and growing solutions to improve efficiency towards sustainability. To expand the area of wind power generation and install more wind farms in future; accurate predictive analytics is mandatory. Due to uncertainty and stochastic nature of wind power time series parameters and outputs, enormous data driven; various machine learning and deep learning approaches have been proposed for the simulation and predictions for wind power predictive analytics. Many approaches have been working towards using Long Short Term Memory (LSTM) and its variants to improve accuracy in wind power predictions. With an aim of easing researchers and applications working in the field of wind power predictive analytics, this study strives to provide critical insights on usage LSTM and associated model in wind power predictions. This study explores at the root level; hence a survey is first made to understand and explore requirements and benefits of time series predictive analytics. Second, a generic exploration of all the different models and performance metrics used over different time series data is performed. Third, a thorough review on WP predictive analysis, based on LSTM as a whole or part of the model is presented. This will also include decomposition techniques, normalization methods, performance metrics, experimented datasets and dependent variable used for wind power predictive analytics. These approaches have been thoroughly seeking to improve the results; however certain challenges still persist due to variability and uncertain nature of wind parameters. Therefore, the major objective of presenting this paper is to learn (i) requirements and benefits of time series predictive analytics, (ii) state of art models and metrics used in time series predictive analytics, (iii) role of LSTM and associated models in wind power predictive analytics, (iv) different decomposition techniques, normalization methods, performance metrics, experimented datasets and predictive frequency used in wind power predictive analytics; and (v) challenges persisting in wind power predictive analytics and usage of LSTM.

Keywords Sustainable energy · Time-series data · Wind power predictive analytics · Deep learning · LSTM

Extended author information available on the last page of the article

Abbreviations

GBM	Gradient boosting machine
DBN	Deep belief network
IoT	Internet of Things
AWNN	Adaptive wavelet neural network
NREL	National renewable energy laboratory
EFG	Enhanced forget gate
CSO	Cuckoo search optimization algorithm
NIWE	National institute of wind energy
MASE	Mean absolute scale error
MAE	Mean absolute error
RMSPE	Root mean square percentage error
RMSE	Root mean square error
MSE	Mean squared error
MAPE	Mean square percentage error
R^2	Coefficient of variation
SATCN	Self-attention temporal convolutional network
EA	Evolutionary attention
CRS	Competitive random search
BP	Back propagation
ENN	Elman neural network
ED	Encoder decoder
GA	Genetic algorithm
TC	Tropical cyclone
FTSNN	Feedback time series neural network
IFTSNN	Input feedback time series neural network
NFMP: Networks	Friends, money, and bytes
NI: Networks	Friends, money, and bytes
KPI	Key performance indicator
RF	Random forest
BBN	Bayesian belief networks
DE	Differential evolution
WNN	Wavelet neural network
RNN	Recurrent neural network
ELM	Extreme learning machine
DWT	Discrete wavelet transformation
FFT	Fast Fourier transformation
LR	Linear regression
NWP	Numerical weather prediction
BR	Bayesian ridge
LSSVM	Least squares support vector machines
BI	Business intelligence
EL	Ensemble learning
GPR	Gaussian process regression
SVR	Support vector regression
BO	Bayesian optimization
CPCB	Central pollution control board
NCRB	National crime records bureau
MFFNN	Multilayer feed-forward neural network

TDNN	Time-delay neural network
RBFNN	Radial basis function neural networks
GRU	Gated recurrent unit
CNN	Convolution neural network
PSBF	Pattern sequence based forecasting
MLP	MultiLayer perceptron
VAE	Variational AutoEncoder
VAR	Vector autoregression
GWEC	Global wind energy council
PSO	Particle swarm optimization
SOA	Swarm optimization algorithm
ACO	Ant colony optimization
BSO	Brain storm optimization
DE	Differential evolution
MWdc	MegaWatts defined conditions
LSSVM	Least-squares support vector machines
MTL	Multitask learning
GMM	Gaussian mixture model
MFO	Moth-flame optimization
CSO	Cuckoo search algorithm
ABC	Artificial bee colony
FA	Firefly algorithm
DE	Differential evolution
RMT	R-matrix with time
MIMO	Multiple input and multiple output
DFF	Deep feed forward
EMD	Empirical mode decomposition
PSR	Phase space reconstruction
AR	Autoregressive
ARMA	Autoregressive moving average
ARIMA	Autoregressive integrated moving average

1 Introduction

The rapidly changing environment, extreme climatic condition and to be aware of the environment protection, intervention from humans is required. The work should progress towards reduction of carbon footprints, limit the usage of fossil fuels and their deliverables (Hossain et al. 2018). Non-renewable resources of power are replenishing, and sustainable energy is the need of the hour. Wind power (WP) is growing as a promising alternative of green energy resource to fossil fuel-generated electricity and recently started to boom and gained attention as one of the cleanest and readily available resources (Report 2020; Global wind energy council 2022). However, the nature of wind is instantaneous and highly variable even in an ultra-short span of time. To make a robust dependency of the grid over WP production resources, to create a balance between scheduling the power from wind resource or fossil resource, to optimize the energy reservoir usage and to plan the power distribution from different energy generation resources; accurate WP predictive analytics is needed (Chang 2014). Also, to plan the day-ahead schedule of power trading against power generation resources WP predictions play an important role. Hence, various studies have been devoted by experts in the field and extensive researches are proposed to improve the WP predictions. Variety of assessment models and metrics are developed. According to literature, there exist two types of forecasting namely qualitative and quantitative as shown in Fig. 1. Wind energy (WE) data represents continuous sequential type of time series data and falls under the quantitative type of predictive analytics. Also, the time series forecasting in the perspective of WE forecasting includes various research paradigms.

To begin with time series data, this section presents research paradigms available in time series data and WP. Further, one of the research paragons, WP forecasting, is elaborated w.r.t. state of art. Motivation, scope and objective of this study are stated. To understand different aspects of WP and its research areas, a detailed taxonomy of WP forecasting is discussed.

1.1 Time series data

Definition: A sequential ordered observations collected regularly at a fixed frequency of time like sensors data denoted as, X_t where ($t = 1,..., N$) (Maçaira et al. 2018)

1.1.1 Type of time series

The data can be univariate time series data or multivariate time series data.

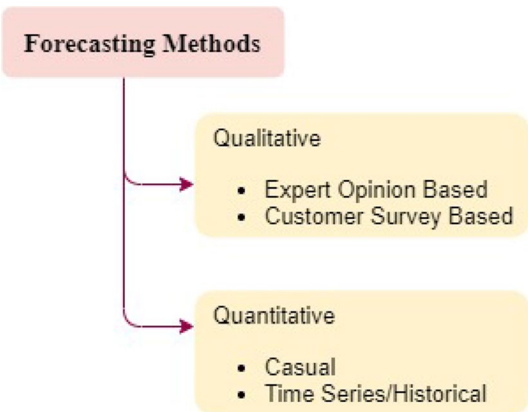
- **Univariate time series:** A time series of the form $x_i \in R$ with a single attribute dependent over time.
- **Multivariate time series:** A time series of the form $x_i \in R \times R$ with multiple attributes dependent over time

1.1.2 Research areas in time series

There are different research areas that can be explored in time series:

- **Time series classification (TSC):** Classifying the current time sequence into its designated class C_i where total number of classes $n \geq 2$.

Fig. 1 Types of forecasting



- **Time series regression (TSR):** Predicting univariate or multivariate values for future timestamps.
- **Time series clustering (TSCl):** Clustering time subsequences into its design nated cluster Ci, where total number of clusters n ≥ 2, based on similarity measure. Number of clusters can be pre-decided or can be calculated on the go.
- **Time series anomaly detection (TSA):** Assigning a time sequence to one of the classes $Ci = C_{normal}$, $C_{anomaly}$, where C_{normal} represents the majority class of normal state observations, while $C_{anomaly}$ represents the class of rare observations, i.e., anomalies.
- **Time series forecasting (TSF):** Predicting next or future value(s) of the timestamp in a univariate way or multivariate way of predictive analytics also known as Time Series Predictive Analytics.

1.2 Motivation and contribution

As per GWEC 2022 statistics (Global Wind Report 2022) both U.S. and India has achieved growth in WP capacity installation and thereby reduced the tariffs over the years. A positive rise in global wind industry can be stated from the fact that despite COVID-19 continued its second year, 94 GW of new capacity was still added to the growth in 2021. Also, annual global market is expected to grow from 21.1 GW to 31.4 GW from 2021 to 2026. This will increase the contribution of WE from 22.5 to 24.4% in the year 2026. However, in offshore sector more than 90 GW of is expected to be added globally from 2022 to 2026. As per Figs. 2 and 3, U.S. installed capacity has gone up from 848.84 MWdc in 2010 to 23,564.63 MWdc in 2021 and reduced the tariffs from 5.79 to 1.38 \$/W. Similarly, India has reduced tariffs for both WP and solar power by 50–68%.

GWEC mentions in a published report that by 2030 WE systems could provide one-fifth or 20% of the global electricity demand (Energy 2016). WP capacity increased by 7% in 2019 to reach 645 GW from 599 GW in 2018 (Schmela et al. 2018; Pitteloud 2020). WE is

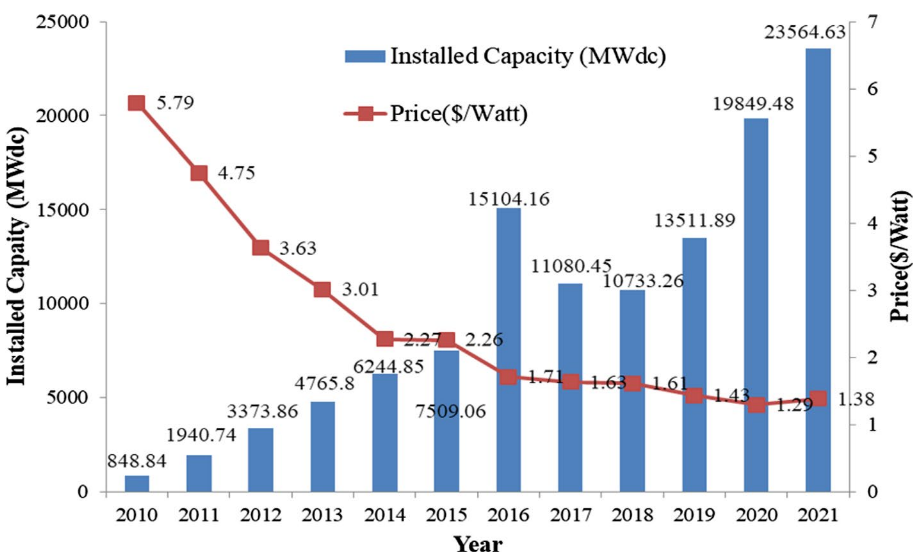


Fig. 2 U.S. Solar Power Tariffs and Installed Capacity from 2010 to 2021

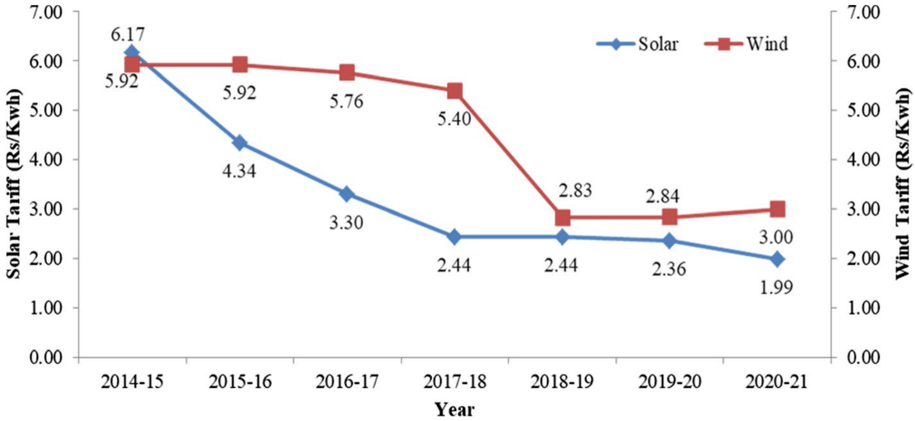


Fig. 3 India Solar and Wind Power Tariffs from 2014–2015 to 2020–2021

estimated to take over 18% of global power by 2050 (Pitteloud 2020). The facts like globally 30 countries have installed more than 1 GW of wind energy capacity and 9 countries have installed more than 10 GW capacity (Schmela et al. 2018; Pitteloud 2020) are the proof of WP global expansion.

These statistics are enlightening about the growing sector WP energy and requirement of better predictive analytics for more installations and lower tariffs. Also, better grid synchronization across all sectors of power system can become more efficient. A latest, single point critical survey can be a great help for developers and researchers in enhancing the existing literature without reinventing the wheel. Hence, these facts motivated this study to contribute in terms of trying to cover the maximum amount of literature for a better understanding of the problem domain and its proposed solutions as inspired from (Zafar et al. 2023).

1.3 Scope and objectives

To meet the demand and growth ratio, the inclination towards WP generation research is increased so that power generation through wind can be maximized (Ramadan 2017). The research paradigms in WE have wide branches namely, (i) WP prediction (Study et al. 2017), (ii) wind speed prediction (Han et al. 2017), (iii) wind farm layout optimization (Ramadan 2017), concerning maximized power and minimized cost. From these areas, WE prediction, which in turn has distributed areas like spot prediction or deterministic point prediction (Yu et al. 2018; Osório et al. 2018; Fu et al. 2019), probabilistic interval prediction (Du et al. 2019; Hong et al. 2016; Nowotarski and Weron 2018) and interval prediction (Li and Jin 2018; Zou et al. 2019), is the urging need for expanding the sustainable energy. Hence, out of all the optimization areas, this survey will concentrate on WP predictive analytics. Further, in WE predictive analytics, there are various branches in terms of intervals depending on wind farm historic data, models proposed and evaluated and, methods used. The whole widened taxonomy of WP predictive analytics is presented in Fig. 4.

The non-linear nature of data embedded with time factors is inevitable in real-time wind data. The WP signal is quite random in nature due to variable high-frequencies. This random nature of WP, if predicted well in time, can lead to growth of sustainable energy

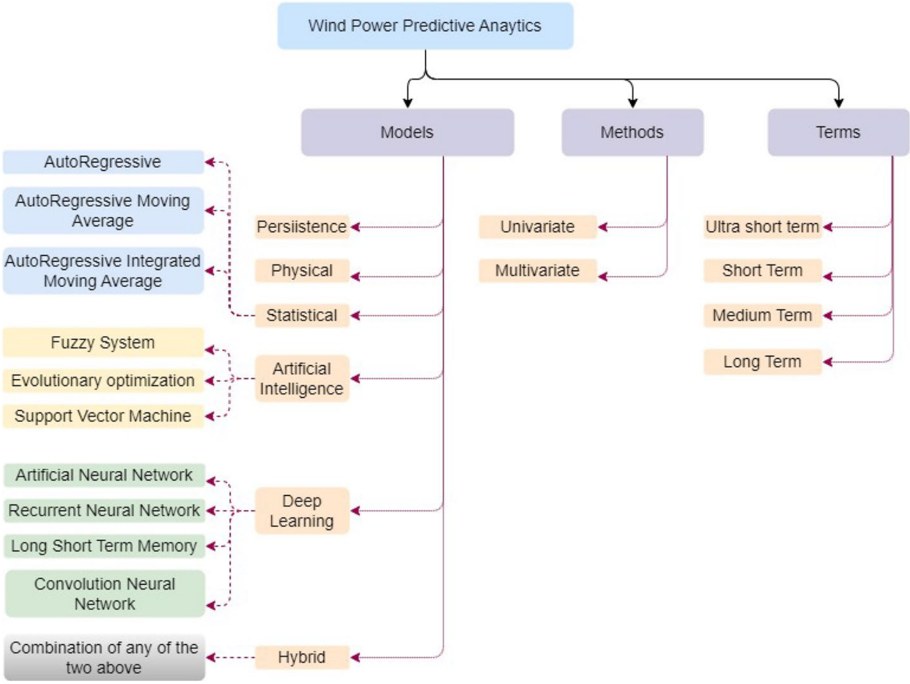


Fig. 4 Wind Power Predictive Analytics Taxonomy

development. Hence, there is a need to understand various aspects of WP forecasting. It involves study of (i) time series attribute of WP data, (ii) predictive analytics requirement on time series data as Table 1, (iii) different models, techniques and approaches existing for predictive analytics of WP data as Table 2, and (iv) detailed state of art study of LSTM and its associated models with the performance evaluation as Tables 3 and 4

1.4 Key contributions

The key contributions of this paper are:

- To extensively understand the requirement and benefits of time series data and time-series predictive analysis.
- Exhaustive study of various time-series datasets, different models applied over these different datasets and their quantitative performance.
- A rigorous taxonomy of WP predictive analytics based on literature since 2016.
- Comprehensive role of LSTM and its associated models over wind time series data to forecast power, their performance, contributions and demerits.
- Detailed list of challenges in state of art WP predictive analytics and LSTM.

