Results in Engineering 26 (2025) 104597

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Designing empirical fourier decomposition reinforced with multiscale increment entropy and deep learning to forecast dry bulb air temperature

Mohammed Diykh [a,d], Mumtaz Ali [b,f,*], Abdulhaleem H. Labban [c], Ramendra Prasad [e], Mehdi Jamei [f], Shahab Abdulla [b], Aitazaz Ahsan Farooque [f,g]

[a] University of Thi-Qar, College of Education for Pure Science, Iraq
[b] UniSQ College, University of Southern Queensland, QLD, 4305, Australia
[c] Department of Meteorology, King Abdulaziz University, Jeddah, 21589, Saudi Arabia
[d] Technical Engineering College, Department of Cybersecurity, Al-Ayen Iraqi University, Thi-Qar 64001, Iraq
[e] Department of Science, School of Science and Technology, The University of Fiji, Saweni, Lautoka, Fiji
[f] Canadian Centre for Climate Change and Adaptation, Faculty of Sustainable Design Engineering, University of Prince Edward Island, St Peters, PE, Canada
[g] Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE, Canada

ARTICLE INFO

Keywords:
Air temperature
Dry bulb
Forecasting
Deep learning
MSIE
MEFD

ABSTRACT

Accurate prediction of dry bulb air temperature (DBTair) is significant to determine the state of humid air and supporting experts in the environmental sector. Traditional machine learning based approaches struggle to deliver accurate predictions when temperature is suddenly fluctuated during extreme weather conditions. This paper aims to design an intelligent model namely MEFD-MSIE-FCNN to forecast DBT_{air} which integrates multivariate empirical Fourier decomposition (MEFD), multiscale increment entropy (MSIE), and FCSM model that integrates a fully connected neural network FCNN with long short-term memory (LSTM) to forecast DBT_{air}. The multivariant time series of each predictor variable is passed through the MEFD to extract mutual features across multivariant time series and deliver multivariable-aligned modes. Then, the MSIE is extracted to form a feature final matrix to represent mutual information from multivariant time series. Finally, the features set is sent to the FCSM to forecast multistep ahead DBT_{air} using goodness-of-fit statistical metrics for two regions in Saudi Arabia. The proposed model showed highest accuracy for Jazan station (RMSE=2.120, MAE=2.912, RSE=0.123, ECC=0.971, WIA=0.981, CC=0.982), and Jeddah station (RMSE=2.131, MAE=2.921, RSE=0.113, ECC=0.969, WIA=0.979, CC=0.980). A comprehensive comparison is made against state-of-the art benchmarking models, concluding that there is a noticeable improvement in model's performance in terms of AME, ECC, CC, WIA, RMSE and correlation coefficient. The proposed FCSM can be helpful for many applications such as improving weather prediction, preventing climate risks, energy consumption, water resources management and agricultural industry. Additionally, the proposed model can support decision makers and industries in the environmental sector to make informed decisions to mitigate the effects of climate change.

List of acronyms

Wind prevailing direction	WPD
Wind mean speed	WMP
Wind maximum direction	WMD
Wind maximum speed	WMS
Wind maximum time (day)	WMTD
Pressure mean station level	PMSL
Pressure mean sea level	PMSLE
Relative humidity mean	RHM
Vapor pressure mean	VPM
Sky cover oktes mean	SCOM
Pressure maximum station level	PMASAL
Pressure maximum sea level	PMASEL
Air temperature maximum DB	$DBTmax_{air}$
Air temperature maximum WB	$WBTmax_{air}$
Relative humidity maximum	RHM
Pressure minimum station level	PMISL
Pressure minimum sea level	PMISLE
Air temperature minimum DB	$DBTmin_{air}$
Air temperature minimum WB	$WBTmin_{air}$
Relative humidity minimum	RHMIN

* UniSQ College, University of Southern Queensland, QLD, 4305, Australia

https://doi.org/10.1016/j.rineng.2025.104597
Received 13 November 2024; Received in revised form 28 February 2025; Accepted 6 March 2025
Available online 18 March 2025

Rainfall total RAINFT
Synops Hrs SHR
Main 4 syn M4SY
Syn Obsrvn SOB
Air temperature mean WB WBT_{air}
Air temperature mean DB DBT_{air}

1. Introduction

The changing climate due to anthropogenic greenhouse gas emissions (GHG) has been causing an increase in the frequency and intensity of disasters such as floods, droughts, storms, heatwaves, and wildfires with consequent sea level rises, saltwater intrusion, coastal erosion and biodiversity loss [1–3]. The GHG emissions essentially disrupt the natural energy balance, causing the global mean surface temperature to rise. The knowledge of ambient air temperature or dry bulb temperature (DBT_{air}) is crucial for many applications, such as short-term electrical load forecasting [4] and subsequent energy generation in the energy sector. The water resources management and agricultural industry modulates the evaporation, evapotranspiration, and soil moisture [5,6]. For the health and tourism sectors, extreme DBT_{air} magnitudes are a precursor to cardiovascular disease [7] and cerebrovascular mortality [8], while extremely low temperatures cause hypothermia [9]. The health indicators such as energy or heat index are derived from DBT_{air}. In addition, the wildfire risk is contingent upon the DBTair. Moreover, infrastructure planning and urban development depend on dry bulb temperature, as DBT_{air} is exceptionally relevant in determining the urban heat island effect [10,11]. Hence, for appropriate informed decision-making and mitigation actions, it is imperative to have an accurate DBT_{air} forecasting and early warning system to avert the looming issues. Designing a an accurate model for dry bulb air temperature (DBT) prediction plays a vital role in supporting decision-makers in various sectors such Climate monitoring, agricultural planning, energy consumption managemen.

Accurate analysing of climate data remains a challenging task, particularly in many developing countries that are most affected by temperature change [12,13]. In addition, several studies have showed there was a link between climate change and agriculture issues, particularly food production in different regions [14,15]. The forecasting of DBT_{air} is a relatively complex yet important field of study as DBT_{air} displays considerable spatial, temporal, and seasonal variability [16]. The time series of DBT_{air} is non-stationary and displays non-linear and stochastic characteristics, which the statistical modelling processes are unable to emulate these nonlinear dynamics aptly and incapable of achieving an acceptable level of forecasting accuracy [6,17]. Subsequently, Apaydin, Yumuş et al. [18] compared the performance of 12 different regression methods with two machine learning methods (ML) i.e., boosting-based Extreme gradient boosting regression (XGBoost) and Light gradient boosting machine (LightGBM) for one-day ahead maximum and minimum air temperature is forecasted. Their study found that the performance of ML models was better, indicating that ML data-driven models are suitable alternatives. Additionally, scholars have developed and studied several classical ML models for DBT_{air} forecasting, including gradient Boosting Tree (GBT), Random Forest (RF), Linear regression (LR) and artificial neural network (ANN) with various architectures such as multi-layered perceptron, radial basis function [6]. The ANN models were found to work better in this study. However, in a study by Mellit, Pavan et al. [17], the least square support vector machine (LSSVM) produced significantly better results than ANN architectures, including recurrent neural network, multi-layered perceptron, radial basis function and probabilistic neural network. In another study, while testing the forecasting capability of three ML models i.e., LSSVM, group method and data handling neural network (GMDHNN) and classification and regression trees (CART) for monthly DBT_{air} forecasting, Adnan, Liang et al. [5] found that LSSVM model outperformed the other models (GHMDN and CART).

With advancements in the ML and artificial intelligence modelling approaches over recent years, the deep learning (DL) has gained popularity mainly for classification problems and computer vision applications. It is well established that the DL approaches can recognize complex patterns and uncover highly nonlinear relations from complex datasets [19], and DBT_{air} is no exception. Different hybrid models with ML and DL approaches have been developed to improve the model performances. Zhou, Wang et al. [20] developed a hybrid ANN model with a powerful hetaeristic Honey Badger Algorithm (HBA-ANN) and compared the performance of classical ANN and Gene Expression Programming (GEP) to forecast DBT_{air} in the hottest place (Furnace Creek, Death Valley, USA) and the coldest place (Vostok, Antartica). In all the forecast horizons (one to three months ahead), the HBA-ANN model outperformed the comparative ones [20]. Regarding DL hybrid mode, Thi Kieu Tran, Lee et al. [21] applied the DL-based model long short-term memory (LSTM) in a 15-day ahead DBT_{air} forecasting in a long-term forecasting horizon approach. The performance of LSTM was compared with that of ANN, a recurrent neural network (RNN), while the genetic algorithm (GA) was used for meta-learning to select the optimum architecture of the networks. The results demonstrated that the hybrid model of an LSTM (i.e., GA-LSTM) outperformed the other models. Two deep-learning methods i.e., a convolutional neural network (CNN) and long short-term memory (LSTM) has been integrated into a hybrid modelling network (CNN–LSTM) and found to outperform other models. The CNN reduced the dimensionality of the time-series data, while LSTM captured the long-term memory of the massive temperature time-series data. Roy [22] found a better performance of CNN-LSTM in comparison to standalone LSTM and standalone CNN in forecasting daily DBT_{air} for John F. Kennedy International Airport, NY, while Hou, Wang et al. [23] demonstrated a better hybrid CNN-LSTM model performance in hourly DBT_{air} predictions in Yinchuan, China in comparison to their standalone counterparts. The main essence of all these hybrid DL learning approaches is integrating a data pre-processing technique that improves the model performance.

However, in order to extract and unveil pertinent embedded features within the complex DBT_{air} time series, the data decomposition techniques or the multiresolution analysis (MRA) have been less popular with DBT_{air} forecasting. Yet, the MRA pre-processing techniques have been applied in other applications such as significant wave height [24,25], non-invasive glucose detection [26], hydrological [3,27] and energy [28] with excellent outcomes. The commonly used univariate MRA includes empirical mode decomposition (EMD) [29,25], ensemble-EMD (EEMD) [30,27], in Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) [24,31,26] and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [32]. Since DBTair forecasting requires multivariate inputs, the common multivariate MRA included multivariate-EMD (MEMD) [33–35]. This MEMD has been used in energy [28], crude oil price forecasting [36] and hydrological [37] applications. In this study, the application of a new MRA utility named the Empirical Fourier decomposition (EFD) performs signal processing in two steps: i) an improved segmentation technique and ii) the construction of a zero-phase filter bank [38]. The benefit is that the first phase eliminates any possible trivial residual in the first decomposed component, while the second phase, with the use of a zero-phase filter bank, facilitates accurate signal analysis of closely spaced modes and eliminates the transition phases, overcoming the mode mixing issue [38]. Empirical results demonstrated that EFD yields decomposition results with high accuracy and consistency. EFD also yielded accurate time–frequency representations (TFRs) for non-stationary signals and is computationally efficient [38]. In this study the EFD is applied in a multivariate approach, whereby the MEFD can demarcate the multivariate input time series into sub-series called the multiscale increment entropy (MSIE).