This paper is organized in sections in the sequence of these key contributions. Section 2 describes (i) areas/applications in the literature that are benefitted from the time series predictive analysis and the associated benefits, (ii) study of various time-series

Table 1 Requirement and importance of predictive analysis in time-series data

SN	References	Publication detail	Time series data used	Requirement/reason of predictive analysis
1	Yang et al. (2017)	T-Yen Yang, 2017	Two MOOC courses' CFA average score of students	Predictions on student grades and analysis on grade progression outcomes to implement any preventive measures required for extreme graded students
2	Javid et al. (2020)	Alireza M., 2020	Cumulative number of reported COVID-19 cases collected daily for 12 different countries by Johns Hopkins University	Prediction on next 14 days COVID cases for risk mitigation by health departments
3	Wang et al. (2018)	Chih-Hsuan Wang, 2017	Financial and non-financial, total 25 indicators	Predictive analysis on computer firms to characterize depending on business models like brand/design/electronic service manufacturing
4	Filonov et al. (2016)	P Filonov, 2016	Sensor and control, continuous and discreet channels fault test set for analysis for industrial fault detection	Predictive analysis on fault behaviour in the industrial plant to generate roots and causes of fault generation
5	Redondo-Bravo et al. (2019)	Lidia MD, 2019	19 years hospital discharge data from January 1997 to December 2016	To estimate Dengue cases in Japan shortly, location and time-wise
6	Mustaffa et al. (2018)	Zuriani M., 2018	Daily data recorded at Kuala Lumpur, Malaysia, from April 2016 to May 2017	Five days advance weather prediction to take predictive measures against extremes
7	Geetha (2016)	A.Geetha, 2016	Six years of modeler data from 2007–2012	Tropical cyclone (TC) has hazardous effects on human lives and economy in coastal areas. Predicting a TC well in advance can help risk mitigation and disaster management
8	Mehdiyev et al. (2017)	Nijat M, 2017	89 variables and 3 terabytes sensor data from steel	Predictive analysis on next production process step of a steel industry using the quality of the semi-finished products
9	Zhang et al. (2019)	Q. Zhang, 2019	15-min interval based waste water flow rates of year 2015–2016	Predictive analysis on amount of incoming wastewater in wastewater treatment plants (WWTPs) and do efficient management of WWTPs

Table 1 (continued)

SN	References	Publication detail	Time series data used	Requirement/reason of predictive analysis
10	Pavlyshenko (2019)	Bohdan M., 2019	store sales historical data from Rossmann Store Sales	To incorporate business intelligence for efficient sales patterns. To take prescriptive measures for exogenous factors impacting sales
11	Sarveswararao et al. (2023)	Vangala Sarv, 2023	Daily cash withdrawal data from 21–11-2017 to 20–11-2019 of 100 ATMs	ATM cash withdrawal of a large Indian commercial bank and forecast the withdrawals using DL and hybrid DL methods
12	Deep et al. Jun. (2022)	B. Deep, 2021	From CPCB of India, 3 years Delhi's IGI Airport air monitoring station's pollution data	To plan a less polluted route enhanced pollution prediction ahead of time at specific locations in urban cities and to know AQI well in advance
13	Chen et al. (2022)	C.Chen, 2022	monthly rainfall data for a period of 41 years (1980–2020) from two meteorological stations in Turkey, namely Rize and Konya	Due to inadequate rainfall Turkey is drought prone which leads to water scarcity and affect agricultural industry. So, accurate prediction of rainfall over the concerned region is helpful to be prepared for the measures well in advance
14	Liu et al. (2021)	M. Liu, 2021	Daily price of Bitcoin and the cryptocurrency market determinant collected from 'www.coindesk.com' and 'BTC.com'	Economic growth and trading help is developed by forecasting crypto currency price to invest with a calculated risk
15	Rogachev (2022)	AF Rogachev, 2022	1950–2018 data of grain crops in the conditions of the Volgograd region of Russia	Impact of solar and hydro-meteorological conditions of Lower Volga region of Russia for planning and management of cultivated crops
16	Park et al. (2022)	H. Jun Park, 2022	S&P500, SSE, and KOSPI200–stock markets with diverse patterns, five primary indicators (open, high, low, close prices, and total volume [OHLCV]) and 38 auxiliary indicators	predicts the next day's stock return and return direction using 43 technical indicators with 50 time-steps to increase the profit in investments

Table 1 (continued)

SN	References	Publication detail	Time series data used	Requirement/reason of predictive analysis
17	Nguyen et al. (2021)	Le Quyen Nguyen, 2021	Vietnam's tourism data (1) monthly number of international tourists from Jan 2008 to Dec 2020 (2) annual number of international tourists from Jan 1995 – Dec 2020 (3) number of international tourists by mode of transport, from 1995 to 2020 (4) number of international tourists by region, from 2008 to 2020; and (5) average length of stay and expenditure of international tourists from 2005 to 2020	Due to pandemic COVID-19, Vietnam tourism was bearing losses. To help recover the tourism industry specially international and increase employment and incomes of tourism industry predictions for monthly number of international tourists to the country are done
18	Lazzari et al. (2022)	F. Lazzari,2022	500 residential south-eastern region of Spain users consumption, weather and calendar data	To forecast residential households day-ahead electricity consumption with highly irregular human behaviour and plan electricity distribution accordingly
19	Essien and Giannetti (2019)	Aniekan Essien, 2019	3 univariate time series datasets: (i) The Chickenpox, (ii) minimum daily temperatures and (iii) PeMS daily traffic flow	Uni and multi variate power predictive analytics help planning the power stages efficiently
20	Ma and Faye (2022)	Tai-Yu Ma, 2022	Dundee, UK, EV station data charger identifiers, start and end times of charging sessions, amount of energy charged, the power of chargers, and the geographical locations of chargers from March 5, 2018, to June 4, 2018 (91 days)	To implement smart charging stations to EV users and operators, predictive analytics is required
21	Nagaraj (2022)	P. Nagaraj, 2022	12 years from 2005 to 2017 of cyber hacking operations data involving malware assaults	To understand how the cyber threats and attacks are evolving and implement risk mitigation measures well in time forecasting of cyber-attacks is needed
22	Muthamizharasan and Ponnusamy (2022)	M. Muthamizharasan, 2022	3 years crime dataset taken from the NCRB	forecasting crime rate, location and time can set up timely suitable vigilance alerts to reduce the crime rate

Table 1 (continued)

SN	References	Publication detail	Time series data used	Requirement/reason of predictive analysis
23	Livieris et al. (2020)	Ioannis E. Livieris, 2020	Daily gold price collected from Jan 2014 to Apr 2018 from http://finance.yahoo.com	predictive modelling to predict gold price has a significant impact on financial activities. This could offer insights in gold price fluctuations, behavior and dynamics and ultimately could provide the opportunity of gaining significant profits

Table 2 Models and performance metrics for time-series predictive analysis since 2016

SN	References	Publication detail	Type of prediction	Time series dataset used	Model used	Performance metric computed
1	Do et al. (2019)	Loan N. N. Do, 2019	Survey of Models on Short-term traffic state prediction	Real time travel data collected via different methods like (i) loop detectors, (ii) probe vehicles, (iii) GPS equipment, and (iv) RTMS	MLFNN, TDNN, Basic RNN, LSTM, GRU, CNN, DBN, RBFNN, WNN and FNN, GLU	Average error 4.3% in ARIMA, in MLFNN + 1 hidden layer was 14% higher than deep model, PCA + 3 hidden layers MAPE 2.8, RBFNN MAPE 3.37%, WNN + GA + clustering MAE 11.986, TDNN average errors is 3–4%, LSTM MAPE 5.89% and GLU-CNN: MAPE at 5.4%
2	Bao (2017)	Wei Bao, 2017	Yearly	8 years daily finance data of six stock market indices from Jul. 2008 till Sep, 2016	Combination of WT, SAEs, and LSTM	Among six stock indices **Average MAPE** from 0.018 to 0.066
3	Qin (2019)	Hongqian Qin, 2019	Hourly	64,411 data points at 30 min intervals of dissolved oxygen from Aug, 2012 till May, 2016	CNN, TCN, LSTM, GRU, BiRNN	Best Performance metric is of GRU Worst Performance metric is of LSTM
4	Zhu and Laptev (2017)	Lingxue Zhu, 2017	Daily taking 28 sliding window	8 cities daily completed trips data collected for four years	LSTM layered encoder-decoder with 128 and 32 hidden states and 3 FC layers – (128, 64, and 16 hidden units with tanh activation)	SMAPE LSTM: 9.2 and Proposed model: 5.9

Table 2 (continued)

SN	References	Publication detail	Type of prediction	Time series dataset used	Model used	Performance metric computed
5	Cui et al. (2018)	Zhiyong Cui, 2018	Short term	Two traffic datasets: (i) Loop detector dataset, data of 323 sensor stations in 2015, timestep 5 min, (ii) INRIX data, the year 2012 interval 5 min	SBU-LSTMs comprised of 1-input layer and **N** middle layers of BDLSTM layers with an LSTM layer	Best performance is of **SBU-LSTMs with MAE 2.42.** MAE of comparative models is listed as: SVM 9.23, RF 2.64, FFNN (2-hidden layers) 2.63, GRU 3.43
6	Torres et al. (2019)	J. F. Torres, 2019	Daily with a sliding window of 1 day	2 years 7 a.m. to 5 p.m. dataset, from Jan 1, 2015, to Dec 31, 2016, in 30 min intervals. 14,620 samples	DL with no. of hidden layers 1 to 5 and no. of neurons from 10 to 40	NN PSF **DL** RMSE 154.16 149.52 **148.98** MAE 116.64 119.17 **114.76**
7	Akita (2016)	Ryo Akita, 2016		Nikkei 225, 10 companies data published from 2001 to 2008	LSTM: minibatch size is 30, LSTM with one layer and were unrolled 20 steps, and 50% dropout on the non-recurrent connections, adam as an optimizer for 50 epochs	Trade Gains SVR MLP RNN **LSTM** −46.96 −561.14 262.54 **1211.90**
8	Zeroual et al. (2020)	Abdelhafid Zeroual, 2020	17 days- forecasts based on 148 days window	Daily confirmed and recovered cases collected from Spain, Italy, China, USA, and Australia from 22 Jan 2020 to Jun 17, 2020	Comparison of 5 DL methods namely, RNN, LSTM, BiLSTM, GRUs and VAE, to forecast count of new and recovered COVID cases	VAE model outperformed when compared with RNN, LSTM, BiLSTM, GRU on the basis of error measure
9	Hua et al. (2018)	Yuxiu Hua, 2019	Daily taking 100 days window	Traffic data sampled every 15 min during four months with 10,772 traffic data points	RCLSTM	RMSE of SVR 0.075, ARIMA 0.078, FFNN 0.073, LSTM30 0.07, RCLSTM 0.064 and **LSTM300 0.047**

Table 2 (continued)

SN	References	Publication detail	Type of prediction	Time series dataset used	Model used	Performance metric computed
10	Vicente et al. (2017)	J.F. Torres, 2017	Short term – 4 h	from January 2007 to June 2016 every 10 min, time-series is composed of 497,832 measurements	DL with three hidden layers and 10 to 100 neurons, a step of 10, and a learning rate of 0.001	MRE of the proposed hybrid model is 1.84%
11	Li et al. (2019)	Youru Li, 2019	–	(i) Beijing PM2.5 hourly data from Jan 1, 2010, to Dec 31, 2014. (ii) SML2010 40 days of data sampled every minute, (iii) MSR Action3D data with 567 depth map sequences	Multivariate time series predictive analytics using Evolutionary LSTM (attention-based) with random search	**MAE** Beijing PM2.5 SML2010 SVR 2.6779 0.0558 GBRT 0.9909 0.0253 RNN 0.8646 0.0261 GRU 0.6733 0.0231 LSTM 0.6168 0.0178 Attention-LSTM 0.2324 0.0190 DA-RNN – 0.0150 EA-LSTM 0.1902 0.0103
12	Qian (2017)	Xin-Yao Qian, 2017		S&P 500, Dow 30, and Nasdaq stock indices weekly data ranges from Jan 2, 2012, to Dec 26, 2016	ARIMA, LR, MLP, SVM and DAE-SVM	Hit Ratio: ARIMA 0.593. LR 0.623, MLP 0.63, SVM 0.642 and DAE-SVM 0.672
13	Kim and Kang (2019)	Sangyeon Kim. 2019	Weekly	Financial: the 200 most traded securities in the Korean stock exchange	MLP, 1D CNN, SLSTM, attention networks, and weighted attention networks	Attention Networks Hit Ratio 0.715 and Weighted attention Networks 0.763
14	Eapen et al. (2019)	Supriyo Chakraborty, Jithin Eapen, 2019		The daily closing price of the S&P 500 index from Jan 2, 2008, to Nov 27, 2018, for five days a week for ten years	Multiple pipeline CNN and bi-directional LSTM model	Mean Test Score Mean Train Score 0.000281317 0.000204378

Table 2 (continued)

SN	References	Publication detail	Type of prediction	Time series dataset used	Model used	Performance metric computed
15	Xiao et al. (2019)	Changjiang Xiao, 2019	Short and mid-term	Daily OISST-V2-AVHRR data from Jan 01, 1982, to Dec 12, 2017 (13,149 days in total)	LSTM and LSTM-AdaBoost	The biggest RMSE improvement compared to LSTM is 2.26%, and AdaBoost is 4.61%
16	Heidari and Khovalyg (2020)	Amirreza Heidari, 2020	Short term	solar-assisted water heating hourly data used from Feb 2015 to Feb 2016 for one year, a total of 8400 samples	Combination of attention layer with time series decomposition with LSTM layer	The maximum value of absolute error by proposed model lowered by 8.2% from Attention-based LSTM, 4.8% from LSTM, and 17.8% from FF
17	Liu et al. (2019)	Chien-Liang Liu, 2019		prognostics and health management 2015 challenge data	MVCNN for multivariate time series classification	MVCNN Accuracy 0.9740 and Random Forest 0.9505

Table 3 LSTM and associated models used in wind power predictive analytics to date

SN	Paper	Author	Type of WP prediction frequency	Performance metric	Model used	Dataset used	Decomposition method (if any)
1	Zhou et al. (2020)	Baobin Zhou (2020)	Ultra short-term	MAE, RMSE, MAPE	LSTM with SARIMA PCA hybrid model and EEMD data decomposition	35,040 samples (from 15 Jan 2017 to 27 Feb 2018)	EEMD
2	Duan et al. (2020)	Jiandong Duan (2020)	Short term	RMSE, MAE, TIC, and Time	hybrid forecasting model based on LSTM with Correntropy combining an IVMD and Sample Entropy (SE)	China Wind farm Dataset sampled every 5 min from July 1 to July 9, 2019	IVMD
3	Srivastava et al. (2020)	Tushar Srivastava (2020)	Short term		Basic RNN, GBM contrasted with LSTM	Over a span of one year 1st Jan to 31st December of 2014, hourly data of Kolkata wind turbine	None
4	Han et al. (2019)	Li Han (2019)	Long term		LSTM	2 years of ELIA data is from 2015 to 2017 and NREL data is from 2011 to 2012 at an interval of 15 min	VMD
5	Liu et al. (2020)	Bingchun Liu(2020)	Short term	MAPE	WD + LSTM	power generation data of 40 data points from the seventh month of 2009 to the fourth month of 2019 are selected	WD

Table 3 (continued)

SN	Paper	Author	Type of WP prediction frequency	Performance metric	Model used	Dataset used	Decomposition method (if any)
6	Sun et al. (2019)	Zexian Sun (2019)	Short-Term(15 min and 2 h)		Data decomposed using VMD + K-Means Clustering + LSTM	Two wind farms data from China for the last quarter of 2012 and the first quarter of 2013	VMD
7	Liu et al. (2019)	Y Liu (2019)	Short-Term (15 min)	MAE/MW MAPE/% RMSE/MW	DWT and LSTM (DWT_LSTM)	3-time series wind power data at 15 min period of 12 months each	DWT
8	Choi and Lee (2018)	Jae Young Choi (2018)			LSTM ensemble forecasting algorithm	From publicly available data at Time Series Data Library (TSDL), River flow, Vehicles, Wine, and Airline, datasets	–
9	Shi et al. (2018)	Xiaoyu Shi (2018)	Hourly day-ahead wind power predictions	MAE, MAPE, and RMSE	recursive LSTM and direct-VMD-LSTM models	Two wind power series of hourly data: summer and spring data from (1–31 August 2017) and (1–31 March 2017) respectively	VMD
10	Woo et al. (2018)	Seongcheol Woo (2018)		MAPE, and RMSE	Conv-LSTM MTL hybrid model	Wind field and the corresponding two response time series data. In A total 22,000 pairs of data points	

Table 3 (continued)

SN	Paper	Author	Type of WP prediction frequency	Performance metric	Model used	Dataset used	Decomposition method (if any)
11	Wu et al. (2016)	Wenzu Wu (2016)	Short term	RMSE and MAE	CNN-LSTM, CNN-RNN, CNN-FF, FNN, and ARIMA	Year 2014 North East China wind farm, wind power and wind speed data recorded every 15 min	
12	Zhang et al. (2019)	Jinhua Zhang (2019)	Short term	RMSE	LSTM and GMM	7393 data points from a wind farm of 123 units containing wind speed were collected from one unit	
13	Xu and Xia (2018)	Gang Xu (2018)		accuracy and MAPE	adaptive LSTM model using an improved genetic algorithm	Daily 144 data points from May 2016 to May 2018 at intervals of 10 min	
14	Yu et al. (2019)	Ruiguo Yu(2019)		MSE	LSTM-EFG (LSTM-enhanced forget-gate). Clustering based on feature correlation	NREL data set of 32,043 wind turbines with time intervals of 10 min between 2004 and 2006	
15	Lu et al. (2018)	Kuan LU (2018)	Short term	RMSE	LSTM network-based encoder-decoder (ED) model	From January 1, 2016, to May 31st with an interval 15 min, a total of 14,496 observations	

Table 3 (continued)

SN	Paper	Author	Type of WP prediction frequency	Performance metric	Model used	Dataset used	Decomposition method (if any)
16	Zhou et al. (2019)	Bowen Zhou (2019)	Spot prediction	MAE, RMSE, MAPE	K-Means-LSTM and a non-parametric KDE model	From Dec. 2017 to Dec. 2018, sampled with a gap of 10 min interval	Clustering of abnormal data points using K means, DBSCAN and 3σ
17	Zhou et al. (2017)	Tushar Srivastava (2020)		MSE, RMSE, MAE, MAPE	RNN, GBM, and LSTM,	Hourly data of year 2014 of Kolkata region in India is considered	–
18	Duan et al. (2022)	Jiandong Duan (2022)	Short term	MAE, RMSE	VMD based LSTM and PSO-DBN	Dingbian, Shaanxi, 10 min.interval 2880 datapoints	VMD
19	Zhang et al. (2022)	W. Zhang (2022)	Short Term	Accuracy	DWT based SARIMA and LSTM	SCADA data of off-shore wind turbine in Scotland	DWT
20	Devi et al. (2020)	A. Shobana Devi (2020)	Short Term	MAPE, RMSE, MAE, MASE	EEMD based LSTM-EFG-CSO	NIWE, 15 min interval 24,192 data points	EEMD
21	Sun et al. (2022)	Yiyang Sun (2022)		MAPE, MAE	Modified PSO based A-LSTM	NREL, hourly data from 2007 and 2012	–
22	Ma and Mei (2022)	Zhengjing Ma (2022)		Accuracy	Attention based CNN-BiLSTM	Turkey wind farm data	–
23	Garg and Krishnamurthi (2022)	S Garg (2022)	Short Term	MSE, RMSE, Time	CNN ED LSTM	NREL, hourly data from 2007 and 2012	–
24	Garg and Krishnamurthi (2023)	S Garg (2023)	Short Term	MSE, RMSE, MAE, Time	Powernet	NREL, hourly data from 2007 and 2012	–

Table 4 Merits and Demerits of LSTM and associated models used in wind power predictive analytics to date

SN	Paper	Author	Merits	Demerits
1	Zhou et al. (2020)	Baobin Zhou (2020)	SARIMA helps in additional seasonal information extraction and in turn adds on to prediction capability of LSTM. MAPE is improved using proposed method by 10.31% due to periodicity improvement added to LSTM using SARIMA, redundancy removal using PCA and decomposition using EEMD	• Quite slower than PCA-EEMD-LSTM, • it takes approximate 3 h to train the models • shows efficient results only in 15 min interval predication
2	Duan et al. (2020)	Jiandong Duan (2020)	IVMD-SE data pre-processing makes LSTM insensitive against outliers and noise by improving RMSE from 169.64 in LSTM to 58.77 in proposed model	Performance improvement is at the stake of higher computational time from 1439.31 units in LSTM to 3205.72 in proposed one
3	Srivastava et al. (2020)	Tushar Srivastava (2020)	Application of 2 models namely Basic RNN and GBM is well contrasted with LSTM model, over a common predictions using wind velocity as input and wind power as output. LSTM is proven best with MAE of 0.176683 as compared to the worst performing GBM with MAE 0.320215	Improvement in the Neural system's accuracy by using various other hybrid techniques and these models and datasets is limited to a single city
4	Han et al. (2019)	Li Han (2019)	Data decomposition on the basis of different characteristics when combined with memory component of LSTM is generating better performance when tested for multistep as well as real time predictions. VMD_LSTM reduces error to 0.019 as compared to LSTM with error of 0.024. Also it has MAPE as less as 0.014 as compared to highest of 0.132 of BP	Time taken increases to as high as 17.33 s as compared to lowest in LSTM 10.56 s. excessive layers of decomposition add to the computational time. Hence excessive decomposition might lead to adverse results
5	Liu et al. (2020)	Bingchun Liu (2020)	MAPE of WD-LSTM is 5.831 as compared to LSTM having 13.715, performing better than other models because of data optimization by WD, which improves the prediction accuracy	LSTM takes 32 min to converge as compared to 44 min taken by WD-LSTM But proposed model still needs to verify whether the time series with different scales can be used as the same model's input index

Table 4 (continued)

SN	Paper	Author	Merits	Demerits
6	Sun et al. (2019)	Zexian Sun (2019)	VMD-KMeans-LSTM improves performance in terms of MRE as high as 66.67% as compared to LSTM. Hence, decomposition of data on different scales and clustering using K means improves the predictive measures as compared to standalone LSTM	
7	Liu et al. (2019)	Y Liu (2019)	DWT-LSTM when analysed over 3 different wind farms performed with an MAE of 10.12 against LSTM having 28.31 and this performance is consistent for all the 3 farms. Also, amongst other models LSTM,RNN, and BP, LSTM outperformed others for all wind farms	It still needs to optimize parameters like weight, learning rate, etc., to achieve higher accuracy
8	Choi and Lee (2018)	Jae Young Choi (2018)	Ensemble LSTM improves over basic LSTM by as high as 46% over MSE and as low as 21% tested over 4 different travel time series datasets	
9	Shi et al. (2018)	Xiaoyu Shi (2018)	MAE is improved for both R-LSTM and D-LSTM. MAE for R-LSTM is improved to 0.28 over LSTM MAE 2.81. MAE for D-LSTM is reduced to 0.54 over LSTM with 5.04	The model is quite limited to Wind power forecasting application. The logic of Recursion and direct is strictly adhered to application in hand
10	Woo et al. (2018)	Seongcheol Woo (2018)	It can take the Spatio-temporal characteristics wind flow input, multi-tasks model reduce the complexity of model. The MAPE for power prediction is reduced to 0.841	Indirect wind power prediction via wind speed. a direct application should be tested for accuracy and efficiency
11	Wu et al. (2016)	Wenzu Wu (2016)	Works better than different hybrid methods and ARIMA with optimized cost. CNN-LSTM reduces the MAE to 5.62 as compared to 6.65 in FNN, 9.72 in ARIMA, and 12.81 in persistence	Tested only for point forecasting

Table 4 (continued)

SN	Paper	Author	Merits	Demerits
12	Zhang et al. (2019)	Jinhua Zhang (2019)	LSTM model when compared with other DL models performed best with the RMSE improved by 6.37% when applied on 16 units. GMM describes the uncertainty and confidence intervals	LSTM model when compared with other DL models, took highest time that is 540 s. Classification of data and then predicting its uncertainty might improve on the forecasting
13	Xu and Xia (2018)	Gang Xu (2018)	Adaptive LSTM when compared to BP, LR, SVR, BR improved the MAPE to 2.7 in contrast to the worst performer LR with MAPE approx. 19%	
14	Yu et al. (2019)	Ruiguo Yu(2019)	LSTM-EFG reduces the error to 6.3648 as compared to 6.9322 in LSTM tested over same environment. Forget gate when enhanced using either spectral or agglomerative clustering perform optimally	At certain times prediction does not give results as expected. The consistency of LSTM-EFG with either of the enhancements is missing
15	Lu et al. (2018)	Kuan LU (2018)	The RMSE of the LSTM E-D was 2.6%,5.2%, 8.5% and 11.8%, lower than LSTM	Dataset can be more wide and robust for testing accuracy
16	Zhou et al. (2019)	Bowen Zhou (2019)	K LSTM and LSTM performs better by reducing prediction error to 33.445 and 49.559 as compared to BP, elman and SVR where worst is 58.446 of BP	Selecting some unreasonable value for K influence and deteriorate the results and hence out of 2200 total sampled 2000 samples are required first just to predict the value of K
17	Zhou et al. (2019)	Tushar Srivastava (2020)	LSTM performs better than the RNN and GBM in wind power forecasting with MSE of 0.0078 as compared to 0.0080 of RNN and 0.1623 of GBM	Moreover, LSTM needs to be compared with more hybrid models and their variants to see if it outperforms them or not
18	Duan et al. (2022)	Jiandong Duan (2022)	MAE and RMSE for the proposed LSTM based model is reduced upto 210% against state of art DL models and upto 95% against other decomposition techniques	Time taken by proposed model is highest of all the models, decomposition methods used for comparison i.e. 2254 s
19	Devi et al. (2020)	A. Shobana Devi (2020)	Error metrics are improved when compared against ARMA, BPNN, LSTM, LSTM-EEG, EEMD-CSO-SVM, EEMD-GA-BPNN, EEMD-LSTM	Increased i.e. (2–3 years) in training set in showing improvements in accuracy as compared to lesser i.e. (1 year) one
20	Sun et al. (2022)	Yiyang Sun (2022)	Quantitative improvement of MAPE and MAE against LSTM, A-LSTM and PSO-A-LSTM	Data cleaning used is an extra step when using DL models with attention mechanism

Table 4 (continued)

SN	Paper	Author	Merits	Demerits
21	Ma and Mei (2022)	Zhengjing Ma (2022)	A temporal embedding layer, Time2Vec replaces decomposition overhead Attention based CNN-stacked BiLSTM improves accuracy than LSTM and GRU	Comparative analysis should be done with more relevant models
22	Garg and Krishnamurthi (2022)	S Garg (2022)	CNN-ED-LSTM is 9.6% better than VanillaLSTM, 3.2% better than StackedLSTM, 5.3% better than CNN-LSTM and 7.8% better than Bi-LSTM	Time is a trade-off with error and accuracy
23	Garg and Krishnamurthi (2023)	S Garg (2023)	Powernet is working better than 5 DL models VanillaLSTM, StackedLSTM, BiLSTM, CNN-LSTM, and CNN ED LSTM, and five conventional statistical models, namely, AR, ARMA, ARIMA, and ARIMA-LSTM	Time taken by proposed model is highest of all

datasets and different models applied over these different datasets, (iii) taxonomy of WP predictive analytics, elaborating different state of art classifications, evolution of this branch, decomposition methods proposed and implemented, and (iv) models with LSTM as whole or part applied and assessed over WP time series dataset. This part of the paper has a wide assessment of WP predictive analytics via different LSTM variants and their combinations. Section 3 discusses challenges faced in WP time series predictive analytics in a categorical fashion. It also lists the challenges associated with LSTM in WP analytics. Section 4 finally concludes the different aspects of this study.

2 Literature

2.1 Time series predictive analytics

Time series data has led to a wide diversified set of applications solving different set of problems as referred in Table 1. State of art represents different dimensions where time series predictive analytics engender awareness and strengthen with the meaningful insights.

2.1.1 Teaching learning industry

In Yang et al. (2017), the predictive analysis is applied over MOOC online courses. Two subjects' Dataset is taken from the past student assessment data and video watching trend. The major issues faced in the data collection were (i) per-student assessment response sparsity and (ii) individuality in every student's responses. Predicting struggling students before starting the course gives a fair idea to the course instructor before starting the course. This instructor can plan the course proceedings accordingly and make the best interest class concerning the students enrolled. Also, students will learn maximum out of the course, benefitting MOOC to grow as a business and teaching–learning industry.

2.1.2 Health industry

Other predictive analytics on time series data helped predict future 14 days situations in COVID cases. Javid et al. (2020) worked on COVID-19 data and in May 2020 proposed four different models for predicting upcoming days' number of cases. The training was done on Johns Hopkins University COVID-19 dataset having 12 different countries. That was the perfect timing to help the government make the best decisions in all the odds of the COVID situation.

In health spectrum another study Redondo-Bravo et al. (2019) is to show immigration and travel history impacts on increase in dengue specially in those travelled areas. Also, predictive analysis to estimate dengue incidents till 2025 is done. Dengue reported and discharged cases of hospitals from 1997 to 2016 are used for this study. The observation seen says a direct relation between travelled and pandemic affected areas. Due to critical increase estimation of 65% by 2025 risk mitigation measures are proposed. Awareness to general public for preventive measures is another prescription made.

2.1.3 Business intelligence and analysis (BI&A) industry

Different business models need prescriptive and responsive action plans for growth. To identify key performance indicators (KPIs), use these KPIs for formulating different strategies and segregate to include the top indicators suitable for a respective business model are some of the critical issues in BI&A. Wang et al. (2018) presented a framework consisting of RF, BBN, and tested on three product companies' business models.

2.1.4 Production industries

Industrial challenges like (i) fault detection, (ii) fault monitoring, (iii) quality monitoring, and (iv) quality control are few of those areas which have gained attention of predictive analytics. In Filonov et al. (2016), LSTM neural network helped in fault detection and specifically the areas where fault has occurred, root cause and the reasons behind it. In a similar study Mehdiyev et al. (2017) different challenges associated with process industry are researched. Also post processing activities like quality check and control are studied. Predictive analytics in this time series sensors data is the need of industrial processes and business related to it.

2.1.5 Weather forecasting and disaster management

A significant use case of time series predictive analytics is weather forecasting. Likewise, Mustaffa et al. (2018) searches for the best model to predict weather out of five previously proposed models independently. The same benefit of this predicted close to "completely accurate" makes daily life planning easier in many areas. One of the most hazardous and disastrous natural calamities towards human lives and economic wealth is Tropical cyclones. Though it is inevitable but correct predictions about its happening, intensity, and track can prepare disaster analysis, mitigation, and management quite effectively and, in turn, save many lives living near those areas (Geetha 2016).