This paper proposes a novel hybrid model called FCSM to predict one day, one week, and on month ahead DBTair using historical multivariate time series data. A multivariate empirical Fourier decomposition

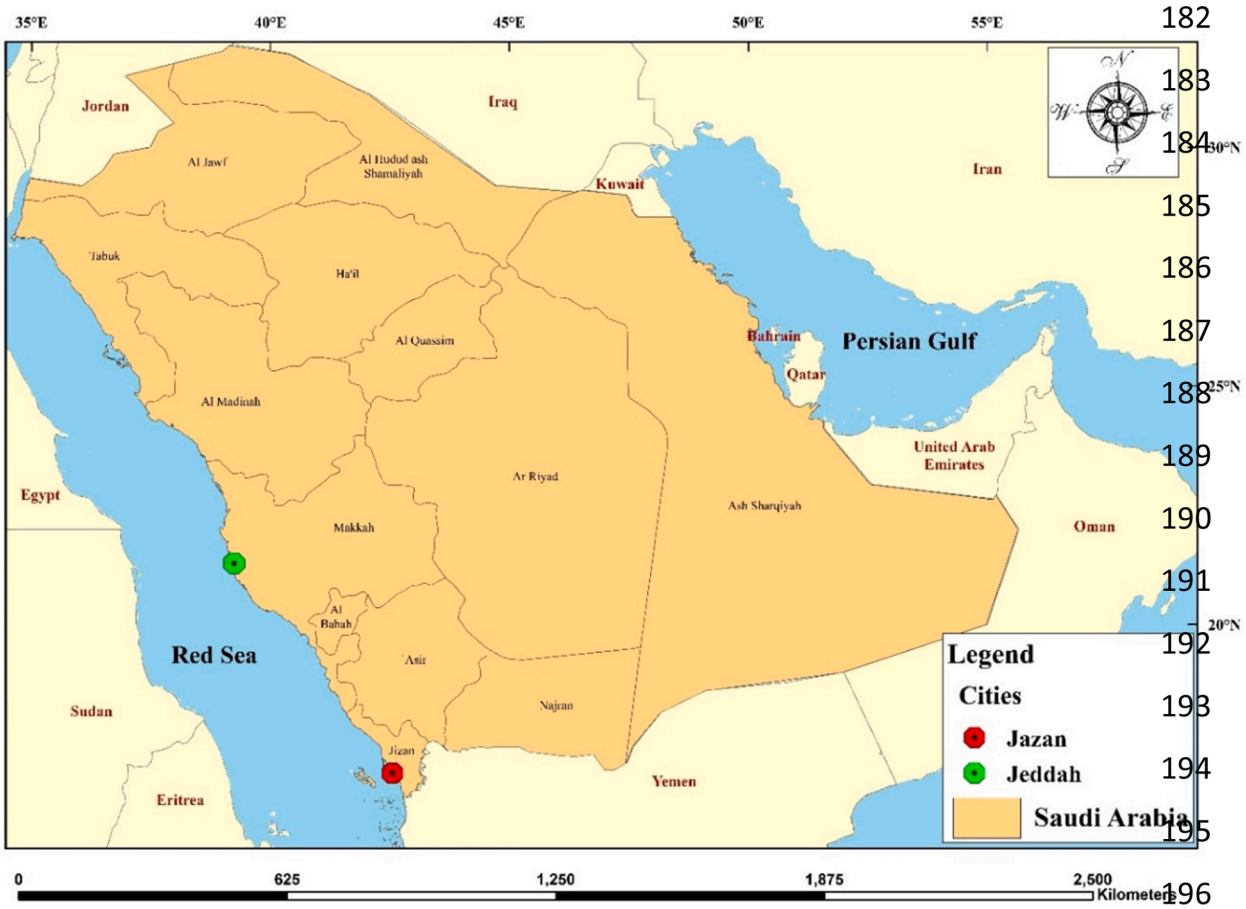


Fig. 1. Map of the study locations in Saudi Arabia.

Table 1
Summary and descriptive statistics of the data.

Station	Jazan [16.8892°N; 42.5511°E]						Jeddah [21.2854°N; 39.2376°E]					
Input	Min.	Max.	Mean	Std.	Skew.	Kurt.	Min.	Max.	Mean	Std.	Skew.	Kurt.
Wind mean speed	0.00	19.00	6.10	1.62	0.62	1.75	0.00	72.00	6.91	2.59	3.46	70.14
Win max. direction	0.00	350.00	27.15	26.14	8.51	75.81	0.00	360.00	35.01	43.09	5.54	30.31
Wind max. speed	6.00	75.00	16.86	4.74	2.63	15.29	4.00	50.00	16.78	4.05	0.96	2.74
Wind max time (day)	0.00	90.00	11.62	2.68	4.12	130.83	0.00	29.00	11.30	2.49	-0.01	5.56
Pressure mean station level	910.20	1018.80	1006.93	8.27	-46.31	3238.67	910.20	1106.50	1006.40	8.40	-45.08	3172.57
Pressure mean sea level	0.00	1053.30	1004.61	59.78	-16.49	265.63	0.00	1022.40	1005.04	59.82	-16.48	265.44
Relative humidity mean	18.00	999.00	67.66	13.33	51.71	3616.76	18.00	97.00	59.73	10.37	-0.47	0.22
Vapor pressure mean	0.00	65.00	28.89	4.37	-1.48	8.60	0.00	293.00	22.75	6.77	9.57	384.61
Sky cover oktes mean	0.00	91.00	1.97	2.28	17.36	563.29	0.00	57.00	1.00	1.59	12.21	367.99
Pressure max station level	100.40	1021.60	1008.32	22.53	-33.78	1177.35	101.50	1022.20	1007.73	16.02	-41.17	1851.81
Pressure max sea level	0.00	4008.40	1006.04	75.08	-0.24	496.44	0.00	1126.80	1005.87	64.02	-15.25	227.79
Air temperature max DB	3.00	46.30	35.10	3.54	-1.57	9.71	3.00	52.00	34.56	4.58	-0.76	2.88
Air temperature max WB	-0.80	85.00	27.03	3.58	-3.50	38.10	-0.80	40.00	23.31	5.92	-2.41	7.00
Relative humidity max	0.00	100.00	79.26	7.06	-0.86	5.96	8.00	100.00	80.51	10.14	-0.90	2.05
Pressure min station level	800.00	1019.80	1005.20	8.60	-44.31	3080.21	906.20	9997.10	1006.25	111.48	79.58	6339.17
Pressure min sea level	0.00	1017.40	1002.28	65.38	-14.87	216.53	0.00	1060.10	1003.16	63.25	-15.41	232.36
Air temperature min DB	-8.00	37.80	26.07	3.76	-1.46	9.24	-8.00	42.00	22.81	4.14	-0.89	3.58
Air temperature min WB	-20.00	68.00	22.73	3.60	-2.64	27.05	-6.50	61.00	18.53	5.45	-1.53	4.01
Relative humidity min	6.00	601.00	54.21	11.17	17.49	871.02	3.00	98.00	37.37	12.06	-0.10	-0.31
Rainfall total	0.00	777.70	0.58	13.90	53.30	2962.88	0.00	777.70	0.27	9.82	75.80	5934.04
Synops Hrs	10.00	24.00	23.97	0.50	-19.88	1673.36	10.00	24.00	23.97	0.51	-19.22	1546.18
Main 4 syn	0.00	8.00	7.94	0.66	-11.38	165.13	0.00	8.00	7.99	0.21	-21.56	4312.61
Syn Obsrvn	0.00	4.00	3.96	0.37	-9.28	172.49	1.00	4.00	3.99	0.18	-14.92	2580.43
Air temperature mean WB	-2.00	72.00	25.04	3.26	-3.49	33.10	-2.00	32.10	21.13	5.51	-2.17	5.93
Air temperature mean DB	-0.30	39.10	30.22	3.31	-1.78	11.52	-0.30	39.00	28.29	4.02	-0.99	3.31

(MEFD) is performed on the multivariate input signals to decompose the data into several modes. Then, from each mode, the MSI is extracted to form the final feature array, which yields multiscale increment entropy (MSIE) components. These sub-series are channelled as inputs to the long short-term memory LSTM-based fully connected neural network (FCNN) to predict DBTair. The performance of hybrid proposed model

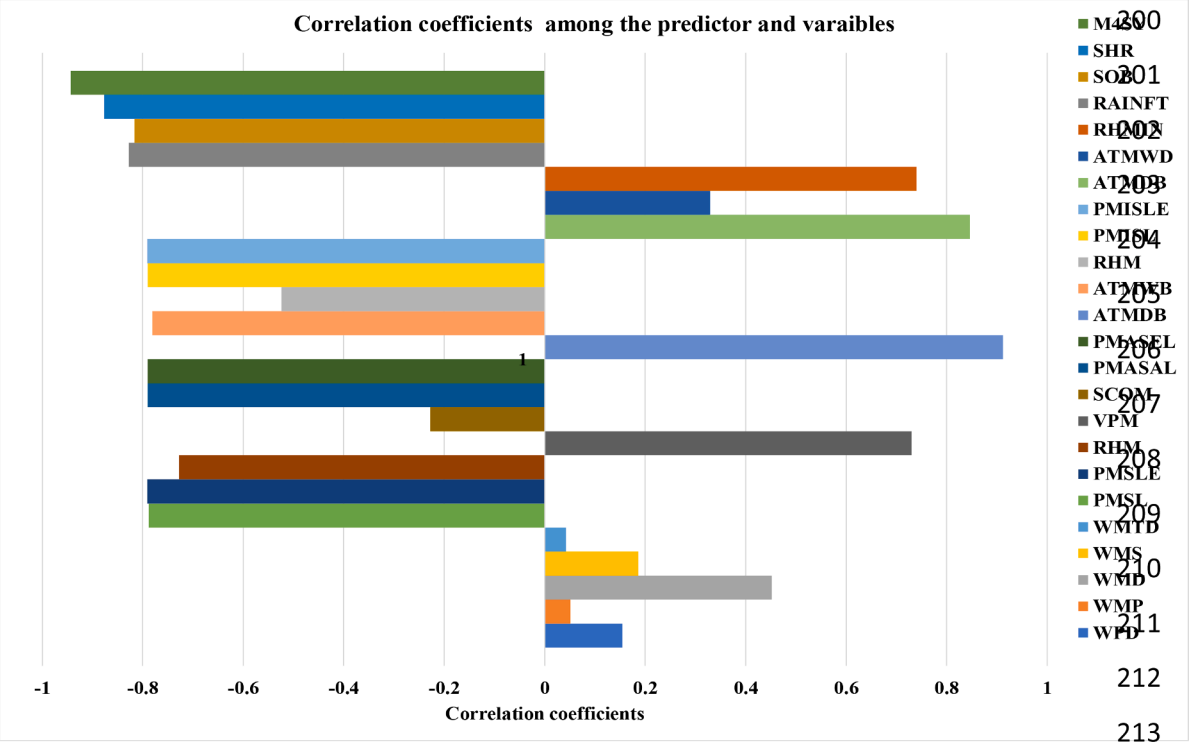


Fig. 2. Correlation coefficients of input variables with DBT_{air}.

FCSM is compared with seven other benchmark models. The obtained results showed that the proposed model provides important implications for the environmental and industrial sectors, as accurate prediction of DBTair can support decision makers in optimising energy efficiency, improve agricultural strategies.

2. Study area and data description

The city of Jazan is situated in the southwest of Saudi Arabia (approximately 16.8892°N latitude and 42.5511°E longitude), experiencing a tropical climate. The temperatures are higher during the year, with moderately consistent humidity levels. The summers in Jazan City are hot and humid, with temperatures frequently rising above 40°C while winters experience milder climates with temperatures hardly plummeting below 20°C. The rainfall is also significant during the monsoon season (June to September) in Jazan, which brings greenery and agricultural production. The topography of Jazan is diverse, comprising coastal plains, mountains, and valleys. The sandy beaches and low-lying areas are situated along the coast of the Red Sea. The beautiful Sarawat Mountains run parallel to the coast, influencing the weather by triggering different temperatures and precipitation levels. The strategic location of Jazan makes the city a trading hub and a doorway to the Arabian Peninsula.

Jeddah with coordinates 21.2854°N latitude and 39.2376°E longitude position the city on the western coastline of Saudi Arabia. The city experiences a scorching desert climate with extremely hot summers where temperatures typically soar above 40°C, and the winters are mild with temperatures up to 15°C. Jeddah city receives a small amount of rainfall, which occasionally occurs in the winter, while the summer experiences high humidity levels because of its coastal position. The city faces flash flooding during recent rare heavy rainfall events due to neighbouring wadis (dry riverbeds) and small valleys. The city of Jeddah is a major port and commercial hub and become the centre for trade and commerce, linking the Kingdom to global markets. Fig. 1 shows the study area and the locations of the Jazan and Jeddah stations.

The dataset contains the different input variables recorded at a daily time interval for Jazan and Jeddah stations during 1978 to 2013. The specific input variables are wind mean speed, wind maximum direction, wind maximum speed, wind maximum time (in a day), pressure mean station level, pressure mean sea level, relative humidity mean, vapour pressure mean, sky cover oktas mean, pressure max station level, pressure max sea level, air temperature max DB, air temperature max WB, relative humidity max, pressure min station level, pressure min sea level, air temperature min DB, air temperature min WB, relative humidity min, rainfall total, Synops Hrs, Main 4 syn, Syn Obsrvn, and air temperature mean WB. The target variable is the Air temperature mean DB (i.e., DBT_{air}), which exhibits significant spatial, temporal, and seasonal variability. Table 1 presents a summary, and some descriptive analysis of the data used in this paper. First, a statistical approach was employed to reconstruct missing data points. Fig. 2 depicts the correlation between the DBT_{air} and input predictor variables, it is noticed that some variables showed a higher correlation with the predictor while other variables showed a negative correlation with the predictor. For example, the AIR_TEMPERATURE_MAX_DB (i.e., $DBTmax_{air}$) gained the maximum correlation coefficient, however, the MAIN_4_SYN gained the lowest correlation coefficient. To predict the air temperature dry bulb, the highest correlations variables are considered in our study. In addition, all variables are also tested in term of MSIE to analyse their relationship with the DBT_{air}.

3. Methods

3.1. Multivariate empirical Fourier decomposition

The Multivariate empirical Fourier decomposition (MEFD) is a modified version of empirical Fourier decomposition (EFD) to remove the mode alignment issue in signals and to improve multivariate signals [39,40,38]. One of the main issues of EFD is that EFD adopts an adaptive approach to decompose signals into modes which do not reveal the mutual features among multi-channel signals [41,42]. To address issues associated with the EFD, the MEDF uses the optimal boundary selection and filter bank to reduce noise and improve signal characteristics. The

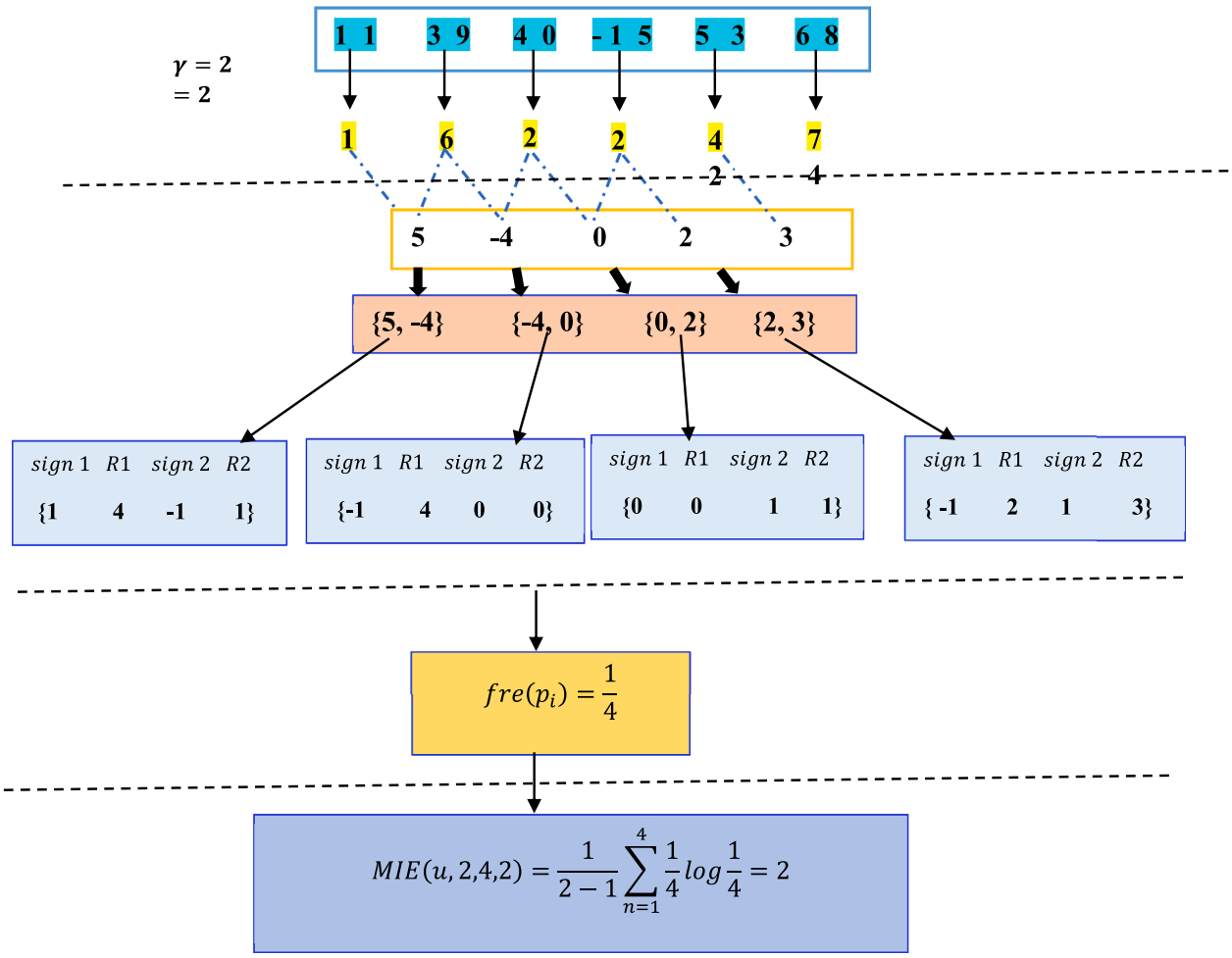


Fig. 3. Diagram of multiscale increment entropy.

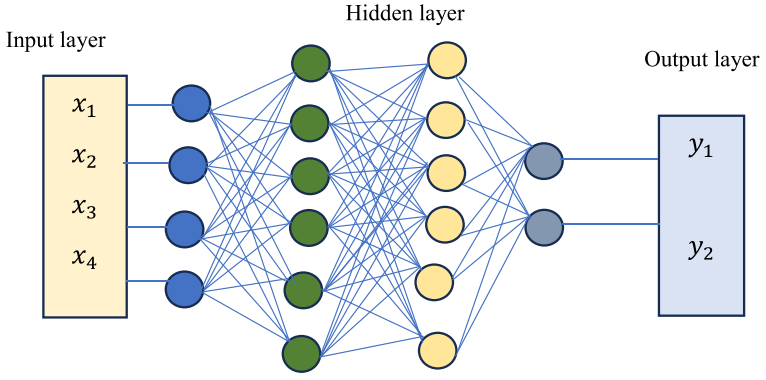
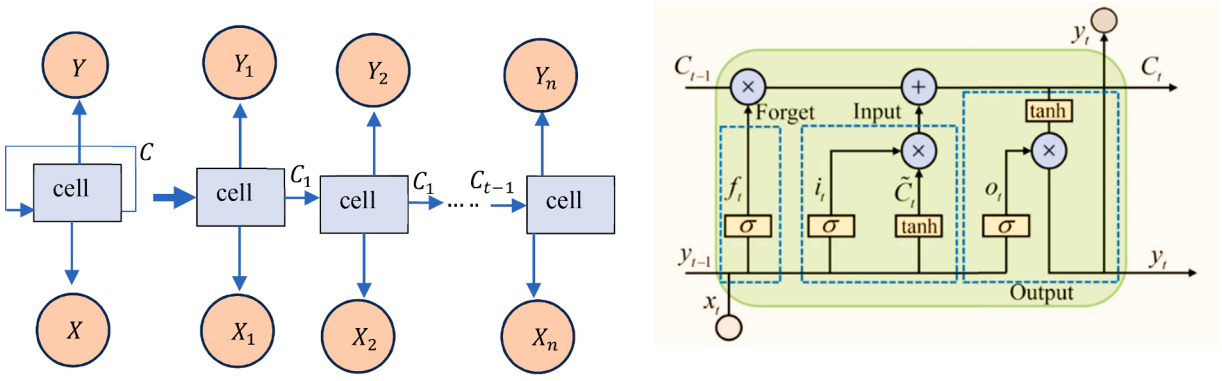


Fig. 4. Schematic diagram of the FCNN model.



(a) Jeddah station **(b) Jazan Station**

Fig. 5. Structural diagram of LSTM model.

Table 2
Performace statistical metrics used in this paper.