2.1.6 Wastewater flow

Predictive analytics in wastewater treatment plants (WWTPs) helps in management and treatment of waste water. But the less and inconsistent data is a prominent challenge in predictive analysis of WWTPs. Another challenge is age old infrastructure in place which makes data collection a difficult task. Addressing all the challenges and using decision support system for performing predictive analytics over WWTPs (Zhang et al. 2019) proposes an optimized time series modelling technique.

2.1.7 Sales

Sales can be considered as a time series. Estimating correct volumes and trends of sales can help business to stock up accordingly and target maximum profits. Pavlyshenko (2019) is one such study that aims at improving accuracy of prediction in sales data.

2.1.8 Demand predictions

Demand in terms of load demand, water demand, inventory demand, gasoline demand,, ATM cash demand and so on are some of areas benefitted from time series predictive analytics. One such study Sarveswararao et al. (2023) works in forecasting ATM cash withdrawal at a particular point in time. ATM cash management can be planned according to the withdrawal predictions made.

2.1.9 Pollution and air quality index (AQI)

Predicting air pollution and AQI for planning various activities is another area of time series predictive analytics. Deep et al. (2022) is a prediction research to plan about knowing air pollution at specific locations of urban cities in advance to plan a less polluted route.

2.1.10 Rainfall

Due to inadequate rainfall Turkey is drought prone which leads to water scarcity and affect agricultural industry. So, accurate prediction of rainfall in Chen et al. (2022) over the concerned region is helpful to be prepared for the measures well in advance.

2.1.11 Bitcoin

Economic growth and trading help is developed by forecasting crypto currency price to invest with a calculated risk. Aditya Pai et al. (2022) forecasts price of Bitcoin, Ethereum, Litecoin and Bitcoin-cash prices using LSTM neural network.

2.1.12 Agricultural industry

Agriculture practices are highly dependent on meteorological parameters like solar activity, rainfall and moisture level, wind direction and speed and so on. Such a predictive analysis in Rogachev (2022) studies solar and hydro-meteorological conditions of Lower Volga region of Russia for the planning and management of cultivated crops.

2.1.13 Stock market

Using three stocks data and neural network predictive modelling (Park et al. 2022) predicts the next day's stock return and return direction using 43 technical indicators with 50 time-steps to increase the profit in investments. This type of predictive analytics help the trading industry and money back assurance upto some extent.

2.1.14 Travel and tourism industry

Due to pandemic COVID-19, tourism was bearing losses. To help recover the tourism industry especially international and increase employment and incomes of travel and tourism industry has to make certain extra efforts. Predictions for monthly number of international tourists to the country are done in Nguyen et al. (2021).

2.1.15 Electricity demand and consumption

Better grid and resources usage can be implemented if energy demand and consumption are estimated beforehand. Using 500 household consumption data in Lazzari et al. (2022), electricity consumed is predicted and used for better electricity distribution.

2.1.16 Sustainable energy industry

Wind, solar, hydro, thermal are the major sustainable energy generation resources. Future predictions made in energy forecasting may lead to better grid usage and planning.Essien and Giannetti (2019) performs single step ahead predictions using three open-source univariate time series datasets. Similarly Zhang et al. (2021) proposed a Multivariate deep learning (DL) algorithm and experimented to evaluate the best model for more accurate predictions.

2.1.17 Electric vehicle demand and consumption

For Dundee, UK, EV station using variables data charger identifiers, start and end times of charging sessions, amount of energy charged, the power of chargers, and the geographical locations of chargers from March 5, 2018, to June 4, 2018 (91 days) in Ma and Faye (2022). To implement smart charging stations to EV users and operators, predictive analytics is required.

2.1.18 Cyber security

A study in Abdullahi et al. (2022) reviews cyber security predictive analytics in IoT from 2016 to 2021 and (Nagaraj 2022) uses 12 years of cyber threats data to understand the evolving cyber-attacks and their trend.

2.1.19 Crime

As the crime rate is increasing data by day, if the rate and time of crime scenes and locations is predicted before time based on previous happenings using predictive tools the stricter and planned vigilance can be installed. Hence, forecasting crime rate can set up timely suitable actions to reduce the crime rate as done in Muthamizharasan and Ponnusamy (2022).

2.1.20 Price prediction

Price prediction in different areas like share, currency, electricity, metal, product, property, vehicle, crop etc. can help in plan the production and stocking. Like Livieris et al. (2020) implements predictive modelling to predict gold price has a significant impact on financial activities. This could offer insights in gold price fluctuations, behaviour and dynamics and ultimately could provide the opportunity of gaining significant profits.

Connected survey is about the evolution of predictive models, different performance metrics used and assessment of the models over these performance metrics. Table 2 describes time series predictive models used since 2016 and their respective performance.

2.2 Generic models and datasets used in time series prediction

Clearly the above listed applications and areas reflect the effect of different data types using time-series predictive analytics applied over it. However, this does not survey the variety of models studied on different datasets to make these most benefit and profit predictions. Hence, Table 2 surveys different models and performance metrics used for the time series type of data and are explained in this section.

Table 2 lists surveys (Do et al. 2019; Qin 2019) comparing MLFFN, ARIMA, TDNN, RBFN, DBN, basic RNN, CNN, GRU, Temporal CNN, Bidirectional RNN, SVR. Also, explores the work done on time series data using different models and comparing the performance metrics amongst different models. The predictive parameters show that LSTM and GRU give efficient results as compared to the six models compared. Bao (2017) Works on a combination of wavelet transform stacked AEs and LSTM. The MAPE comes down to 0.066 when applied on stock market daily data considering six stock indices. Also, (Zhu and Laptev 2017) proposed ED with two-layer LSTM of 128 and 32 hidden states. It is connected with three fully connected dense layers taking tanh as the activation function and 128, 64, and 16 as hidden units, respectively. This hybrid model of LSTM improves MAPE to 5.9 from 9.2 calculated from LSTM. Likewise, (Cui et al. 2018) proposed a deep-stacked bidirectional and unidirectional LSTM NN on traffic dataset collected at 5 min interval, is proved efficient by comparing with SVM, RF, FFNN, GRU-NN with performance metric MAE. Also, normal NN and PSF is compared in Torres et al. (2019) using solar power data over the metric RMSE and MAE with deep NN with 1–5 hidden layers and 10–40 no. of neurons each and proven better with the results.

Using the Stock market dataset, LSTM, SVR, MLP, and RNN are measured over trade gains metric (Akita 2016). The maximum gains are happening when LSTM is applied to the data, and that too only in 50 epochs. Zeroual et al. (2020) evaluates RNN, LSTM, BiLSTM, and VAE over COVID 19 data using RMSE. VAE performed best to performance metrics when compared to the other four models. Once again, (Hua et al. 2018) uses traffic data sampled every 15 min and applies SVR, ARIMA, FFNN, LSTM30 (with 30 layers), RCLSTM, and LSTM300 (with 300 layers). LSTM300 outperformed all other applied on RMSE metrics. Vicente et al. (2017) Uses a three-layered DL architectures with 10–100 neurons to measure electricity load data. The next 24 h predictions are made with just a 2% error. In Li et al. (2019) genetic algorithm combined with A-LSTM is applied and compared with many others over PM2.5 and SML2010 datasets. An efficient MAE proves this evolutionary A-LSTM better and accurate. Refs. (Qian 2017; Kim and Kang 2019) are working on financial data to compare and propose A-LSTM respectively. Weights added to A-LSTM are even more efficient (Kim and Kang 2019). Another model combining multiple CNN and Bi LSTM on stock market data proves it to be the efficient one when applied to this timed data in Eapen et al. (2019). Xiao et al. (2019) uses LSTM and Adaboost ensemble learning model over sea surface temperature prediction time series data. Solar-assisted water heating data is analysed using Attention-based LSTM and data decomposition in Heidari and Khovalyg (2020). Health management data is worked upon using multivariate CNN for multivariate time series classification in Liu et al. (2019).

Different time series datasets used under this survey include stock market, dissolved oxygen at a water monitoring station, daily completed trips, traffic data, COVID-19 dataset, sea surface temperature, financial data, solar-assisted water heating, health management data. Quantitative performance metrics used for comparison are MAPE, MSE, MAE, RMSE. These two table listings give a thorough and broad spectrum of a few observations: (i) the requirement of predictive analytics in the first place, (ii) different types of time series datasets available out there (iii) different performance metrics used to measure the efficiency of predictions (iv) different variety of models and their hybridization. (v) role of LSTM and its associated models in time series predictive analytics. However, wind energy and specifically models for efficient predictions for WP generation are still to be studied. Hence, Table 3 and Table 4 give a detailed and thorough review of WP predictive analysis done to date using LSTM as a whole or part of the model.

2.3 WP predictive analytics

WP predictive analytics is to predict future power of a wind farm or wind turbine to plan energy generation against energy demand. To stabilize grid system WP predictive analytics is an important requirement and effective way. WP is directly dependent on meteorological variables like temperature, pressure, wind speed, wind direction, landscape and most importantly the time. WP, pow, is defined as equation mention in Eq. (1), with ρ as air density, A is swept area of wind and V as wind speed. This establishes a non-linear cubic relationship between WP and wind speed (Wang et al. 2020).

$$\mathrm{pow} = \frac{1}{2}\rho A V^{3} \tag{1}$$

Different approached based on these parameters are proposed in literature for improved prediction accuracy. A wide amount of historical data helps data driven approaches make better power prediction. This section elaborates taxonomy of WP predictive analytics based on different factors. First is based on time frequency, second is variable based and the last one is predicative approached based classification.

2.3.1 Taxonomy

2.3.1.1 Types of term based forecasting Since WE data is a sequential series in the progressive order of time, forecasting is generally done based on time variation namely, (i) ultra-short term forecasting, (ii) short term forecasting, (iii) medium term forecasting and (iv) long term forecasting (Chang 2014; Nowotarski and Weron 2018) as shown in Fig. 4.

Ultra short time forecasting is done From few minutes to an hour ahead (Chang 2014). In WP prediction the ultra-short term, is considered with an interval resolution from 15 min. Applications where intra domain synchronization is required, ultra short term forecasting is useful. Real time power grid operations, electricity market clearing, regulatory actions are few of the major areas of ultra-short term forecasting. Moreover, the predictive analytics in ultra-short term requires more precision since it has shorter time intervals. That makes it difficult as compared with other types of WP prediction (Xiang et al. 2022).

Short term forecasting has a temporal resolution from 1 h onwards till a few hours (Chang 2014; Huang et al. 2022). Majority of the predictive analytics like production industry planning, quality control, finance estimations, sales forecasting are made using short term temporal gaps. Electricity demand and supply, planning and wind energy forecasting are prominent applications under short term WP forecasting (Maldonado-Correa et al. 2021).

Medium-term forecasting is between few hours till a week's time interval (Chang 2014). Weekly strategies or low impact strategic decisions are generally made with medium-term forecasts. In power industry unit commitment decisions, online or offline generator decisions, reserve requirement decisions are few areas where medium term forecasting is applicable and informative.

Long-term forecasting is between one week to more weeks, even for years' intervals (Hossain et al. 2018; Chang 2014). Long term plans, quarterly target prior setting, resource planning, project planning budget plans etc. are some critical forecasts to be made using long-term forecasts. Operational activities like maintenance planning, optimal operational cost, wind farm installation feasibility study are major areas where long term forecasting is feasible.

Depending on the type of forecasting models their performance varies. Also, depending on stationarity, linearity, noise, and various other data features, the predictive model's choice varies.

2.3.1.2 Types of variable based forecasting Time series predictions are classified into univariate and multivariate type of predictive analytics based on single attribute and multiple attributes in consideration (Md Azmi et al. 2022). Different state of art studies have worked and analyzed univariate and multivariate forecasting methods using various predictive models. Some of the studies have also done comparative analysis on these two approaches.

Univariate time series forecasting: In one of the studies (Md Azmi et al. 2022) univariate and multivariate short-term forecasting with a small wind turbine data is assessed and compared. CNN, out of all the evaluated models, is performing best for univariate forecasting. However, gradient boosting and huber regresser are working better for multivariate analysis. Comparative analysis is done using metrics MAE, RMSE, MAPE and R^2. Essien and Giannetti (2019) proposed a WT and bidirectional Convolutional LSTM based stacked AE to have single step ahead predictions. Three open-source univariate time series datasets are used to compare the proposed AE based model with 2 state of art DL models. Another study Akbal and Ünlü (2022) proposes a short to mid-term WP forecasting in Manisa, Turkey using sequence based encoder decoder architecture. Decoder is comprised of stacking of LSTM layers over first dense layer as input layer. Comparative study in Chen et al. (2021) shows machine learning (ML) algorithms work better than persistence on univariate WP forecasting. Also, this deduced that same advanced approaches are working efficiently when applied over multivariate WP forecasting. This was experimented using five wind farms data. Similarly, LSTM is proposed in Xu et al. (2021) to predict ultra-short-term WP, after improved gray correlation analysis method selects a similar day. SD-LSTM, a deterministic prediction model and Bootstrap-Kernel density, an interval prediction model of ultra- short-term WP are proposed in the study. The conclusion is that the models work better for multivariate forecasting when applied long term prediction, however, work better for univariate forecasting when applied for short term predictions. Hu et al. (2020) used a novel convolution-based spatial–temporal highly non-linear and deep architecture for very short term WP predictions. Twenty-eight wind farms as case studies are conducted to show superior performance of proposed model on 5–30 min ahead forecasting than univariate models such as AR, DBN, and ANN. In Alkesaiberi et al. (2022) ML models like GPR, SVR with different kernels, and ES models (Boosted trees and Bagged trees) are investigated on univariate inputs wind speed and WP. BO as hyperparameter tuning algorithm incorporates dynamic behaviour in the models to improve the performance. Three wind farms are used to perform experimentation and assessed using RMSE, MAE, and R2. The results showed that the optimized GPR and ensemble models outperformed the other ML models. Likewise multivariate predictive analytics has been studied, experimented and evaluated using different approaches proposed in literature.

Multivariate time series forecasting: Different approaches in state of art have worked on multivariate time series and specifically WP predictive analytics. To overcome the problem of handling dynamic changes in VAR model, (Messner and Pinson 2019) proposed a time-adaptive lasso estimator and an efficient coordinate descent algorithm for updating the VAR parameters recursively online. This proposed approach works better in the multivariate dynamics and univariate autoregression when compared with non-adaptive Lasso VAR. In a similar study Mishra et al. (2020) performance comparison of five DL models combined with three types of data pre-processing and used for short and long-term multivariate, MIMO architecture, prediction. Five models namely, DFF, Deep CNN, RNN, Attention mechanism and LSTM are assessed with and without DWT and FFT. Attention and DCN perform best with data pre-processing, however, other models perform better when no Wavelet or FFT signal pre-processing is done. Likewise, Sørensen et al. (2023) reviews state of art methods for multivariate wind and solar power forecasts. Linear, adaptive, conditional parametric and combined forecasting are the categories in which literature approaches are categorized. Computational time and accuracy are two trade-offs in which the applications are divided and selection of the apt approach is dependent. Zhang et al. (2021) proposed a Multivariate EMD and attention based CNN and BiLSTM, DL algorithm. Thereafter, 3 experiments were conducted to evaluate the proposed model and indicate that the proposed one is better than other baseline models. Another study Meng et al. (2022) addresses the lesser addressed issue "pre-diction stability". A multi-objective crisscross optimization (MOCSO) based model is proposed to improve prediction stability. Multivariate VMD combined with deep ELM model outperforms three multi-objective state of art optimization algorithms. Du et al. (2020) again used BiLSTM layers combined with temporal attention in the form of encoder decoder for multivariate forecasting. Experiments are performed on 5Beijing PM2.5 multivariate datasets. Comparison is done with baseline models namely, SVR (linear and RBF), ARIMA, VARMA, RNN, CNN, LSTM, GRU, SEQ2SEQ and its two variants SEQ2SEQ-BI and SEQ2SEQ-ATT. Metrics used for comparison are MAE and RMSE. Liu et al. (2019) is another study to boot usage of CNN in time series multivariate classification. The model proposed is named Multivariate CNN (MVCNN). Comparisons are made against Vanilla CNN, Vanilla NN, XGBOOST, RF and LR. Score, precision and sensitivity are quantitative measures for evaluation. Also, an ultra-short-term WP predictive analytics is proposed in Liu et al. (2018) using multivariate PSR and multivariate LR. Moreover, the time series forecasting has another taxonomy based on the interval of prediction. The time-ahead prediction can vary from a few minutes to few days, months and even years. Based on this criterion the types are termed as forecasting terms.