Metrics	Equations
Mean absolute error (MAE)	$MAE = mean\left(\sum_{i\in Atesting_{set}}\left\|T^i_{air} - \overline{T^i_{air}}\right\|\right)$ (24)
Relative percentage error (RPE)	$RSE = \sqrt{\frac{\sum_{i\in testing\_set}\left(T^i_{air} - \overline{T^i_{air}}\right)^2}{\sum_{i\in testing\_set}\left(T^i_{air} - mean(T^i_{air})\right)^2}}$ (25)
Correlation coefficient (CORR)	$CORR = \frac{1}{\left\|SPI_{\text{test}}\right\|}\sum_{i\in t\in sting\_set}\frac{\sum_i\left(T^i_{\text{air}} - mean(T^i_{\text{air}})\right)\left(\overline{T^i_{\text{air}}} - mean\overline{T^i_{\text{air}}}\right)}{\sum_i\left(T^i_{\text{air}} - mean(T^i_{\text{air}})\right)^2\sum_{i=1}^n\left(\overline{T^i_{\text{air}}} - mean\overline{T^i_{\text{air}}}\right)}^2$ (26)
Willmott's index (WI)	$WI = 1 - \left[\frac{\sum^N_{i\in testing_{set}}\left(\overline{T^i_{air}} - T^i_{air}\right)^2}{\sum^N_{i\in testing_{set}}\left(\left\|\overline{T^i_{air}} - T^i_{air}\right\| + \left\|\overline{T^i_{air}} - T^i_{air}\right\|\right)^2}\right]$ (27)
Root square error (RSE)	$NSE = 1 - \left[\frac{\sum^N_{i\in testing_{set}}\left(\overline{T^i_{air}} - T^i_{air}\right)^2}{\sum^N_{i\in testing_{set}}\left(\overline{T^i_{air}} - T^i_{air}\right)^2}\right]$ (28)
Legates & McCabe's (LM)	$LM = 1 - \left[\frac{\sum^N_{i\in testing_{set}}\left\|\overline{T^i_{air}} - T^i_{air}\right\|}{\sum^N_{i\in testing_{set}}\left\|\overline{T^i_{air}} - T^i_{air}\right\|}\right]$ (29)
Relative percentage error (RPE)	$RPE = \frac{1}{N}\sum^N_{\in testing_set}\left\|\frac{\overline{T^i_{air}} - T^i_{air}}{T^i_{air}}\right\|$X100 (30)

Where T^i_{air} denotes the actual values and $\overline{T^i_{air}}$ referees to the predicted air temperature values. A low MAE and RRSE values indicate a high performance.

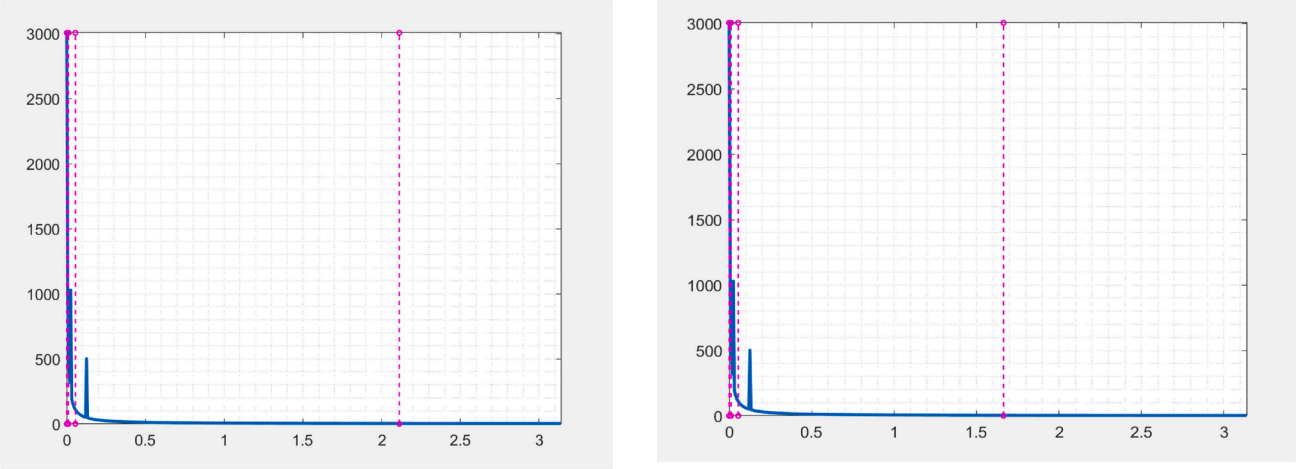


Fig. 6. Detect the boundaries of a sample of data for two stations.

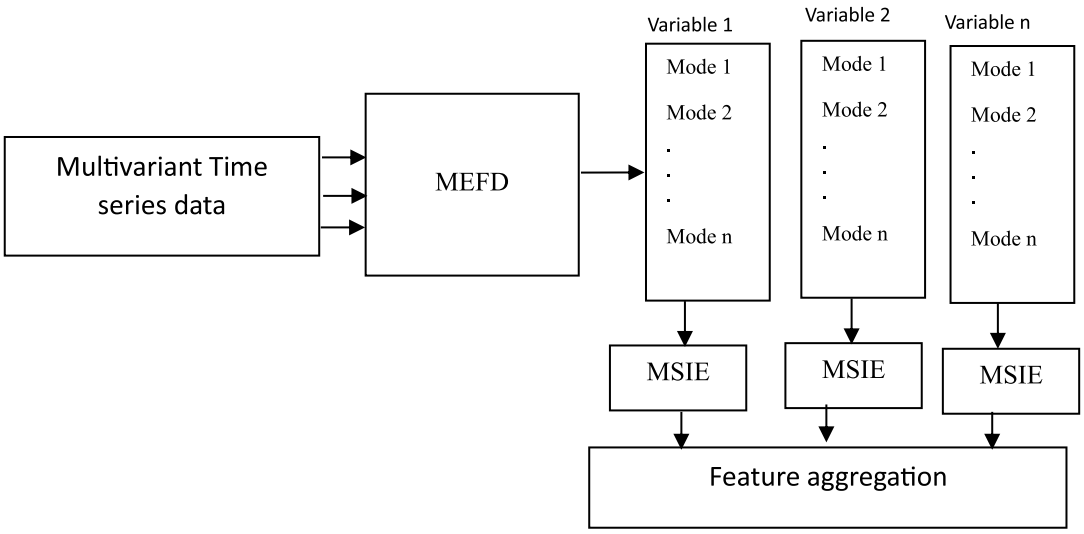


Fig. 7. Diagram of MEFD based MSIE.

Algorithm 1

Input: multipartite signal $X(t) \in R^{mxn}$ where m the number of channel and n refers to the length of signal.
Output: modes $\overline{X}(t)\ R^{mXlXn}$ where l is the number of nodes.
1. **Apply** FFT to obtain the spectrum for all channels of the input signal.
2. **Calculate** the mean spectrum magnitude $F(f)$ from all channels.
3. **Find** the maximum and minimum values of the spectrum.
4. **Sort** the frequencies in ascending order for each channel.
5. **Apply** zero filter bank to filter and create the signal.
6. **Use** inverse FFT to convert the filtered signal into time domain.

MEFD includes three phases as follows:

1. Mean spectrum estimation: suppose a multivariate time series $X(t)$, where

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ . \\ . \\ . \\ x_n(t) \end{bmatrix} \tag{1}$$

Where n is the number of variables (channels) in time series; to align the time series X across different variables, the mean spectrum magnitude is adopted to identify common boundaries across different variables. The mean spectrum of multi-channel data is calculated as follows:

$$F(x) = \frac{1}{m}\sum_{n=1}^{m} |F_n(x) \tag{2}$$

Where $F(x)$ is the Fourier spectrum of channel n, and n *from* 1 *to* m.

2. Improve Spectrum Segmentation Approach (ISSA): The ISSA is based on the lowest minima method. The ISSA divides $[0,\ \pi]$ into N different frequencies and the value of w_0 and w_n are not fixed at 0 and π. The values of w_0 and w_n are determined using an adaptive soring process, which involves Fourier spectrum magnitude. The Fourier spectrum magnitude at w_0 and w_n are identified and grouped in a sequence. Then, all magnitudes in the sequence are sorted in descending order. It is referred to the first largest values in the sorted sequence as $\{\mathscr{M}_1, \mathscr{M}_2, \ldots, \mathscr{M}_n\}$. The boundary w_n of each segment is defined as

$$w_n = \begin{cases} argmin(F_n(x)),\ if\ o \le n \le m\ and\ \mathscr{M}_1 \ne \mathscr{M}_{1+1} \\ \mathscr{M}_n\ \ if\ 0 \le n \le m\ and\ \mathscr{M}_1 = \mathscr{M}_{1+1} \end{cases} \tag{3}$$

Where $\mathscr{M}_0$ is defined as 0 and $\mathscr{M}_{1+1}$ as π.

3. Zero phase filter bank (ZPFB): the ZPFB is a bandpass filter. It functions in each frequency band, and it has no transition band with cut-off frequencies between $\mathscr{M}_{n-1}$, and $\mathscr{M}_n$. The ZPFB keeps each segment's most important Fourier spectrum and removes the components beyond the segment. Fourier transform is employed to gain the spectrum signals $\overline{F}(w)$ of signals $X(t)$. The zero-phase filter bank $\mu(t)$ is defined as:

$$\mu(w) = \begin{cases} 1, & \text{if } w_n - 1 \le |w| \le w_n \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Where $1 \le n \le m$, w_n is determined by Eq.3, the filter signal $\overline{F}_n(w)$ which is corresponding to $\mu(w)$ is calculated using the following question.

$$\overline{F}(w) = \mu(w)\overline{F}(w) = \begin{cases} \overline{F}(w),\ if\ w_n - 1 \le |w| \le w_n \\ 0\ otherwise \end{cases} \tag{5}$$

The inverse FT of $\overline{F}_n(w)$ is constructed by the time domain components $f_n(w)$. The reconstructed $\overline{F}_n(w)$ signals are the sum of all components of the decomposed signal as follows:

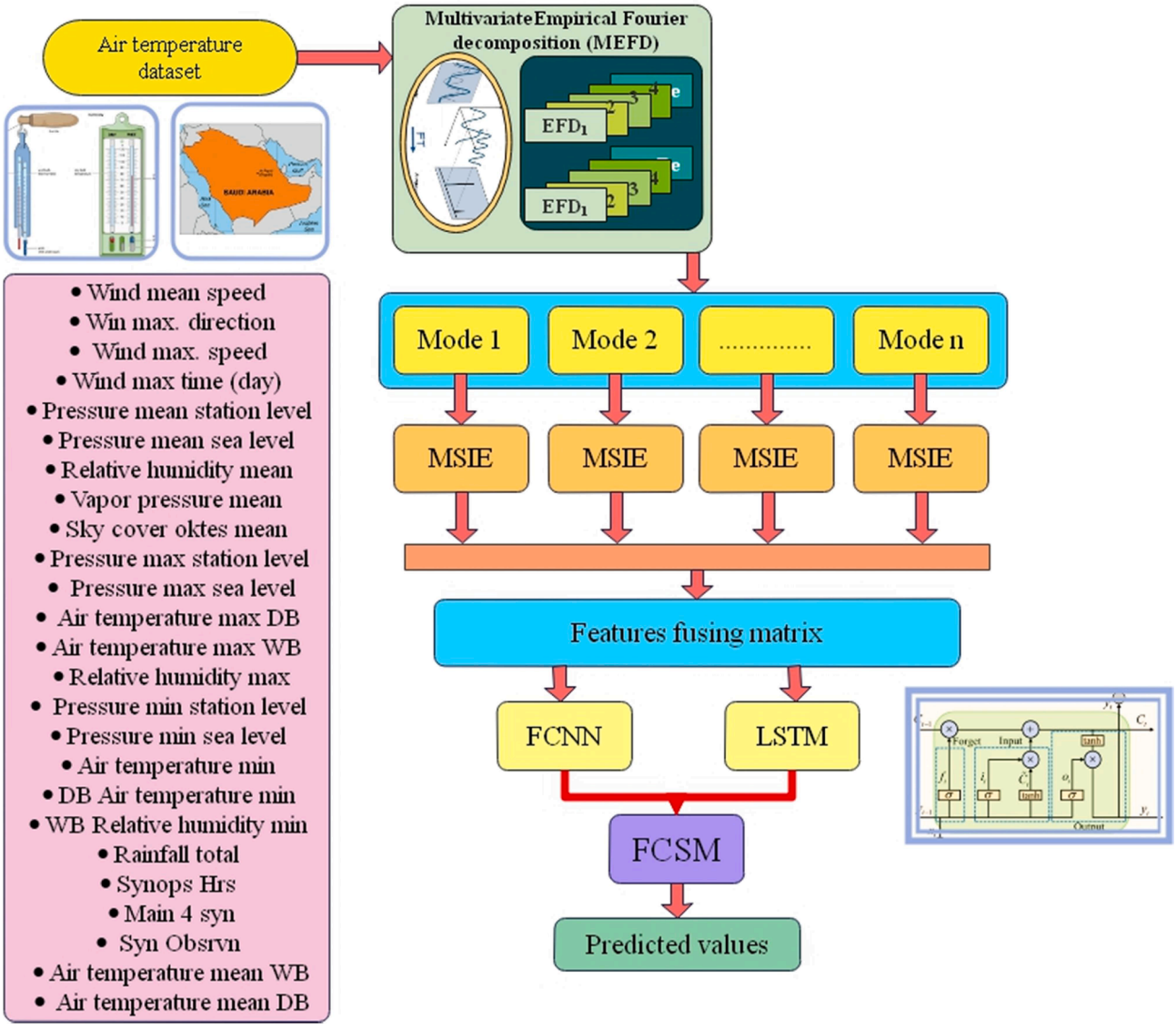


Fig. 8. The framework of the proposed model for DBT_{air} prediction.

Table 3
The turning parameters of the benchmark models.

Model	Parameters
MLP	Number of hidden neurons [10,25,50,100], learning rate [10^{-5}, 10^{-3}, 10^{-3}], number of iterations [1000, 1500].
XG_Boost	Number of gradient boosted tree [100,500,1000], maximum depth of tree [4,5,8], learning tree [0.5,0.75, 1.00].
Lasso_Reg	Maximum number of iterations [1000,2000, 4000, 5000]
SVR	Kernal width [10^{-3}, 10^{-2}, …, 10^{3}]
Elastic_Net	Maximum number of iterations [1000,2000, 4000, 5000] Penalisation weight [0.10, 0.50, 0.70, 0.90, 0.99, 1.00]
RF	Number of trees [100, 500, 1000], Depth of tree [3,5,6], Number of depths of tree [2,5]
FCSM	Batch size 128, 64, epoch=300, 200, learning rate [10^{-5}, 10^{-3},]

$$\overline{F}_n(w) = \sum_{n=1}^{m} f_n(w) \tag{6}$$

as a result, the multivariate signals can be represented as

$$f(w) = \begin{bmatrix} f_{11}(w) & f_{12}(w) & . & . & f_{1n}(w) \\ f_{11}(w) & f_{12}(w) & . & . & f_{21}(w) \\ . & . & . & . & . \\ . & . & . & . & . \\ f_{m1}(w) & f_{m1}(w) & . & . & f_{mn}(w) \end{bmatrix} \tag{7}$$

Where $n \in [1, n]$ is the number of decomposed nodes.

3.2. Multiscale increment entropy (MSIE)

The MSIE takes advantage of multiscale entropy and increment entropy [43–46]. For one-dimensional EEG time series, the main steps of

Table 4
Prediction results using MSIE against other entropy features with the FCSM model.

One day ahead												
Jazan station							Jeddah station					
Entropy	RMSE	MAE	RSE	ECC	WIA	CC	RMSE	MAE	RSE	ECC	WIA	CC
MSDE	2.78	3.121	0.142	0.884	0.873	0.874	2.73	3.151	0.182	0.883	0.872	0.871
RCMDE	3.010	4.983	0.172	0.881	0.891	0.884	3.120	4.991	0.175	0.887	0.898	0.881
RCMPE	4.100	4.971	0.192	0.773	0.782	0.712	4.210	4.984	0.197	0.784	0.761	0.700
MSE	2.861	4.802	0.199	0.801	0.861	0.87	2.871	4.824	0.201	0.854	0.851	0.852
MSIE	**2.120**	**2.912**	**0.123**	**0.971**	**0.981**	**0.982**	**2.131**	**2.921**	**0.113**	**0.969**	**0.979**	**0.980**
Seven days ahead												
MSDE	2.891	3.323	0.193	0.823	0.812	0.813	2.894	3.354	0.197	0.821	0.810	0.811
RCMDE	3.312	4.994	0.193	0.821	0.824	0.816	3.324	5.001	0.198	0.811	0.809	0.810
RCMPE	4.382	4.993	0.199	0.711	0.714	0.704	4.395	5.013	0.299	0.693	0.686	0.693
MSE	2.892	4.875	0.251	0.792	0.813	0.814	2.991	4.899	0.263	0.742	0.806	0.802
MSIE	**2.340**	**2.998**	**0.125**	**0.932**	**0.942**	**0.943**	**2.521**	**3.192**	**0.176**	**0.924**	**0.927**	**0.910**
One-month ahead												
MSDE	3.211	3.642	0.272	0.763	0.793	0.796	3.541	3.921	0.274	0.772	0.781	0.781
RCMDE	3.616	5.021	0.287	0.781	0.783	0.782	3.675	5.324	0.292	0.774	0.776	0.771
RCMPE	4.794	5.232	0.217	0.677	0.673	0.674	4.892	5.731	0.253	0.661	0.668	0.663
MSE	3.103	5.121	0.283	0.713	0.785	0.785	3.103	5.425	0.294	0.741	0.777	0.771
MSIE	**2.821**	**3.211**	**0.178**	**0.912**	**0.921**	**0.910**	**2.876**	**3.312**	**0.189**	**0.901**	**0.891**	**0.890**

MSIE are shown in Fig. 3.

The MSIE is computed based on two steps: the coarse-grained time series and increment entropy, which involves five steps. The details of MSIE calculation are summarised as:

1. Apply Eq. (8) to produce a consecutive coarse-grained time series $\left\{X_j^t,\ 1,\ \leq j \leq N\right\}$.

$$X_{kj}^t = \frac{1}{t}\sum_{i=(j-1)t+k}^{k+jt-1} u_i, 1 \leq j \leq \left[\frac{l}{t}\right] = N, 1 \leq k \leq t \tag{8}$$

2. Calculate the increment entropy value for a consecutive coarse-grained time series according to the following steps:

I. Create an increment time series $IV = \{iv(j),\ 1 \leq j \leq N-1\}$ where $IV(j) = x_{j+1}^{(t)} - x_j^{(t)}$.

II. Divide the IV into $N-m$ increment sets, $IV(k) = [iv(k),\ iv(k+1),\ \ldots,\ vi(k+m-1)],\ 1 \leq k \leq N-m$. Where m refers to the number of embedded dimensions.

III. Transfer each element of the increment vector into a word containing two elements: *size* (l), *sign* (s). The *sign* refers to the direction of volatility between any neighbour elements. The *sign* takes the values $1, 0,\ or\ -1$ where 1 indicates an increase, 0 refers no change, and -1 indicates a decline. The *size* (l) represents the magnitude of variation between these neighbour elements. The magnitude is dependent on the resolution parameter R. Then, each vector is transferred into a pattern vector of size $2m$. With R and m, each increment vector will have $(2R+1)^m$ patterns. As a result, the $sign(k)$ for each element $iv(k)$ in the increment vector, IV is calculated as:

$$Sk = sign(v(k)) \tag{9}$$

and the size l of $iv(k)$ is calculated as

$$l(k) = \begin{cases} 0,\ step = 0 \\ min\left(I, \left\lceil \frac{iv(k)XI}{step} \right\rceil\right),\ step \neq 0 \end{cases} \tag{10}$$

where the step is the standard deviation of the time series.

IV. Calculates the frequency of each unique patterns using the following equation.

$$fre(p_i) = \frac{F(p_i)}{N-m} \tag{11}$$

Where, p_i denotes the i unique pattern, $F(p_i)$, the total number of instances of the p_i pattern vector.

V. define the MSIE value at scale γ as follows:

$$MSIE = (u, m, l, y) = incren(x^y, m, l) \tag{12}$$

$$= -\frac{1}{m-1}\sum n = 1^{(2l+1)^m} fre(p_i)\log fre(p_i)$$

Where $m-1$ is normalisation factor.

3.3. Fully connected neural network model (FCNN)

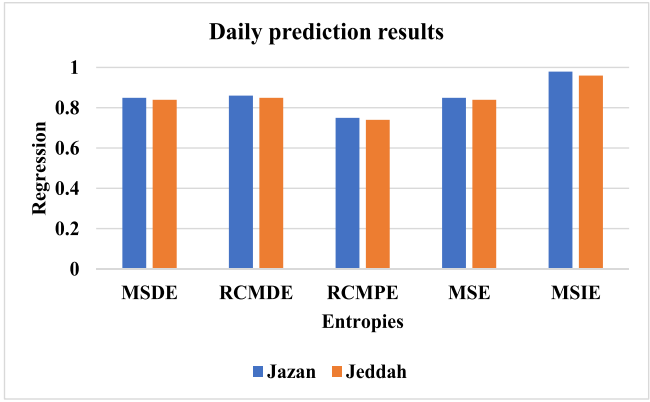
FCNN emphasises the depth of networks compared to traditional neural network algorithms [47]. The FCNN involves several layers: an input, hidden, and output layer. The model's depth is the sum of all hidden and output layers [48,45]. Fig. 5 depicts an example of the FCNN containing two hidden layers. It can be noticed that neurons in a layer are not linked, while neurons between layers are linked together [49]. The neurons accept the output from previous neurons as inputs. Then, neurons manipulate those signals before passing them to the subsequent neurons which are connected to them. The following equations describe how the calculation process of each neuron in each layer.

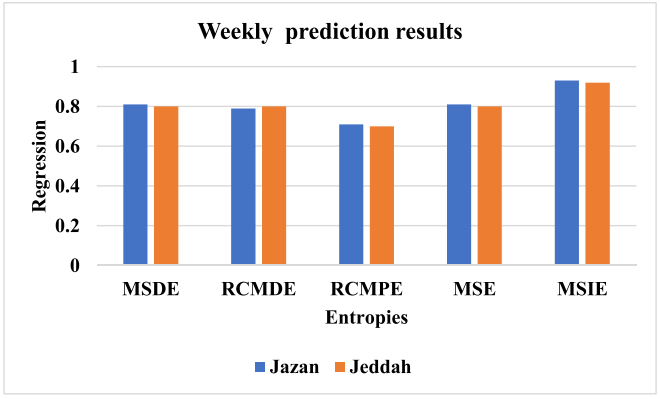
$$\vartheta_1 = f(w_1 x + b_1) \tag{13}$$

$$\vartheta_i = f(w_i \vartheta_{i-1} + b_1),\ 1 < i < b \tag{14}$$

$$y = f(w_d \vartheta_d + b_d) \tag{15}$$

Where f is an activation function which is determined before the training phase, x refers to the input value, y is the output value, b denotes the depth of the model, ϑ_i is the output of i-th layer, w_1 is the wights, $b_1$is biases. Fig. 4 show the schematic diagram of the FCNN model.





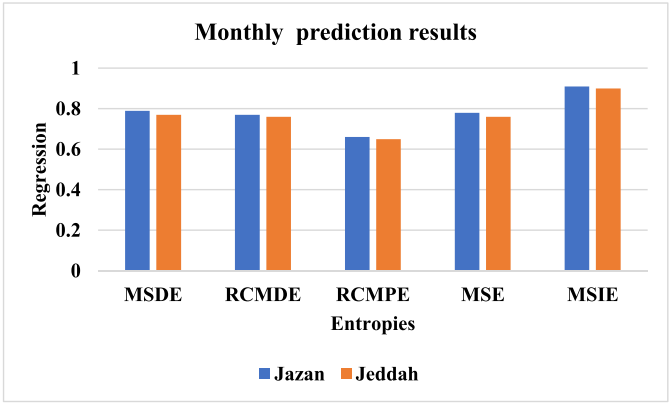


Fig. 9. prediction results based on entropy features using regression.

3.4. Long short-term memory model (LSTM)

The LSTM model is an improved version of RNN [50]. It has been designed to overcome the disappearance and gradient explosion issues associated with the RNN model [50]. The RNN model receives the input data, passes the data through a series of procedures and performs the same method on each element [51]. Fig. 2 depicts the RNN model. In Fig.6, the input data is denoted by $X = \{x_1, x_2, \ldots, x_n\}$, and the output by $Y = \{y_1, y_2, \ldots, y_n\}$, and $C = \{c_1, c_2, \ldots, c_t\}$ are the cells memory. The CNN mode uses the input data x_n, and the cell memory of previous $c_{t-1}$to predict the output y_n . The LSM model reconstructs the RNN model by employing three gates named input gate, forget gate, output gate. Fig. 5 depicts the LSTM model. in literature, we found that the LSTM cell structure have different descriptions, in this paper we choose the simplest one to explain its principles clearly. The activation function σ and *tanh* are calculated using Eq.4 and 5.

$$\sigma = \frac{1}{1+e^{e-t}},\ t \in (-\infty, \infty) \tag{16}$$

$$\tanh(t) = \frac{1-e^{-2t}}{1+e^{e-t}},\ t \in (-\infty, \infty) \tag{17}$$

The input and forget gate are employed to filter the previous cell information by which the LSM mode identifies the data that should be removed or added. This process is carried out using the following equations.

$$c_t = F_t.c_{t-1} + i_t.\bar{c}_t \tag{18}$$

$$F_t = \sigma(we_F.[y_{t-1},\ x_t] + b_F) \tag{19}$$

$$i_t = \sigma(we_F.[y_{t-1},\ x_t] + b_i) \tag{20}$$

$$\bar{c}_t = \tanh(we_c.[y_{t-1},\ x_t] + b_c) \tag{21}$$

Where *we* is the weight matrix, $\bar{c}_t$ refers to the new data that the cell should be remembered. The output gate is then employed to combine the cell memory c_t and the current input data to produce the output based on the following equation.

$$ou_t = \sigma(we_{ou}.[y_{t-1},\ x_t] + b_F) \tag{22}$$

$$y_t = ou_t.\tanh(c_t) \tag{23}$$

3.5. Benchmark models

Our proposed model was compared with seven benchmark models, including eXtreme Gradient Boosting (XG_Boost), ARIMA, Lasso Regression (Lasso_Reg), MultiLayer Perceptron (MLP), Support Vector Regression (SVR), Random Forest (RF), and Elastic_Net [10-12]. The six benchmark algorithms are described as follows:

- ARIMA is a classical model for time series prediction.
- XG_Boost is a gradient-boosting model that employs a tree-building technique to forecast its final value.
- Lasso_Reg is a popular model used in time prediction problems. It estimates the relationships between two variables to make predictions.
- MLP: is a fundamental neural network model. It has been used as an ideal tool to capture nonlinear relationships among time series data.
- SVR: This is a type of machine learning model used to predict complex issues.
- RF: is an ensemble learning model that employs a collection of decision trees to make a final prediction decision.
- Elastic_Net is a linear regression model that analyses a linear relationship between input and target variables.

3.6. Performance evaluation statistical metrics

The dataset was divided into a training set, a validation set, and a test set. We used the training to fit the model. We utilised the validation set to validate the proposed model performance during the training process, and it was not involved in the training phase. The loss of the trained model on the validation phase was adopted to set the final parameters of the proposed model. The test set was used for the evaluation of the final model. Several metrics were employed to evaluate the performance of the proposed model. Table 2 summarises all metrics used to evaluate the proposed model.

4. Model configuration and development

This section presents the development stages of the proposed model to forecast multistep ahead DBT_{air}.

Table 5
Comparison of the proposed FCSM model against the benchmarking models.

One day ahead												
Jazan station							Jeddah station					
Model	RMSE	MAE	RSE	ECC	WIA	CC	RMSE	MAE	RSE	ECC	WIA	CC
ARIMA	2.342	3.011	0.151	0.931	0.921	0.934	2.352	3.312	0.157	0.921	0.918	0.922
XG_Boost	2.351	3.112	0.143	0.921	0.923	0.923	2.331	3.211	0.155	0.913	0.921	0.911
Lasso_Reg	2.352	3.211	1.45	0.904	0.914	0.916	2.342	3.212	1.45	0.904	0.912	0.913
MLP	2.423	3.267	1.47	0.903	0.916	0.909	2.434	3.279	1.48	0.902	0.912	0.901
SVR	2.414	3.282	1.48	0.891	0.902	0.897	2.416	3.295	1.49	0.887	0.900	0.889
RF	2.331	3.212	1.46	0.883	0.891	0.883	2.333	3.224	1.47	0.880	0.884	0.879
Elastic_Net	2.351	3.274	1.49	0.894	0.904	0.896	2.363	3.285	1.48	0.887	0.890	0.894
FCSM	**2.120**	**2.912**	**0.123**	**0.971**	**0.981**	**0.982**	**2.131**	**2.921**	**0.113**	**0.969**	**0.979**	**0.980**
Seven days ahead												
ARIMA	2.421	3.212	0.154	0.901	0.912	0.902	2.451	3.25	0.157	0.892	0.902	0.899
XG_Boost	2.475	3.443	0.163	0.894	0.895	0.897	2.477	3.47	0.165	0.881	0.880	0.880
Lasso_Reg	2.491	3.486	0.175	0.883	0.862	0.882	2.495	3.488	0.178	0.877	0.861	0.864
MLP	2.495	3.493	0.181	0.878	0.871	0.875	2.497	3.498	0.186	0.863	0.859	0.858
SVR	3.021	3.783	0.189	0.845	0.852	0.851	3.037	3.798	0.193	0.838	0.841	0.843
RF	3.123	3.884	0.198	0.824	0.821	0.827	3.167	3.898	0.212	0.816	0.805	0.806
Elastic_Net	2.451	3.386	0.165	0.894	0.882	0.875	2.454	3.394	0.171	0.875	0.868	0.851
FCSM	**2.340**	**2.998**	**0.125**	**0.932**	**0.942**	**0.943**	**2.521**	**3.192**	**0.176**	**0.924**	**0.927**	**0.910**
One month ahead												
ARIMA	2.966	3.645	0.211	0.843	0.836	0.845	2.987	3.663	0.254	0.821	0.832	0.843
XG_Boost	2.954	3.614	0.199	0.853	0.843	0.852	2.954	3.614	0.199	0.853	0.843	0.852
Lasso_Reg	3.721	3.861	0.276	0.796	0.786	0.783	3.781	3.887	0.286	0.765	0.754	0.753
MLP	2.986	3.676	0.245	0.815	0.805	0.817	2.993	3.786	0.287	0.792	0.796	0.793
SVR	2.943	3.612	0.187	0.863	0.864	0.865	2.976	3.645	0.191	0.852	0.854	0.853
RF	3.321	3.778	0.285	0.775	0.765	0.765	3.356	3.798	0.295	0.742	0.759	0.749
Elastic_Net	2.921	3.515	0.182	0.851	0.853	0.854	2.943	3.553	0.187	0.843	0.841	0.834
FCSM	**2.821**	**3.211**	**0.178**	**0.912**	**0.921**	**0.910**	**2.876**	**3.312**	**0.189**	**0.901**	**0.891**	**0.890**

Table 6
Perfomnace evaluation in term of *ENS, KGE, APB.*

One-day ahead						
	Jazan			Jeddah		
Model	*ENS*	*KGE*	*APB*	*ENS*	*KGE*	APB
ARIMA	0.924	0.915	2.91	0.916	0.09	2.95
XG_Boost	0.923	0.921	2.76	0.921	0.918	2.81
Lasso_Reg	0.913	0.905	2.92	0.910	0.901	2.97
MLP	0.925	0.912	2.97	0.913	0.98	2.99
SVR	0.872	0.852	4.21	0.862	0.858	4.52
RF	0.854	0.833	4.65	0.844	0.824	4.72
Elastic_Net	0.842	0.836	4.43	0.837	0.829	4.52
FCSM	**0.982**	**0.982**	**2.21**	**0.978**	**0.971**	**2.23**
Seven-days ahead						
ARIMA	0.903	0.906	2.98	0.897	0.901	2.98
XG_Boost	0.884	0.892	3.98	0.873	0.875	3.98
Lasso_Reg	0.876	0.864	4.32	0.857	0.855	4.56
MLP	0.893	0.880	3.94	0.874	0.873	4.43
SVR	0.846	0.885	4.43	0.832	0.830	4.86
RF	0.821	0.824	4.75	0.814	0.819	4.89
Elastic_Net	0.831	0.826	4.65	0.825	0.819	4.85
FCSM	**0.965**	**0.964**	**2.25**	**0.952**	**0.951**	**2.75**
One-month ahead						
ARIMA	0.831	0.821	5.54	0.823	0.823	5.57
XG_Boost	0.852	0.848	5.32	0.843	0.832	5.45
Lasso_Reg	0.782	0.772	6.43	0.774	0.768	6.67
MLP	0.827	0.824	4.98	0.814	0.821	5.21
SVR	0.853	0.857	4.92	0.842	0.833	5.11
RF	0.773	0.772	6.53	0.765	0.766	6.76
Elastic_Net	0.843	0.832	5.63	0.839	0.825	5.88
FCSM	**0.922**	**0.924**	**2.87**	**0.913**	**0.922**	**2.99**

4.1. Problem formulation

The problem of air temperature prediction is to predict the one day, one week and one month ahead DBT_{air} based on historical time series input datasets. Given a time series $X\{t_1, t_2, \ldots t_{n-1}\}$ which includes information of one-day *dd*, one week *ww*, and one-month *mm* historical data, symbol $t_1, t_2, \ldots t_{n-1}$ represent the previous observation data points.