2.3.1.3 Models: conventional to the recent evolution The journey of WE predictive analytics from conventional to current learning methods are broadly classified as statistical and learning techniques, as mentioned in Fig. 4. For all forecasting types, whether short-term or long-term, prediction approaches proposed and studied to forecast WE are divided into six broad categories namely (i) persistence, (ii) physical (iii) statistical, (iv) Artificial Intelligence (AI), (v) DL, and (vi) Hybrid (Chang 2014; Zhao et al. 2011).

First is the **persistence method**, one of the simplest methods proven to be quite efficient for ultra-short interval forecasting (Chang 2014). So at any time t if WP is defined as pow(t), then at next timestamp t + i also WP will be generated pow(t) as shown in Eq. (2), which is measured now independent of weather conditions (Chang 2014; Zhao et al. 2011). In ultra-short term forecasting persistence method works accurately. Also, it is simplest and computationally economical of all the other predictive methods. However, the efficiency decreases when the term scale of the forecasting increases.

$$\mathrm{pow}(t + i) = \mathrm{pow}(t) \tag{2}$$

Next is the **physical approach** (Manwell et al. 2010; Artipoli and Durante 2014), in which current weather conditions from weather forecast data called NWP having parameters like temperature, obstacles, pressure are fed into a prediction model to calculate WE. However, again it is adequate for short-term forecasting intervals (Zhang et al. 2016; He et al. 2022). Although accuracy can be achieved in long-term interval forecasting using physical methods if the weather condition is constant (Du et al. 2019). Current commercial WP forecasting methods use NWP wind forecasts as the input data (Liu et al. 2022).

Next are **statistical methods**, which take the history of parameters to forecast future WE (Zhang et al. 2016; Wang et al. 2011; Masseran 2015). They are cost-effective and efficient methods for short forecast horizons (Study et al. 2017). However, accuracy reduces with the increase in time frequency. Statistical methods include the AR, ARMA (Magadum et al. 2023), ARIMA, Seasonal ARIMA (Zhang et al. 2022), Bayesian approach, and gray predictions (Chang 2014). Various improvements are proposed in literature for statistical models. Like for AR (Zhang et al. 2021) proposes an improved AR named ARDA with dynamic adaption added in AR. The experiments prove ARMA better than AR and ARIMA in terms of accuracy and time. ARMA has different improved variants in Liu and Che (2019), Wu et al. (2023). In another study Singh et al. (2019) ARIMA is combined with ANN and (Kim and Hur 2020) proposes ARIMA with exogenous variable ARIMAX to improve WP forecasting.

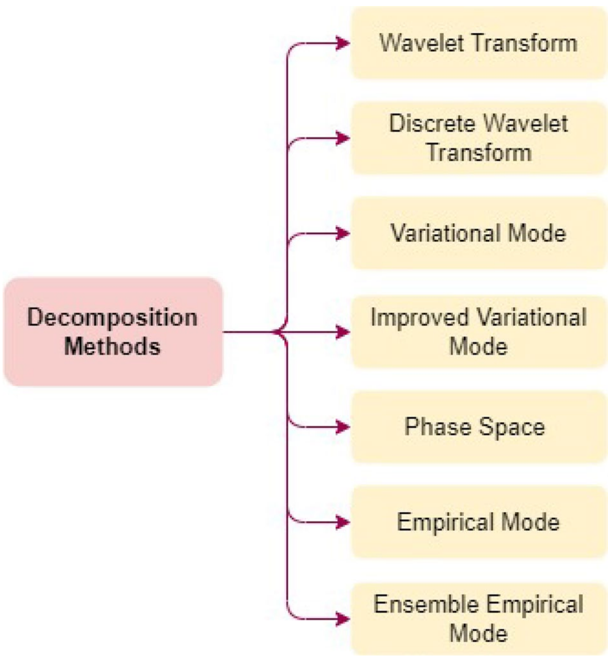
Significant ones are data driven **artificial intelligence (AI)** models (Chang 2014; Shamshirband et al. 2019) that imitate the human brain and predict the energy against the well-trained parameters. Different AI models are proposed and classified into three branches (Wang et al. 2020), (i) Fuzzy system, rule based (Akhtar et al. 2021) (ii) SVM, vector based prediction on samples (Jiang et al. 2020) and (iii) SOA, optimization techniques (Wang et al. 2020).

In WP predictive analytics SVM and LSSVM in different tuned forms, combined with other methods (Mustaffa et al. 2018), decomposition techniques, pre-processing methods, data mining approaches and DL models (Han et al. 2017) are implemented for improved forecasting.

The last ones named **hybrid models** (Hossain et al. 2018; Osório et al. 2018; Du et al. 2019) are taking best out of a particular model and combined with some other model's best. For example, combining any two types of modelling techniques like, physical and artificial intelligence, statistical and artificial intelligence approaches etc. generate hybridization of models. Combination of various artificial intelligence models are proposed under the hybrid category (Study et al. 2017). ANN-ARIMA, SVM-ARIMA (Chang 2014) are real-time models combining AI and statistical models and generating a hybrid model. AI models are more suitable for the non-linear type of data and create training parameters by developing a discreet relationship with historical time series. Combining multiple methods or multiple models of a single way or multiple interval models generates better accuracy than a single model (Chang 2014). From conventional predictive analytics, recently DL models evolved as optimized models for predictive analytics.

Deep Learning and associated hybrid models: The DL technique is considered a class of ML or AI techniques. To handle this data, effective DL techniques include RNN, WNN, Recurrent WNN, DBN, CNN, LSTM, DNN, RMT dependence and single-active electron approximation (Shamshirband et al. 2019). DL methods run deep with several hidden layers abstracted compared to the shallow behaviour of ANN (Vaitheeswaran 2019). LSTM neural networks (Greff et al. 2017) have faster gradient convergence and the power to remember relevant historical data even in long-term intervals due to its structured gated architecture (Yuan et al. 2019; Qing and Niu 2018). Moreover, the forget gates in LSTM keep only relevant data from the past, required in the future times (Greff et al. 2017; Westhuizen and Lasenby 2018). Since WP time series varies similar to LSTM, LSTM has already been applied in many research pieces for WP prediction (Liu et al. 2018; Hu and Chen 2018). Recent researches mainly focused on hybrid forecast models with LSTM, such as the forecasting model based on LSTM network and some other DL neural network (Chen et al. 2019; Santhosh et al. 2020; Qin et al. 2019), the multi-task convolutional LSTM model (Woo et al. 2018), the intelligent multi-step DL model (Liu et al. 2018; Han et al. 2019; Li et al. 2018), hysteretic ELM (Hu and Chen 2018) and differential evolution algorithm (Hossain et al. 2018; Hu and Chen 2018). DL when combined with methods, like pre-processing of data using decomposition methods, combination of DL model with other predictive analytics method, combining multiple DL methods to generate a hybrid approach, are proven to be efficient and optimized.

Fig. 5 Pre-processing data decomposition methods



2.3.2 Decomposition and pre-processing of dataset

Additionally, the pre-processing and optimized performance, are essential requirements of WP predicting. For this purpose various decomposition methods are applied to improve the WP's prediction accuracy (Qian et al. 2019). WP signal is analysed using seven methods as shown in Fig. 5, namely, (i) wavelet decomposition (WD) (Li et al. 2018; Bhaskar and Singh 2012; Liu et al. 2019), (ii) phase space decomposition (Han et al. 2015), (iii) empirical mode decomposition (EMD) (Naik et al. 2018; Guo et al. 2012), (iv) ensemble EMD (EEMD) (Bokde et al. 2018), (v) variation mode decomposition (VMD) (Fu et al. 2019; Liu et al. 2018; Han et al. 2019; Vanitha et al. 2020; Abdoos 2016; Zhang 2019), (vi) discrete wavelet transform (DWT) (Liu et al. 2019), (vii) improved variation mode decomposition (IVMD) (Duan et al. 2020). However, to select an optimized decomposition method like wavelet basis function in the wavelet method is a challenge and hence, is another problem to study altogether. Also, decomposition is not always possible considering the time factor in wind forecast data. The major challenges in phase space decomposition technology are identified as delay time and embedding dimension. A significant historical experience is required to use decomposition techniques EMD and EEMD. Recently study the decomposition of WP signal by utilizing the VMD (Fu et al. 2019; Liu et al. 2018; Han et al. 2019; Vanitha et al. 2020; Abdoos 2016; Zhang 2019) method is done along with neural network and LSTM as well. However, the number of decomposition layers, optimal functions, and the degree of decomposition are still critical to determine.

2.3.3 LSTM and associated models for wind power prediction till date

WP predictive analytics creates fluctuations in the resultant predictions due to randomness and uncertainty involved in the features. The day-ahead or n timestamps-ahead predictions in WP depend on current input as well as previous timestamps. The dependency may go longer depending on current timestamp and weather conditions. LSTM works by maintaining two memory components: short term from the previous state and long term from the cell state. Also, its architecture consists of three gates: input gate, output gate and forget gate. Input gate selectively selects the input to read, forget gate discard the irrelevant information and carry forwards the important timestamps in the cell state and write gate selectively writes the important information to the next state. Hence, in time series predictive analytics LSTM architecture performs better for remembering temporal information and make accurate predictions.

In literature along with standalone LSTM (Wu et al. 2016), hybrid models with LSTM like ARIMA, K Means clustering, Ensemble Network, Convolutional MTL, GMM, genetic and evolutionary optimization algorithms for adaptive learning, are proposed as shown in Fig. 6. Moreover, variants in LSTM by taking an enhanced forget gate or using LSTM in the encoder-decoder model are also some of the proposed approaches in predictive analysis of WP. Additionally, as shown in Fig. 5, decomposition of data as a pre-processing step using VMD, IVMD, WD, DWT are add-ons to improve predictions and analyse the claim that LSTM has properties to work in line with time-series data and help in predicting efficiently and effectively. A detailed and deep analysis of LSTM and its associated models, the application in the study, respective dataset, model specification, quantitative metric, merits and demerits are described in Tables 3 and 4. Exhaustive survey of LSTM as a base model in WP predictive analytics is followed by their tabular representation.

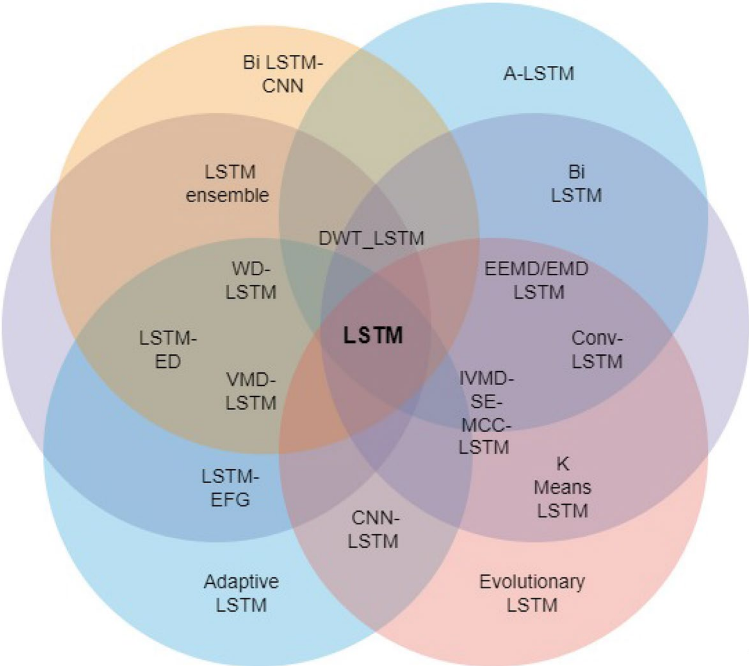


Fig. 6 LSTM state of art

2.3.3.1 LSTM over other standalone time series models Application of two models, over a span of one year 1st January to 31st December of 2014, hourly data of Kolkata wind turbine, namely basic RNN and GBM is well contrasted with LSTM model, over a common predictive environment using wind velocity as input and WP as output in Srivastava et al. (2020). LSTM is proven best with MAE of 0.176683 as compared to the worst performing GBM with MAE 0.320215. Other performance metrics like MAPE, MSE and RMSE are also compared and LSTM outperforms both RNN and GBM. However, no hybrid model is experimented along with the mentioned standalone ones. Also, the dataset is limited to a single city and a single wind farm.

Likewise in Zhang et al. (2019), 7393 data points from a wind farm of 123 units containing wind speed were collected from one unit. LSTM model when compared with other DL models namely RBF, wavelet, DBN, BPNN, and ENN performed best with the RMSE improved by 6.37% when applied on 16 units.

2.3.3.2 LSTM with VMD Two years of ELIA data is from 2015 to 2017 and NREL data is from 2011 to 2012 at an interval of 15 min is taken in Han et al. (2019). Data decomposition on the basis of different characteristics when combined with memory component of LSTM is generating better performance when tested for multistep real time predictions. Also, VMD_LSTM reduces error to 0.019 as compared to LSTM with error of 0.024. Additionally, it has MAPE as less as 0.014 as compared to highest of 0.132 of BP. However, time taken increases to as high as 17.33 s as compared to lowest in LSTM 10.56 s. Excessive layers of decomposition add to the computational time and lead to adverse results.

In another technique (Sun et al. 2019), VMD decomposes the raw WP series from two wind farms taken from China for the last quarter of 2012 and the first quarter of 2013. VMD-KMeans-LSTM improves performance in terms of MRE as high as 66.67% as compared to LSTM. Hence, decomposition of data on different scales and clustering using K means improves the predictive measures as compared to standalone LSTM.

Similarly, Duan et al. (2022) is about facilitating the predictions in WP using subseries generated after VMD decomposition and then PSO based LSTM and DBN are used for predictions. Quantitative comparative analysis is done on two different aspects, (i) two decomposition techniques, WD and EMD; and (ii) four state of art DL models; PSO-DBN, LSTM, ELMAN and BP. Proposed model achieved the metrics as: MAE 35.3776, RMSE 42.9055 and time 2254.06 s. Error is reduced to a great extent at an expense of time. Time taken is highest when compared to all the other base models and decomposition techniques.

2.3.3.3 LSTM with WD In Liu et al. (2020), the non-stationary time series power generation data of 40 data points from the seventh month of 2009 to the fourth month of 2019 are selected. A decomposition method named WD is applied over original data series and then then LSTM model is used for predictive analytics. Training and testing ratio is 80:20. MAPE of WD-LSTM is 5.831 as compared to LSTM having 13.715, performing better than other models because of data optimization by WD, which improves the prediction accuracy. However, LSTM takes 32 min to converge as compared to 44 min taken by WD-LSTM. Also, proposed model still needs to verify whether the time series with different scales can be used as the same model's input index. To boot, (Bhaskar and Singh 2012) proposed a two stage forecasting technique using WD and AWNN. The forecasting is carried out from 1 to 30 h-ahead values using one year (2004) NREL data at 10 min interval for training.

2.3.3.4 LSTM with DWT In another study Liu et al. (2019), a WP short-term forecasting method based on DWT and LSTM (DWT_LSTM) is proposed. Three time series WP data at 15 min period of 12 months each are used. DWT-LSTM when analysed over three different wind farms performed with an MAE of 10.12 against LSTM having 28.31 and this performance is consistent for all the three farms. Also, amongst other models LSTM, RNN, and BP, LSTM outperformed others for all wind farms. However, it still needs to optimize parameters like weight, learning rate, etc., to achieve higher accuracy. DWT_LSTM improves over five models namely DWT_RNN, DWT_BP, LSTM, RNN, and BP. Three models namely LSTM, RNN, and BP, against LSTM in DWT_LSTM and two DWT_RNN and DWT_BP for DWT part in DWT_LSTM are used as benchmarked models. Additionally, Zhang et al. (2022) also used DWT with SARIMA and LSTM to improve on the accuracy of WP prediction. The experiments are conducted on SCADA database of an offshore wind turbine in Scotland.

2.3.3.5 LSTM with EEMD and EMD For WP, LSTM's inclusion in any way is improving performance metrics. (i) GRU, (ii) LSTM, (iii) EEMD-LSTM, (ii) PCA-EEMD-LSTM, and (iv) SARIMA-PCA-EEMD-LSTM are applied in Zhou et al. (2020) over three meteorological datasets and concluded that the performance metrics MAE, MAPE, and RMSE all are reduced in SARIMA-PCA-EEMD-LSTM. However, the computational time taken is approximately 3 h to train the model in the worst-case scenario of all three datasets. Also, the efficient results are shown at the prediction interval of 15 min. Hence a model with efficient performance metrics with lesser computational time will be a great help. IVMD-SE-MCC-LSTM, EMD-SE-MCC-LSTM, MCC-LSTM, DBN, and Elman are various models used over WP data in Duan et al. (2020), and performance metrics are improved. However, only if the data is non-linear and has outliers, then this model will be efficient.

Improved variant of LSTM using EFG with CSO to optimize parameters is improving WP forecasts. The forecasts are happening in subseries generated from EEMD in Devi et al. (2020). Experiments are done on NIWE dataset with 24,192 data points collected from 2017 to 2018 at 15 min interval. EEMD- CSO-LSTM-EFG achieved improved results w.r.t. forecasting related to RMSE values for one-step-ahead to four-step-ahead forecasting. LSTM-EFG is compared with three models: ARMA, BPNN, and LSTM. Additionally, proposed EEMD-CSO-LSTM-EFG is compared with EEMD-CSO-SVM, EEMD-GA-BPNN, and EEMD-LSTM.

2.3.3.6 LSTM ensemble forecasting Then (Choi and Lee 2018) implements the LSTM ensemble forecasting method on publicly available data at Time Series Data Library (TSDL), River flow, Vehicles, Wine, and Airline, datasets. Ensemble LSTM improves over basic LSTM by as high as 46% over MSE and as low as 21% tested over 4 different travel time series datasets. Shi et al. (2018) uses two WP series of hourly data: summer and spring data from (1–31 August 2017) to (1–31 March 2017) respectively. MAE is improved for both R(recursive)-VMD-LSTM and D(direct)-VMD-LSTM. MAE for R-LSTM is improved to 0.28 over LSTM MAE 2.81. MAE for D-LSTM is reduced to 0.54 over LSTM with 5.04. Comparison is done with models BP, ELM, SVM, LSTM, EMD-ELM, EMD-SVM, EMD-LSTM, VMD-ELM and VMD-SVM.

2.3.3.7 LSTM with CNN Another hybrid approach adopted Woo et al. (2018) is convolution operation to extract spatial features, and the LSTM layer enables the modelling temporal dynamics of wind field and the corresponding two response time series data in a total 22,000 pairs of data points. It takes the spatio-temporal characteristics of wind flow input and multi-tasks learning reduces the complexity of model. The MAPE for power prediction is reduced to 0.841. However, indirect WP prediction via wind speed. a direct application should be tested for accuracy and efficiency. In Wu et al. (2016), one layer CNN + 2 layers LSTM model, one layer CNN + 2 layers RNN, 1 CNN + 4 layers FNN, and four layers FNN hybrid models are used over Year 2014 North East China wind farm, WP and wind speed data recorded every 15 min. Besides these hybrid DL methods, for comparison basic persistence method and the ARIMA method are used. Combination of LSTM and CNN works better than different hybrid methods and ARIMA with optimized cost. CNN-LSTM reduces the MAE to 5.62 as compared to 6.65 in FNN, 9.72 in ARIMA, and 12.81 in persistence. However, the models are tested only for point forecasting rather than interval one, which is the practical need in WP forecasting. Powernet model proposed in Garg and Krishnamurthi (2022) uses CNN-1D layers combined with BiLSTM model for WP predictive analytics.

2.3.3.8 Adaptive LSTM and GA The multivariable ALSTM WP prediction model is proposed in Xu and Xia (2018). Daily 144 data points from May 2016 to May 2018 at intervals of 10 min are adopted and Adaptive LSTM is applied over it. When compared to BR, LR, BP, and SVR improved the MAPE to 2.7 in contrast to the worst performer LR with MAPE approximately 19%.

2.3.3.9 Attention based LSTM (A-LSTM) SATCN-LSTM model in Xiang et al. (2022) predicts WP for ultra-short time interval. Taking California's fourth quarter WP forecast data, the proposed method has RMSE reduction of 17.56, 10.99, 11.34 and 3.68%

compared with LSTM, TCN, CNN-LSTM, and TCN-LSTM. Li et al. (2019), Sun et al. (2022), Xiong et al. (2022) are two evolutionary optimized A-LSTM models proposed. In the study Li et al. (2019), EA LSTM model used CRS genetic algorithm to optimize the weights of attention layer. Also, Sun et al. (2022) proposed a two stage attention based mechanism for improving the stability in input meteorological parameters and modified PSO for hyper-parameter tuning in LSTM.

2.3.3.10 LSTM encode decoder (LSTM ED) Another variant of LSTM, LSTM ED, in a hybrid fashion is proposed and compared with a model LSTM without AE in Lu et al. (2018). This LSTM ED performs better and more sensitive towards WP data and highly concentrated error distribution. The RMSE of the LSTM ED was 2.6, 5.2, 8.5 and 11.8%, lower than LSTM. However, the comparison is made with just one model and a compact dataset of a wind farm from January 1, 2016, to May 31st with an interval 15 min, a total of 14,496 observations. Another study Garg and Krishnamurthi (2023) uses a CNN-1D and LSTM in the ED architecture respectively. Using NREL dataset collected at hourly interval for 6 years, experiments are conducted to assess the proposed model with statistical models: AR, ARMA, ARIMA, ARIMA-LSTM and four conventional DL models: VanillaLSTM, StackedLSTM, BiLSTM and CNN-LSTM. Error metrics MSE and RMSE prove that CNN ED LSTM works better for WP predictive analytics as compared to other base models.

2.3.3.11 K-Means-LSTM Zhou et al. (2019) Proposes a K-Means-LSTM network model over a dataset, from Dec 2017 to Dec 2018, sampled with a gap of 10 min interval. K-Means-LSTM is proven to be better as it is reducing prediction error to 33.445 and 49.559 as compared to BP, ENN and SVR where worst is 58.446 of BP. However, selecting some unreasonable value for K influences and deteriorates the results. Hence, out of 2200 total sampled 2000 samples are required first to predict the value of K.

2.3.3.12 Evolutionary LSTM EA-LSTM: for time series prediction (Li et al. 2019; Sun et al. 2022) are two evolutionary optimized A-LSTM models proposed. In the study Li et al. (2019), EA-LSTM model used CRS genetic algorithm to optimize the weights of attention layer. Also, Sun et al. (2022) proposed a two stage attention based mechanism for improving the stability in input meteorological parameters and modified PSO for hyper-parameter tuning in LSTM. Vaitheeswaran (2019) used GA for optimum window size and number of units calculation and LSTM for short term and medium term predictions RMSE is reduced to 0.0957993 and 0.0929905 respectively.

2.3.3.13 Bidirectional LSTM As pet the study Ma and Mei (2022), decomposition and multistage analysis generate better results as compared to standalone models. However, decomposition is an extra overhead that leads to increases computational complexity. To overcome this issue, this paper implements a different DL model is applied to a specific part of the input series using Time2Vec layer and attention method. Combination of CNN and stacked BiLSTM create the required hybrid model. Similarly, in Garg and Krishnamurthi (2022) a model named Powernet is proposed using CNN-1D and BiLSTM layers. Extra overhead of decomposition is not incorporated in the analysis. MAE and MSE are improved against statistical and conventional DL models using Powernet.

3 Challenges

There are different researches to optimize LSTM and WP forecasting. However, there are challenges which always need to be addressed for better performance. Major challenges are, (i) data related, (ii) model related, (iii) computational time, (iv) relevant comparisons and, (v) measurement of forecasting error.

3.1 Data related

The first and foremost requirement while assessing WP and making predictions is the data. Underfitting and overfitting of data are the most common issues which need special care. Some of the prevalent challenges that are seen while applying predictive models are listed below.

3.1.1 Uncertainty of data (Zhang et al. 2019; Balanchine 2018)

Uncertainty of the data is one of the challenging, striking features of chaos despite a deterministic evolution. Uncertainty is directly proportional to unpredictability. This leads to difficulty in building a dependency relation between sequences and instability in memorization part of time series predictive analytics. This is a challenge to resolve uncertainty in data. Sometimes resolving uncertainty may lead to information loss.

3.1.2 Voluminous structured data (Do et al. 2019; Balanchine 2018)

A time series is a collection of sequences that track the progress of an activity through time. Sequences are records made at evenly spaced intervals in a consistent way of measurement or collection of data. This collection is generally made via some IoT device like sensors with varied number of data items. Equally spaced interval can be as small as a second. Hence, this process generated a voluminous data sequences. To manage such volumes of data and then to decide how much data, what all features are required for making optimal predictions. Hence, the volume of data is a challenge not only in terms of computation time but relevant feature selection as well.

3.1.3 Data quality (Mehdiyev et al. 2017; Ismail et al. 2020)

Raw collected data needs pre-processing to handle inconsistency and redundancy. Hence, cleaning of data involves removing duplicates, null values and missing values, redundant values, transforming categorical data to numerical and normalized one. But, the challenge is to decide on the level of cleansing to be done. Sometimes data becomes over clean and leads to overfitting during training.

3.1.4 Data preparation (Li and Jin 2018)

Raw data is another challenge that needs attention and processing for efficient predictions. There are missing values, null values, NAN problem, categorical data, outliers etc. So pre-processing of data to decompose, cleanse, normalize, reduce, filter is important before applying time series predictive analytics.

3.1.5 Not enough/too much data (Balanchine 2018)

The models are fed with a small subset of data available for predictive analytics, despite the fact that time-series data are sometimes huge in size and have a significant number of features. Model reliability is greatly influenced by the quantity of the training dataset, among other things. Hence, a larger dataset is sometimes more dependable than a smaller one during training. Cold start is one of the ML difficulties in which the model is unusable owing to a lack of data. In terms of statistics, supplying the model with data only improves predictions by a small amount. As a result, computational and time resources may be wasted.

3.1.6 Non linear data (Zhang and Thearling 1994, 2004)

The measuring methodology, sampling strategy, and the shape and interpretation of the data influence modelling and prediction. Nonlinear time series analysis is a relatively new approach to complex dynamical phenomena. The analyst reviews the data using numerous approaches, including a visual assessment of several plots that can be obtained before processing, before applying any modelling or forecasting approaches. Traditional linear stochastic models—specifically, AR, MA, ARMA, and ARIMA models—were used in early attempts (Bradley and Kantz 2015). However, thoroughly reviewing and analysing data before using predictive analytics remains a hurdle.

3.1.7 Spatio-temporal aspect in data (Balanchine 2018; Wang et al. 2019)

Deciding on data space dimensionality is another crucial challenge to estimate. Moreover, both spatio-temporal features make the dimension size increase to a great extent. This is even more challenging for non-linear type of data sequences to estimate data length horizontally.

3.1.8 An appropriate sampling of data (Camilleri 2004)

Appropriate sampling means taking consistent samples multiple times with a trend in selection. Strategy to select samples involve selecting samples at the same time every day to include seasonality, if hourly samples are to be drawn then fix the time intervals, for daily samples the timing of sample selection should be kept same. These tricks impact data and in turn prediction.

3.2 Model related (Balanchine 2018)

A model formulation first in mathematical terms, second in terms of dependent and independent variables, third in terms of optimization is challenging. The varied list of challenges concerning the predictive models is listed below.

3.2.1 Complexity of model (Hu and Chen 2018; Balanchine 2018)

The model selected can neither be too simple nor be too expensive. Moreover, predictive modelling performance depends on choosing and adjusting different hyper parameters, challenging while applied on other datasets.

3.2.2 Oversimplified

In terms of computational power, simple linear models are always better. But when there is trade-off between time and error and performance is the key it is a challenge to choose one over other. Best is always not the simple one. But, how simple or complex a model can be is another difficult challenge when building a model.

3.2.3 Too much expensive model (Zhang et al. 2017)

In contrast to a simple model, how complex a model should be is again a variable parameter. Various complexities may include (i) Number of input variables, (ii) number of timestamps, (iii) depth of a model, (iv) level of hybridization, (v) number of neurons, (vi) optimization function, (vi) loss and activation function and so on. Depending on the data and type of predictive analysis in hand model and different parameters associated with model need to be selected.

3.2.4 Wrong features selection (Ismail et al. 2020)

In depth knowledge of domain and knowledge of modelling techniques is the key to prediction efficiency. Selection of appropriate input and output features sometimes need extra work. Features might turn into noise if not selected optimally.

3.2.5 Hyper-parameter selection and tuning

To optimize performance of a model over given problem in hand hyper-parameters needs to be selected wisely. Hyper-parameter tuning is one of the solutions but computationally very expensive. Hence, opting and performing hyper-parameter tuning and the algorithm to be used is another challenge in the model architecture.

3.3 Computational time (Bokde et al. 2018)

Complex and deep models sometimes add to the cost in terms of resources and computational power instead of improving the prediction efficiency. However, most of the time, this performance metric is ignored while proposing a better predictive analytical model.

3.4 Relevant comparison (Qin 2019)

An apple should always be compared with an apple. The same is applicable when a new efficient approach is taken and proven against previous work done. The last work comparison should be with the same league of models or pre-processing techniques instead of random comparisons.

3.5 Measurement of forecasting error (Zhang et al. 2019; Maiti and Bidinger 1981)

After a model is used to make predictive analytics some measurement method is required to quantify the performance the respective model. This quantification is generally by calculating how far predicted value from the actual one is. The aim is to reduce this gap between predictions and actual ones as much as possible. At the end, the requirement is get a single measure for the model to see if it is an improvement or not. Various measures available are error based rather than accuracy based namely; MAE, RMSE, MAE, MAPE, MRE, Normalized RMSE (NRMSE), sum of squared error, error percentage, MPE. This quantification can be done separately on training and test parts.

3.6 LSTM: challenges

LSTM has the property to extract temporal features of the input series and enhance the predictive analytics by incorporating this property into future timestamp predictions. However, since LSTM works better for long term dependencies combined with short term ones, the length of input dimensions should be long enough while designing model architecture. Hence, there are a few challenges associated while performing WP predictive analytics using LSTM and its associated hybrid models:

- LSTM based DL models are quite complex because of their depth and number of neurons. This complex structure affects computational complexity and hence become less suitable for ultras short term predictive analytics.
- Requirement of long term dependency state tend to increase the input dimensions, which increases the number of trainable parameters.
- With increase in depth of the model architecture, the optimization may go into local optima and hence reduce the predictive accuracy.
- Choosing fixed set of hyper-parameters in LSTM may prove to be optimum for a certain set of experiments but not for all.
- Choice of activation function needs to be consistent with type of data, its frequency, prediction interval and the model architecture.
- Use of data decomposition as pre-processing step might lead to extra overhead. It is another challenge like hyper-parameter selection to include as an extra layer.