The air temperature time series contains 24 variables collected from two stations. We represent the time series X as a two-dimension matrix as follows:

$$\mathrm{X}(w) = \begin{bmatrix} x_{11}(t) & x_{12}(t) & . & . & x_{1n}(t) \\ & & & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{m1}(t) & x_{m1}(t) & . & . & x_{mn}(t) \end{bmatrix} \rightarrow \begin{matrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{matrix}$$

matrix $\mathrm{X}(w)$ Where each row represents one variable of weather data that includes air temperature, relative humidity, vapour pressure, wind speed and wind direction, clouds, etc., while the last vector refers to the target value for forecasting $[y_1, y_2, \ldots, y_n]$. The main challenge is to find the most significant variables that are highly correlated with the target values $[y_1, y_2, \ldots, y_n]$ to predict the future air temperature. The proposed model aims at implementing a sequence of steps from applying the transformation technique to the prediction model to find the actual values that correspond to the target $[y_1, y_2, \ldots, y_n]$. The dataset is split into the training set X_{train}, the testing set $X_{test}$and validation set X_{vali}. The model is trained using X_{train} set, and then the model is validating and tested using set $X_{test}$and validation set X_{vali}, respectively. The proposed model combined MEFD, MSIE, and LSFC models (long short-term memory LSTM-based FCNN fully connected neural network) to predict multistep ahead DBT_{air}.

Step 1: Input decomposition *via* MFFD

The MEFD decreases noise in the input data as a purpose to improve the prediction accuracy. Firstly, the dataset is pre-processed to remove unwanted variables with zero values. Then, the multivariate time series are passed through MEFD to decompose the data into several modes (i.e., signals).Fig. 7 illustrates the detected boundaries of time series at Jazan and Jeddah stations whereas algorithm 1 describes the MEFD model.

Step 2: Extraction of features *via* MSIE

Then, from each mode, the MSIE is extracted to form the final feature array. We extracted several distinct types of entropy features for comparison and employed several benchmark prediction models. These well-known entropy features named Multiscale dispersion entropy (MSDE), refined composite multiscale dispersion entropy (RCMDE), refined composite multiscale permutation entropy (RCMPE), and multiscale entropy (MSE). The extracted features from entropy were used for the final forecasting of multistep ahead DBT_{air}. Fig. 7 shows the diagram

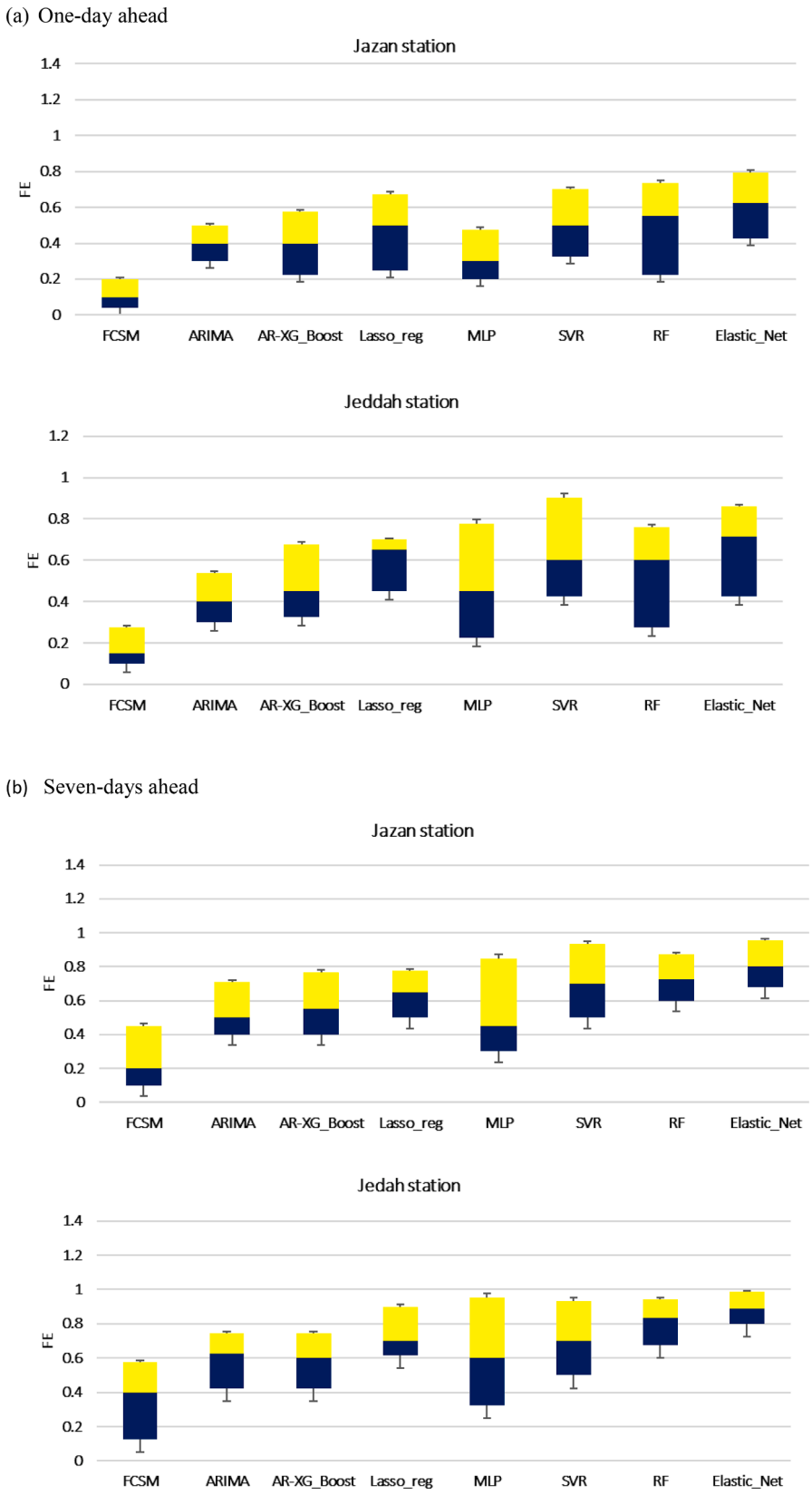


Fig. 10. Forecasting error of all the models in Jazan and Jeddah stations.

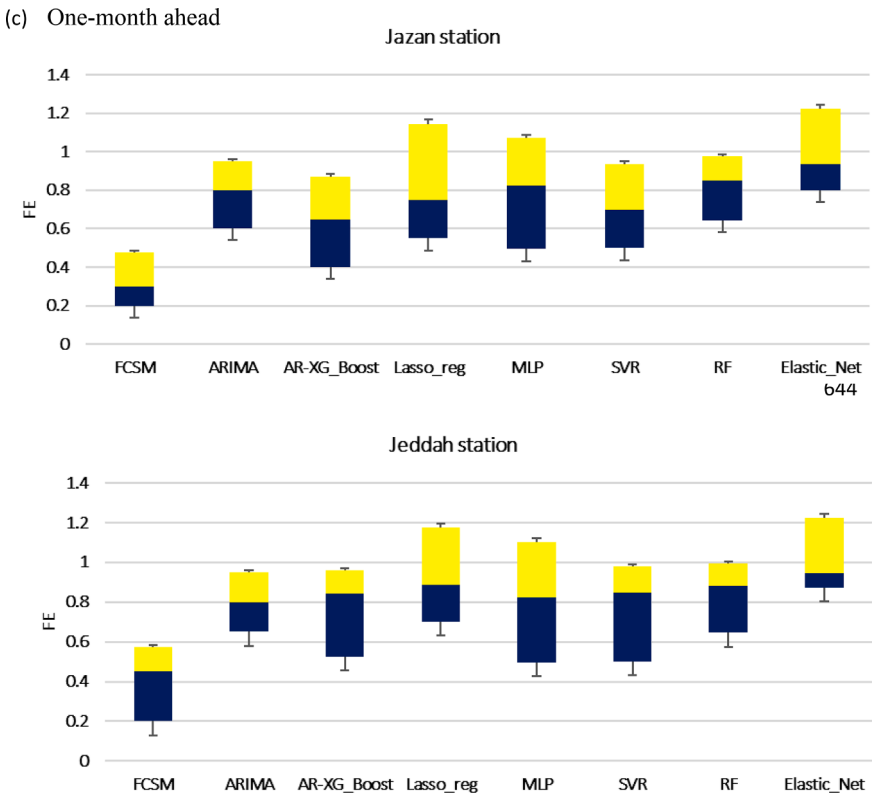


Fig. 10. (*continued*).

of MEFD.

Step 3: Combining LSMT and FCNN to construct FCSM model

The outputs of the LSTM and FCNN models are combined to form FCSM network to produce the final predicted value. The extracted features in step 2 were sent into the FCSM model to predict the multistep ahead DBT_{air}. Fig. 8 displays the schematic view of the proposed FCSM model. The proposed model avoid overfitting, minimise prediction error, ensuring an effective solution for multistep ahead DBT_{air} prediction. The FCSM combined the strength of FCNN, and LSTM offering more accurate and adaptable predictions than the baseline method.

4.2. Parameters setting

The depth and the number of neurons in each layer influence the prediction model's performance. We made a thorough investigation to select the optimal parameters using hit and trail method. The model under each parameter was trained using the training set and recorded all the testing results. The batch size was set to 128 where the loss function was set according to the MSIE. The number of epochs were 300 and the depths of LSTM and FCNN were selected empirically.

For the benchmark models, all parameters were selected empirically. Table 3 reports all tuning parameters of the eXtreme Gradient Boosting (XG_Boost), ARIMA, Lasso Regression (Lasso_Reg), MultiLayer Perceptron (MLP), Support Vector Regression (SVR), Random Forest (RF), and Elastic_Net models.

5. Results and discussion

Table 4 reports the prediction rate-based entropy features with different time intervals. Table 4 demonstrates the performance of the proposed model for prediction 1 day, 7 days, and one month. Based on the obtained results in Tables 3, the prediction rate is changed with the length of the time interval. In Table 4, the MSIE gained the highest rates of CC=0.982, WIA=0.981, and the lowest MAE=2.120 for the 1-day prediction with both stations. Similarly, the performance of MSIE was higher with one-week and one-month prediction compared with other entropy features, as shown in Table 4, respectively. The results showed the ability of MSIE to analyse complex time series in different intervals and extract the actual values to predict future data.