4 Conclusion

In the field of sustainable energy, WP predictive analysis is a booming and challenging requirement. To expand the area of wind power generation and install more wind farms in future; accurate predictive analytics is mandatory. However, the sequential WP time series data collected is non-linear, uncertain/unpredictable, non-stationary, dirty, and voluminous. Due to uncertainty and stochastic nature of wind power time series parameters and outputs, enormous data driven; various machine learning and deep learning approaches have been proposed in the state of art. This non-linear type of data has successfully adopted LSTM and its associated hybrid models as efficient modelling approach in many studies. This study provides critical insights on usage LSTM and associated model in wind power predictions. For this, firstly review of requirement of time series predictive analytics is presented. Different business and social values from different recent areas, where time series predictive analytics is applied, are discussed. This step explains the need of predictive analytics in time series kind of data. Second is the study of models for variety of time series data and their performance metrics. This step gives a broader picture of different areas which generate time series data and different models and metrics available to measure performance of time series predictive analysis. Third and final study is specifically for WP predictive analytics, role of LSTM and its associated hybrid models, studies conducted for the same and their pros and cons. This focuses into a comprehensive survey of various researches done on WP predictive analytics using LSTM, as full or part. The merits and demerits of respective study are identified. Also, different associations of LSTM with other models or/and with decomposition methods are critically explored. These partial studies when combined together, give a brief idea about importance of LSTM, its variants and its associations in WE predictive analytics. This article provides extensive literature on the (i) time series predictive application areas and its benefits, (ii) modelling techniques utilized for time series predictive analytics and their quantitative analysis, (iii) rigorous classification of WP predictive analytics, (iv) role of LSTM and its associated hybrid models in WP predictive analytics, (v) merits and demerits of state of art LSTM studies, and (v) challenges that still need attention from the researchers.

Author contributions Conception and design of study, acquisition of data, analysis and/or interpretation of data, drafting the manuscript: SG;Revising the manuscript critically for important intellectual content: RK.

Declarations

Competing interest The authors declare no competing interests.

References

Abdoos AA (2016) A new intelligent method based on combination of VMD and ELM for short term wind power forecasting. Neurocomputing 203:111–120. https://doi.org/10.1016/j.neucom.2016.03.054

Abdullahi M et al (2022) Detecting cybersecurity attacks in internet of things using artificial intelligence methods: a systematic literature review. Electronics 11:198. https://doi.org/10.3390/ELECTRONICS11020198

Aditya Pai B, Devareddy L, Hegde S, Ramya BS (2022) A time series cryptocurrency price prediction using LSTM. Lect. Notes Electr. Eng. 790:653–662. https://doi.org/10.1007/978-981-16-1342-5_50/COVER

Akbal Y, Ünlü KD (2022) A univariate time series methodology based on sequence-to-sequence learning for short to midterm wind power production. Renew Energy 200:832–844. https://doi.org/10.1016/J.RENENE.2022.10.055

Akhtar I, Kirmani S, Ahmad M, Ahmad S (2021) Average monthly wind power forecasting using fuzzy approach. IEEE Access 9:30426–30440. https://doi.org/10.1109/ACCESS.2021.3056562

Akita R (2016) Deep learning for stock prediction using numerical and textual information. 2016 IEEE/ACIS 15th Conf Comput Inf Sci 128:1–6

Alkesaiberi A, Harrou F, Sun Y (2022) Efficient wind power prediction using machine learning methods: a comparative study. Energies 15(7):2327. https://doi.org/10.3390/EN15072327

Artipoli G, Durante F (2014) Physical modeling in wind energy forecasting. Dewi Mag. (44):10–15. https://pdfs.semanticscholar.org/baaa/0abdb12157ebe381ec4ffc1d1621b15b3a3f.pdf

Balanchine S (2018) Challenges requirements for building a predictive analysis model. https://www.cetrixcloudservices.com/blog/challenges-requirements-for-building-a-predictive-analysis-model. accessed March 22, 2021

Bao W (2017) A deep learning framework for financial time series using stacked autoencoders and long-short term memory. Int Commun Red Cross. https://doi.org/10.6084/m9.figshare.5028110

Bhaskar K, Singh SN (2012) AWNN-Assisted wind power forecasting using feed-forward neural network. IEEE Trans Sustain Energy 3(2):306–315. https://doi.org/10.1109/TSTE.2011.2182215

Bokde N, Feijóo A, Kulat K (2018) Analysis of differencing and decomposition preprocessing methods for wind speed prediction. Appl Soft Comput J 71:926–938. https://doi.org/10.1016/j.asoc.2018.07.041

Bradley E, Kantz H (2015) Nonlinear time-series analysis revisited. Chaos. https://doi.org/10.1063/1.4917289

Camilleri M (2004) Forecasting using non-linear techniques in time series analysis: an overview of techniques and main issues. Univ Malta Comput Sci Anu Res Work, 19–28

Chang W (2014) A literature review of wind forecasting methods. J Power Energy Eng 2:161–168

Chen Y, He Z, Shang Z, Li C, Li L, Xu M (2019) A novel combined model based on echo state network for multi-step ahead wind speed forecasting: a case study of NREL. Energy Convers. Manag. 179:13–29. https://doi.org/10.1016/j.enconman.2018.10.068

Chen H, Birkelund Y, Anfinsen SN, Yuan F (2021) Comparative study of data-driven short-term wind power forecasting approaches for the Norwegian Arctic region. J Renew Sustain Energy. https://doi.org/10.1063/5.0038429/926618

Chen C et al (2022) Forecast of rainfall distribution based on fixed sliding window long short-term memory. Eng App Comput Fluid Mech 16(1):248–261. https://doi.org/10.1080/19942060.2021.2009374

Choi JY, Lee B (2018) Combining LSTM network ensemble via adaptive weighting for improved time series forecasting. Math Probl Eng 2018:1–8. https://doi.org/10.1155/2018/2470171

Cui Z, Ke R, Pu Z, Wang Y (2018) Stacked bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. arXiv, pp. 1–11

Deep B, Mathur I, Joshi N (2022) An approach to forecast pollutants concentration with varied dispersion. Int J Environ Sci Technol 19(6):5131–5138. https://doi.org/10.1007/S13762-021-03378-Z/FIGURES/6

Devi AS, Maragatham G, Boopathi K, Rangaraj AG (2020) Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTM-EFG deep learning technique. Soft Comput 24(16):12391–12411. https://doi.org/10.1007/s00500-020-04680-7

Do LNN, Taherifar N, Vu HL (2019) Survey of neural network-based models for short-term traffic state prediction. Wiley Interdiscip Rev Data Min Knowl Discov 9(1):1–24. https://doi.org/10.1002/widm.1285

Du P, Wang J, Yang W, Niu T (2019) A novel hybrid model for short-term wind power forecasting. Appl Soft Comput J 80:93–106. https://doi.org/10.1016/j.asoc.2019.03.035

Du S, Li T, Yang Y, Horng SJ (2020) Multivariate time series forecasting via attention-based encoder–decoder framework. Neurocomputing 388:269–279. https://doi.org/10.1016/j.neucom.2019.12.118

Duan J et al (2020) Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and Correntropy long short -term memory neural network. Energy 214:118980. https://doi.org/10.1016/j.energy.2020.118980

Duan J, Wang P, Ma W, Fang S, Hou Z (2022) A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting. Int J Electr Power Energy Syst 134:107452. https://doi.org/10.1016/j.ijepes.2021.107452

Eapen J, Bein D, Verma A (2019) Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. 2019 IEEE 9th Annu Comput Commun Work Conf CCWC, 2019 128:264–270. https://doi.org/10.1109/CCWC.2019.8666592

Energy GW (2016) Global wind energy opening up new markets for business

Essien A, Giannetti C (2019) A deep learning framework for univariate time series prediction using convolutional LSTM stacked autoencoders. IEEE Int. Symp. Innov. Intell. Syst. Appl. INISTA 2019, Proc. 128:1–6. https://doi.org/10.1109/INISTA.2019.8778417

Filonov P, Lavrentyev A, Vorontsov A (2016) Multivariate industrial time series with cyber-attack simulation: fault detection using an LSTM-based predictive data model, pp. 1–8. http://arxiv.org/abs/1612.06676

Fu W, Wang K, Li C, Tan J (2019) Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM. Energy Convers Manag 187(Feb):356–377. https://doi.org/10.1016/j.enconman.2019.02.086

Garg S, Krishnamurthi R (2022) Powernet: a novel method for wind power predictive analytics using Powernet deep learning model. J Renew Sustain Energy. https://doi.org/10.1063/5.0090126/2848645

Garg S, Krishnamurthi R (2023) A CNN encoder decoder LSTM model for sustainable wind power predictive analytics. Sustain Comput Inf Syst 38:100869. https://doi.org/10.1016/J.SUSCOM.2023.100869

Geetha A (2016) Nasira DGM (2016) Time series modeling and forecasting tropical cyclone prediction using ARIMA model. Int J Soc Syst Sci 128:3080–3086

Global wind energy council (2022) Annual wind report

Global Wind Report (2022) https://gwec.net/global-wind-report-2022/. Accessed June 27, 2022

Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J (2017) LSTM: a search space Odyssey. IEEE Trans Neural Netw Learn Syst 28(10):2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924

Guo Z, Zhao W, Lu H, Wang J (2012) Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renew Energy 37(1):241–249. https://doi.org/10.1016/j.renene.2011.06.023

Han L, Romero CE, Yao Z (2015) Wind power forecasting based on principle component phase space reconstruction. Renew Energy 81:737–744. https://doi.org/10.1016/j.renene.2015.03.037

Han Q, Meng F, Hu T, Chu F (2017) Non-parametric hybrid models for wind speed forecasting. Energy Convers Manag 148:554–568. https://doi.org/10.1016/j.enconman.2017.06.021

Han L, Zhang R, Wang X, Bao A, Jing H (2019) Multi-step wind power forecast based on VMD-LSTM. IET Renew Power Gen 13(10):1690–1700. https://doi.org/10.1049/iet-rpg.2018.5781

He B et al (2022) A combined model for short-term wind power forecasting based on the analysis of numerical weather prediction data. Energy Rep 8:929–939. https://doi.org/10.1016/j.egyr.2021.10.102

Heidari A, Khovalyg D (2020) Short-term energy use prediction of solar-assisted water heating system: application case of combined attention-based LSTM and time-series decomposition. Sol Energy 207:626–639. https://doi.org/10.1016/j.solener.2020.07.008

Hong T, Pinson P, Fan S, Zareipour H, Troccoli A, Hyndman RJ (2016) Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond. Int J Forecast 32(3):896–913. https://doi.org/10.1016/j.ijforecast.2016.02.001

Hossain M et al (2018) Application of the hybrid ANFIS models for long term wind power density prediction with extrapolation capability. Plos One 13(4):e0193772

Hu YL, Chen L (2018) A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm. Energy Convers Manag 173:123–142. https://doi.org/10.1016/j.enconman.2018.07.070

Hu T, Wu W, Guo Q, Sun H, Shi L, Shen X (2020) Very short-term spatial and temporal wind power forecasting: a deep learning approach. CSEE J Power Energy Syst 6(2):434–443. https://doi.org/10.17775/CSEEJPES.2018.00010

Hua Y, Zhao Z, Li R, Chen X, Liu Z, Zhang H (2018) Deep learning with long short-term memory for time series prediction. IEEE Commun Mag. 128:114–119

Huang L, Li L, Wei X, Zhang D (2022) Short-term prediction of wind power based on BiLSTM–CNN–WGAN-GP. Soft Comput 26(20):10607–10621. https://doi.org/10.1007/s00500-021-06725-x

Ismail AA, Gunady M, Bravo HC, Feizi S (2020) Benchmarking deep learning interpretability in time series predictions. No. NeurIPS. http://arxiv.org/abs/2010.13924

Javid AM, Liang X, Venkitaraman A, Chatterjee S (2020) Predictive analysis of COVID-19 time-series data from Johns Hopkins University. arXiv, pp. 1–18

Jiang G, Chen Z, Li X, Yan X (2020) Short-term prediction of wind power based on EEMD-ACS-LSSVM. Taiyangneng Xuebao/acta Energiae Solaris Sin 41(5):77–84

Kim S, Kang M (2019) Financial series prediction using attention LSTM. arXiv

Kim Y, Hur J (2020) An ensemble forecasting model of wind power outputs based on improved statistical approaches. Energies 13(5):1071. https://doi.org/10.3390/EN13051071

Lazzari F et al (2022) User behaviour models to forecast electricity consumption of residential customers based on smart metering data. Energy Rep 8:3680–3691. https://doi.org/10.1016/J.EGYR.2022.02.260

Li R, Jin Y (2018) A wind speed interval prediction system based on multi-objective optimization for machine learning method. Appl Energy 228:2207–2220. https://doi.org/10.1016/j.apenergy.2018.07.032

Li Y, Wu H, Liu H (2018) Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction. Energy Convers Manag 167:203–219. https://doi.org/10.1016/j.enconman.2018.04.082

Li Y, Zhu Z, Kong D, Han H, Zhao Y (2019) EA-LSTM: evolutionary attention-based LSTM for time series prediction. Knowl-Based Syst. 181:104785. https://doi.org/10.1016/j.knosys.2019.05.028

Liu Y, Che P (2019) Short-term wind power prediction based on dynamic STARMA model. Proc 31st Chinese Control Decis Conf CCDC 2019. https://doi.org/10.1109/CCDC.2019.8832755

Liu R, Peng M, Xiao X (2018) Ultra-short-term wind power prediction based on multivariate phase space reconstruction and multivariate linear regression. Energies 11(10):2763. https://doi.org/10.3390/EN11102763

Liu H, Mi X, Li Y (2018) Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. Energy Convers Manag 159:54–64. https://doi.org/10.1016/j.enconman.2018.01.010

Liu C-L et al (2019) Time series classification with multivariate convolutional neural network. IEEE Trans Ind Electron 66(6):4788–4797

Liu Y et al (2019) Wind power short-term prediction based on LSTM and discrete wavelet transform. Appl Sci 9(6):1108. https://doi.org/10.3390/app9061108

Liu CL, Hsaio WH, Tu YC (2019) Time series classification with multivariate convolutional neural network. IEEE Trans Ind Electron 66(6):4788–4797. https://doi.org/10.1109/TIE.2018.2864702

Liu B, Zhao S, Yu X, Zhang L, Wang Q (2020) A novel deep learning approach for wind power forecasting based on WD-LSTM model. Energies 13(18):1–17. https://doi.org/10.3390/en13184964

Liu M, Li G, Li J, Zhu X, Yao Y (2021) Forecasting the price of Bitcoin using deep learning. Financ. Res. Lett. 40:101755. https://doi.org/10.1016/J.FRL.2020.101755

Liu C et al (2022) Numerical weather prediction enhanced wind power forecasting: rank ensemble and probabilistic fluctuation awareness. Appl Energy 313:118769. https://doi.org/10.1016/J.APENERGY.2022.118769

Livieris IE, Pintelas E, Pintelas P (2020) A CNN–LSTM model for gold price time-series forecasting. Neural Comput Appl 32(23):17351–17360. https://doi.org/10.1007/S00521-020-04867-X/TABLES/8

Lu K et al (2018) Short-term wind power prediction model based on encoder-decoder LSTM. IOP Conf Ser Earth Environ Sci 186(5):012020. https://doi.org/10.1088/1755-1315/186/5/012020

Ma TY, Faye S (2022) Multistep electric vehicle charging station occupancy prediction using hybrid LSTM neural networks. Energy 244:123217. https://doi.org/10.1016/J.ENERGY.2022.123217

Ma Z, Mei G (2022) A hybrid attention-based deep learning approach for wind power prediction. Appl. Energy 323:119608. https://doi.org/10.1016/j.apenergy.2022.119608

Maçaira PM, Tavares Thomé AM, Cyrino Oliveira FL, Carvalho-Ferrer AL (2018) Time series analysis with explanatory variables: a systematic literature review. Environ Model Softw 107:199–209. https://doi.org/10.1016/j.envsoft.2018.06.004

Magadum RB, Bilagi S, Bhandarkar S, Patil A, Joshi A (2023) Short-term wind power forecast using time series analysis: auto-regressive moving-average model (ARMA). Lect Notes Electr Eng 979:319–341. https://doi.org/10.1007/978-981-19-7993-4_26/COVER

Maiti, Bidinger (1981) TIME series forecasting: a non linear dynamics approcach. J. Chem. Inf. Model. 53(9):1689–1699

Maldonado-Correa J, Solano JC, Rojas-Moncayo M (2021) Wind power forecasting: a systematic literature review. Wind Eng 45(2):413–426. https://doi.org/10.1177/0309524X19891672

Manwell AL, McGowan JF, Rogers JG (2010) Wind energy explained: theory design and application. Wiley

Masseran N (2015) Evaluating wind power density models and their statistical properties. Energy. https://doi.org/10.1016/j.energy.2015.03.018

Md Azmi CSA et al (2022) Univariate and multivariate regression models for short-term wind energy forecasting. Inf Sci Lett 11(2):465–473. https://doi.org/10.18576/isl/110217

Mehdiyev N, Lahann J, Emrich A, Enke D, Fettke P, Loos P (2017) Time series classification using deep learning for process planning: a case from the process industry. Procedia Comput Sci 114:242–249. https://doi.org/10.1016/j.procs.2017.09.066

Meng A et al (2022) A novel wind power prediction approach using multivariate variational mode decomposition and multi-objective crisscross optimization based deep extreme learning machine. Energy 260:124957. https://doi.org/10.1016/J.ENERGY.2022.124957

Messner JW, Pinson P (2019) Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting. Int J Forecast 35(4):1485–1498. https://doi.org/10.1016/J.IJFORECAST.2018.02.001

Mishra S, Bordin C, Taharaguchi K, Palu I (2020) Comparison of deep learning models for multivariate prediction of time series wind power generation and temperature. Energy Rep 6:273–286. https://doi.org/10.1016/J.EGYR.2019.11.009

Mustaffa Z, Sulaiman MH, Rohidin D, Ernawan F, Kasim S (2018) Time series predictive analysis based on hybridization of meta-heuristic algorithms. Int J Adv Sci Eng Inf Technol 8(5):1919–1925. https://doi.org/10.18517/ijaseit.8.5.4968

Muthamizharasan M, Ponnusamy R (2022) Forecasting crime event rate with a CNN-LSTM model. Lect Notes Data Eng Commun Technol 96:461–470. https://doi.org/10.1007/978-981-16-7167-8_33/COVER

Nagaraj P (2022) Forecasting cyber attacks using machine learning. J Optoelectron Laser 41(7):550–556. http://www.gdzjg.org/index.php/JOL/article/view/746. Accessed May 28, 2023

Naik J, Bisoi R, Dash PK (2018) Prediction interval forecasting of wind speed and wind power using modes decomposition based low rank multi-kernel ridge regression. Renew Energy 129:357–383. https://doi.org/10.1016/j.renene.2018.05.031

Nguyen LQ, Fernandes PO, Teixeira JP (2021) Analyzing and forecasting tourism demand in Vietnam with artificial neural networks. Forecast 4:36–50. https://doi.org/10.3390/FORECAST4010003

Nowotarski J, Weron R (2018) Recent advances in electricity price forecasting: a review of probabilistic forecasting. Renew Sustain Energy Rev 81:1548–1568. https://doi.org/10.1016/j.rser.2017.05.234

Osório GJ, Lotfi M, Shafie-khah M, Campos VMA, Catalão JPS (2018) Hybrid forecasting model for short-term electricity market prices with renewable integration. Sustain 11(1):1–15. https://doi.org/10.3390/su11010057

Park HJ, Kim Y, Kim HY (2022) Stock market forecasting using a multi-task approach integrating long short-term memory and the random forest framework. Appl Soft Comput. 114:108106. https://doi.org/10.1016/J.ASOC.2021.108106

Pavlyshenko BM (2019) Machine-learning models for sales time series forecasting. Data 4(1):1–11. https://doi.org/10.3390/data4010015

Pitteloud J (2020) Global wind installations. Wind Energy Int. https://library.wwindea.org/global-statistics/

Qian X (2017) Financial series prediction: comparison between precision of time series models and machine learning methods. arXiv, pp. 1–9

Qian Z, Pei Y, Zareipour H, Chen N (2019) A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl Energy 235:939–953. https://doi.org/10.1016/J.APENERGY.2018.10.080

Qin Y et al (2019) Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. Appl Energy 236:262–272. https://doi.org/10.1016/j.apenergy.2018.11.063

Qin H (2019) Comparison of deep learning models on time series forecasting: a case study of dissolved oxygen prediction. arXiv

Qing X, Niu Y (2018) Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 148:461–468. https://doi.org/10.1016/j.energy.2018.01.177

Ramadan HS (2017) Wind energy farm sizing and resource assessment for optimal energy yield in Sinai Peninsula, Egypt. J Clean Prod 161:1283–1293. https://doi.org/10.1016/j.jclepro.2017.01.120

Redondo-Bravo L, Ruiz-Huerta C, Gomez-Barroso D, Sierra-Moros MJ, Benito A, Herrador Z (2019) Imported dengue in Spain: a nationwide analysis with predictive time series analyses. J Travel Med 26(8):1–9. https://doi.org/10.1093/jtm/taz072

Report W (2020) Global offshore wind report 2020, pp. 1–102

Rogachev A et al. (2022) Systematic analysis of retrospective crop yields time series based on their structure identification. https://iopscience.iop.org/article/https://doi.org/10.1088/1755-1315/1069/1/012014/meta. Accessed: May 28, 2023

Santhosh M, Venkaiah C, VinodKumar DM (2020) Current advances and approaches in wind speed and wind power forecasting for improved renewable energy integration: a review. Eng Rep 2(6):1–20. https://doi.org/10.1002/eng2.12178

Sarveswararao V, Ravi V, Vivek Y (2023) ATM cash demand forecasting in an Indian bank with chaos and hybrid deep learning networks. Expert Syst. Appl. 211:118645. https://doi.org/10.1016/J.ESWA.2022.118645

Schmela R, Beauvais M, Chevillard A, Paredes N, Heisz MG, Rossi M (2018) Global market outlook for solar power. Glob. Mark. Outlook, p. 92. http://www.solarpowereurope.org/wp-content/uploads/2019/05/SolarPower-Europe-Global-Market-Outlook-2019-2023.pdf

Shamshirband S, Rabczuk T, Chau K (2019) A survey of deep learning techniques: application in wind and solar energy resources. IEEE Access 7:164650–164666. https://doi.org/10.1109/ACCESS.2019.2951750

Shi X, Lei X, Huang Q, Huang S, Ren K, Hu Y (2018) Hourly day-ahead wind power prediction using the hybrid model of variational model decomposition and long short-term memory. Energies 11(11):1–20. https://doi.org/10.3390/en11113227

Singh PK, Singh N, Negi R (2019) Wind power forecasting using hybrid ARIMA-ANN technique. Adv Intell Syst Comput 904:209–220. https://doi.org/10.1007/978-981-13-5934-7_19/COVER

Sørensen ML, Nystrup P, Bjerregård MB, Møller JK, Bacher P, Madsen H (2023) Recent developments in multivariate wind and solar power forecasting. Wiley Interdiscip Rev Energy Environ 12(2):e465. https://doi.org/10.1002/WENE.465

Srivastava T, Vedanshu, Tripathi MM (2020) Predictive analysis of RNN, GBM and LSTM network for short-term wind power forecasting. J Stat Manag Syst 23(1):33–47. https://doi.org/10.1080/09720510.2020.1723224

Study GC, De Alencar DB, Affonso CDM, Roberto C (2017) Different models for forecasting wind power. Energies. https://doi.org/10.3390/en10121976

Sun Z, Zhao S, Zhang J (2019) Short-term wind power forecasting on multiple scales using VMD decomposition, k-means clustering and LSTM principal computing. IEEE Access 7:166917–166929. https://doi.org/10.1109/ACCESS.2019.2942040

Sun Y, Wang X, Yang J (2022) modified particle swarm optimization with attention-based LSTM for wind power prediction. Energies 15(12):4334. https://doi.org/10.3390/EN15124334

Torres JF, Troncoso A, Koprinska I, Wang Z, Martínez-Álvarez F (2019) Deep learning for big data time series forecasting applied to solar power. Adv Intell Syst Comput 771:123–133. https://doi.org/10.1007/978-3-319-94120-2_12

Vaitheeswaran SS (2019) Wind power pattern prediction in time series measuremnt data for wind energy prediction modelling using LSTM-GA networks. In 2019 10th International Conference on Computer and Communication Network Technology, pp. 1–5.

van der Westhuizen J, Lasenby J (2018) The unreasonable effectiveness of the forget gate, pp. 1–15. http://arxiv.org/abs/1804.04849

Vanitha V, Sophia G, Resmi JR, Raphel D (2020) Artificial intelligence-based wind forecasting using variational mode decomposition. Comput Intell 37:1–13. https://doi.org/10.1111/coin.12331

Vicente JMF, Álvarez-Sánchez JR, de la Paz-López F, Moreo JT, Adeli H (2017) Deep learning-based approach for time series forecasting with application to electricity load. Lect Notes Comput Sci. https://doi.org/10.1007/978-3-319-59773-7

Wang X, Guo P, Huang X (2011) A review of wind power forecasting models. Energy Procedia 12:770–778. https://doi.org/10.1016/j.egypro.2011.10.103

Wang CH, Cheng HY, Deng YT (2018) Using Bayesian belief network and time-series model to conduct prescriptive and predictive analytics for computer industries. Comput Ind Eng 115:486–494

Wang S, Cao J, Yu PS (2019) Deep learning for spatio-temporal data mining: a survey. IEEE Trans Knowl Data Eng 14(8):3681–3700. https://doi.org/10.1109/tkde.2020.3025580

Wang Y, Yu Y, Cao S, Zhang X, Gao S (2020) A review of applications of artificial intelligent algorithms in wind farms. Artif Intell Rev 53(5):3447–3500. https://doi.org/10.1007/S10462-019-09768-7/TABLES/10

Woo S, Park J, Park J (2018) Predicting wind turbine power and load outputs by multi-task convolutional LSTM model. IEEE Power Energy Soc Gen Meet. https://doi.org/10.1109/PESGM.2018.8586206

Wu W, Chen K, Qiao Y, Lu Z (2016) Probabilistic short-term wind power forecasting based on deep neural networks. 2016 Int Conf Probabilistic Methods Appl Power Syst PMAPS. https://doi.org/10.1109/PMAPS.2016.7764155

Wu X, Chen N, Du Q, Mao S, Ju X (2023) Short-term wind power prediction model based on ARMA-GRU-QPSO and error correction. J Phys Conf Ser 2427(1):012028. https://doi.org/10.1088/1742-6596/2427/1/012028

Xiang L, Liu J, Yang X, Hu A, Su H (2022) Ultra-short term wind power prediction applying a novel model named SATCN-LSTM. Energy Convers Manag 252:115036. https://doi.org/10.1016/j.enconman.2021.115036

Xiao C, Chen N, Hu C, Wang K, Gong J, Chen Z (2019) Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. Remote Sens. Environ. 233:111358. https://doi.org/10.1016/j.rse.2019.111358

Xiong B, Lou L, Meng X, Wang X, Ma H, Wang Z (2022) Short-term wind power forecasting based on attention mechanism and deep learning. Electr. Power Syst. Res. 206:107776. https://doi.org/10.1016/j.epsr.2022.107776

Xu G, Xia L (2018) Short-term prediction of wind power based on adaptive LSTM. Taiyangneng Xuebao/acta Energiae Solaris Sin 41(5):77–84

Xu HY, Chang YQ, Wang FL, Wang S, Yao Y (2021) Univariate and multivariable forecasting models for ultra-short-term wind power prediction based on the similar day and LSTM network. J Renew Sustain Energy. https://doi.org/10.1063/5.0027130/285201

Yang TY, Brinton CG, Joe-Wong C, Chiang M (2017) Behavior-based grade prediction for MOOCs via time series neural networks. IEEE J Sel Top Signal Process 11(5):716–728. https://doi.org/10.1109/JSTSP.2017.2700227

Yu C, Li Y, Bao Y, Tang H, Zhai G (2018) A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. Energy Convers Manag 178(Oct):137–145. https://doi.org/10.1016/j.enconman.2018.10.008

Yu R et al (2019) LSTM-EFG for wind power forecasting based on sequential correlation features. Futur Gener Comput Syst 93:33–42. https://doi.org/10.1016/j.future.2018.09.054

Yuan X, Chen C, Jiang M, Yuan Y (2019) Prediction interval of wind power using parameter optimized Beta distribution based LSTM model. Appl Soft Comput J 82:105550. https://doi.org/10.1016/j.asoc.2019.105550

Zafar M, Sharif MI, Sharif MI, Kadry S, Bukhari SAC, Rauf H (2023) Skin lesion analysis and cancer detection based on machine/deep learning techniques: a comprehensive survey. Life 13:146. https://doi.org/10.3390/LIFE13010146

Zeroual A, Harrou F, Dairi A, Sun Y (2020) Deep learning methods for forecasting COVID-19 time-series data: a comparative study. Chaos Solitons Fractals 140:110121. https://doi.org/10.1016/j.chaos.2020.110121

Zhang Y (2019) A novel hybrid model for wind speed prediction based on VMD and neural network considering atmospheric uncertainties. IEEE Access 7:60322–60332. https://doi.org/10.1109/ACCESS.2019.2915582

Zhang C, Zeng J, Xie N, Yang P, Zhang Y, Zhang Z (2016) Research on short-term wind power prediction based on combined forecasting models. MATEC Web Conf 70:1–5. https://doi.org/10.1051/matecconf/20167009005

Zhang L, Tan J, Han D, Zhu H (2017) From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today 22(11):1680–1685. https://doi.org/10.1016/j.drudis.2017.08.010

Zhang J, Yan J, Infield D, Liu Y, Lien F-S (2019) Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model. Appl Energy 241:229–244. https://doi.org/10.1016/j.apenergy.2019.03.044

Zhang Q, Li Z, Snowling S, Siam A, El-Dakhakhni W (2019) Predictive models for wastewater flow forecasting based on time series analysis and artificial neural network. Water Sci Technol 80(2):243–253. https://doi.org/10.2166/wst.2019.263

Zhang F, Li PC, Gao L, Liu YQ, Ren XY (2021) Application of autoregressive dynamic adaptive (ARDA) model in real-time wind power forecasting. Renew Energy 169:129–143. https://doi.org/10.1016/J.RENENE.2021.01.003

Zhang S, Chen Y, Xiao J, Zhang W, Feng R (2021) Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism. Renew Energy 174:688–704. https://doi.org/10.1016/J.RENENE.2021.04.091

Zhang W, Lin Z, Liu X (2022) Short-term offshore wind power forecasting—a hybrid model based on discrete wavelet transform (DWT), seasonal autoregressive integrated moving average (SARIMA), and deep-learning-based long short-term memory (LSTM). Renew Energy 185:611–628. https://doi.org/10.1016/J.RENENE.2021.12.100

Zhang X, Thearling K (1994) Non-linear time series prediction by systematic data exploration on a massively parallel computer

Zhao X, Wang S, Li T (2011) Review of evaluation criteria and main methods of wind power forecasting. Energy Procedia 12:761–769. https://doi.org/10.1016/j.egypro.2011.10.102

Zhou B, Ma X, Luo Y (2019) Wind power prediction based on LSTM networks and nonparametric Kernel density estimation. IEEE Access 7:165279–165292. https://doi.org/10.1109/ACCESS.2019.2952555

Zhou B, Liu C, Li J, Sun B, Yang J (2020) A hybrid method for ultrashort-term wind power prediction considering meteorological features and seasonal information. Math Probl Eng 2020:1–12. https://doi.org/10.1155/2020/1795486

Zhu L, Laptev N (2017) Deep and confident prediction for time series at Uber. IEEE Int Conf Data Min Work ICDMW 2017:103–110. https://doi.org/10.1109/ICDMW.2017.19

Zou W, Li C, Chen P (2019) An inter type-2 FCR algorithm based T-S fuzzy model for short-term wind power interval prediction. IEEE Trans Ind Inf 15(9):4934–4943. https://doi.org/10.1109/tii.2019.2910606

Authors and Affiliations

Sherry Garg[1] · **Rajalakshmi Krishnamurthi**[1]

✉ Sherry Garg
sherry.garg@jiit.ac.in

Rajalakshmi Krishnamurthi
k.rajalakshmi@jiit.ac.in

1 Department of Computer Science and Engineering, Jaypee Institute of Information Technology, Noida, India