In addition, the MSDE scored the second highest prediction rate with CC = 0.874, WIA = 0873, the second lowest MAE = 2.78 for one-day prediction with both stations. However, the performance of MSDE suffered with the time interval of 7 days. Its results were decreased by 7% compared with the one-day prediction results. Furthermore, the RCMPE obtained the lowest prediction rate compared with other entropy features. A regression and Legates and McCabe's (LM) were also calculated for further investigation to assess the proposed model. Fig. 9 reports the prediction rates based on regression. The results in Fig. 9 support our finding in Table 4, which shows that the MSIE outperformed the other entropy features. We obtained regression = 0.96, 0.94, 091, with MSIE features for daily, weekly, and monthly prediction. However, the MSDE scored the second highest prediction rate in terms of regression for daily prediction compared with other entropy features. However, the performance of MSDE was degraded when it was applied for monthly and weekly prediction. In addition, the RCMPE showed the lowest performance in air prediction compared with other entropy features.

To evaluate the performance of the proposed model in prediction of air temperature, several benchmark models were utilised to conduct a comprehensive comparison. We selected seven benchmark models names eXtreme Gradient Boosting (XG_Boost), ARIMA, Lasso Regression (Lasso_Reg), MultiLayer Perceptron (MLP), Support Vector Regression (SVR), Random Forest (RF), and Elastic_Net models. All turning parameters were reported in the previous section. It is worth mentioning that these methods have been widely employed in several prediction problems as benchmark models and have shown good results. The extracted entropy features matrix was sent to the proposed FCSM model

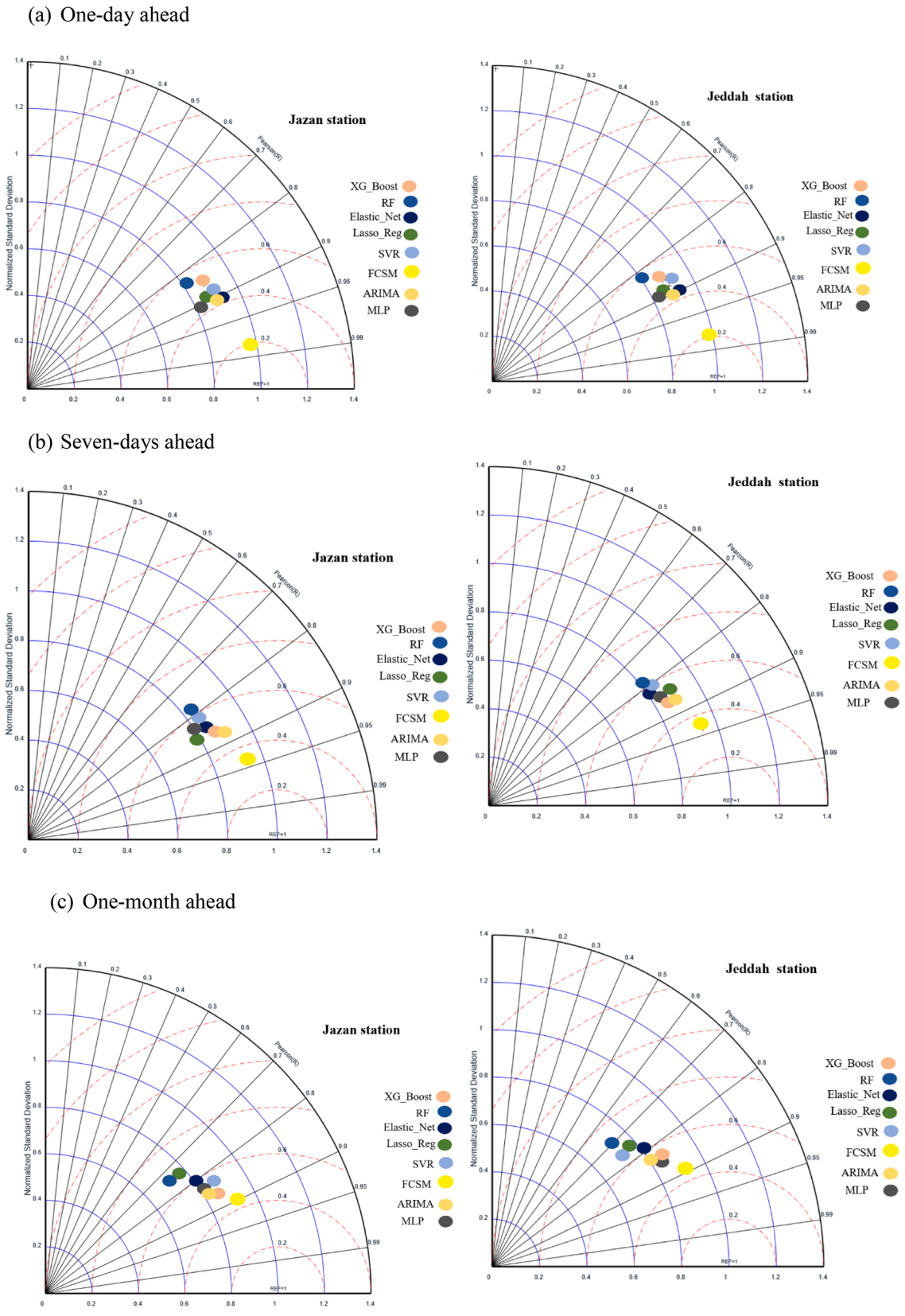


Fig. 11. Performance evaluation of each model using Taylor plots.

as well as the benchmark models. Tables 5-6 report the prediction rate in terms of RMSE, MAE, RSE, ECC, WIA, and CC using different time intervals.

In the first scenario, we investigated the capability of all models to predict values for daily temperature. Based on the training phase, not all models predicted the daily, weekly, and monthly temperature well. The evaluation of the proposed model performance based on the testing set was crucial. For the training phase, we trained all models with complete

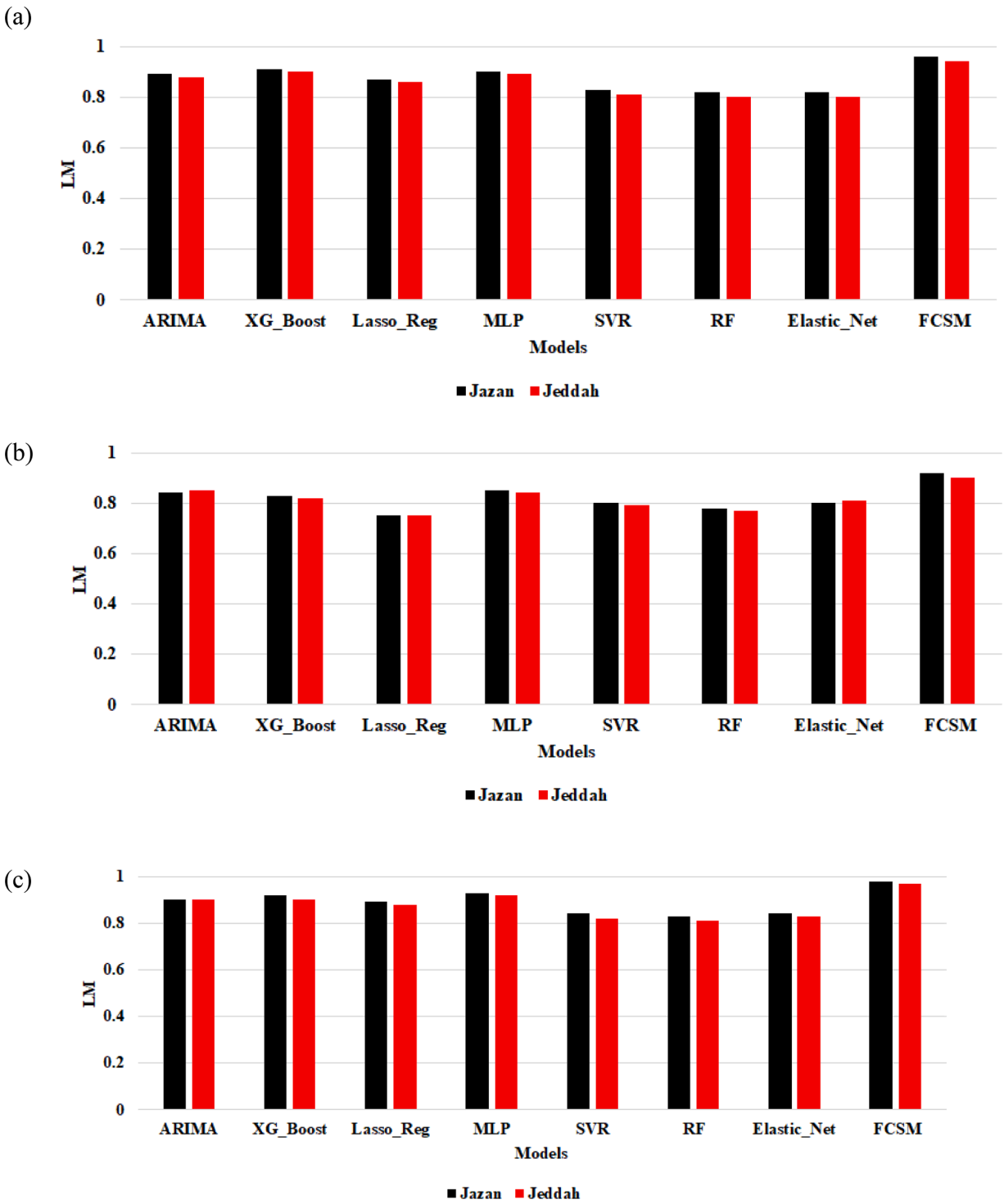


Fig. 12. Legate's and MacCabe's index of all the models in Jazan and Jeddah stations for (a) one-day, (b) seven-days, and (c) one-month DBT_{air} prediction.

data.

Table 5 lists the results of one-day, seven days and one-month ahead prediction for all models. From the results, we can notice that the proposed FCSM model gained the highest prediction rate compared with other models, and it did not show a high fluctuation when it was tested with different time intervals. The proposed FCSM model recorded the highest CC=982 and lowest RMSE=2.120 for daily temperature prediction. In addition, the XG_Boost and ARIMA showed acceptable performances with one-day prediction, and they obtained CC=923, RMSE=2.351, CC=934, and RMSE=2.342, respectively. We can observe that the results showed no significant differences between the average CC and MSE values obtained by XG_Boost and ARIMA, although the average results achieved by ARIMA were better. However, there were noticeable differences in the average metrics for these two models with respect to other models.

In the second scenario, we investigated the capability of all models to predict values for weekly and monthly temperatures. The monthly and weekly temperature prediction results are presented in Tables 5. We observed that most models performed poorly and struggled to predict air temperature one month ahead. For example, the RF model showed a

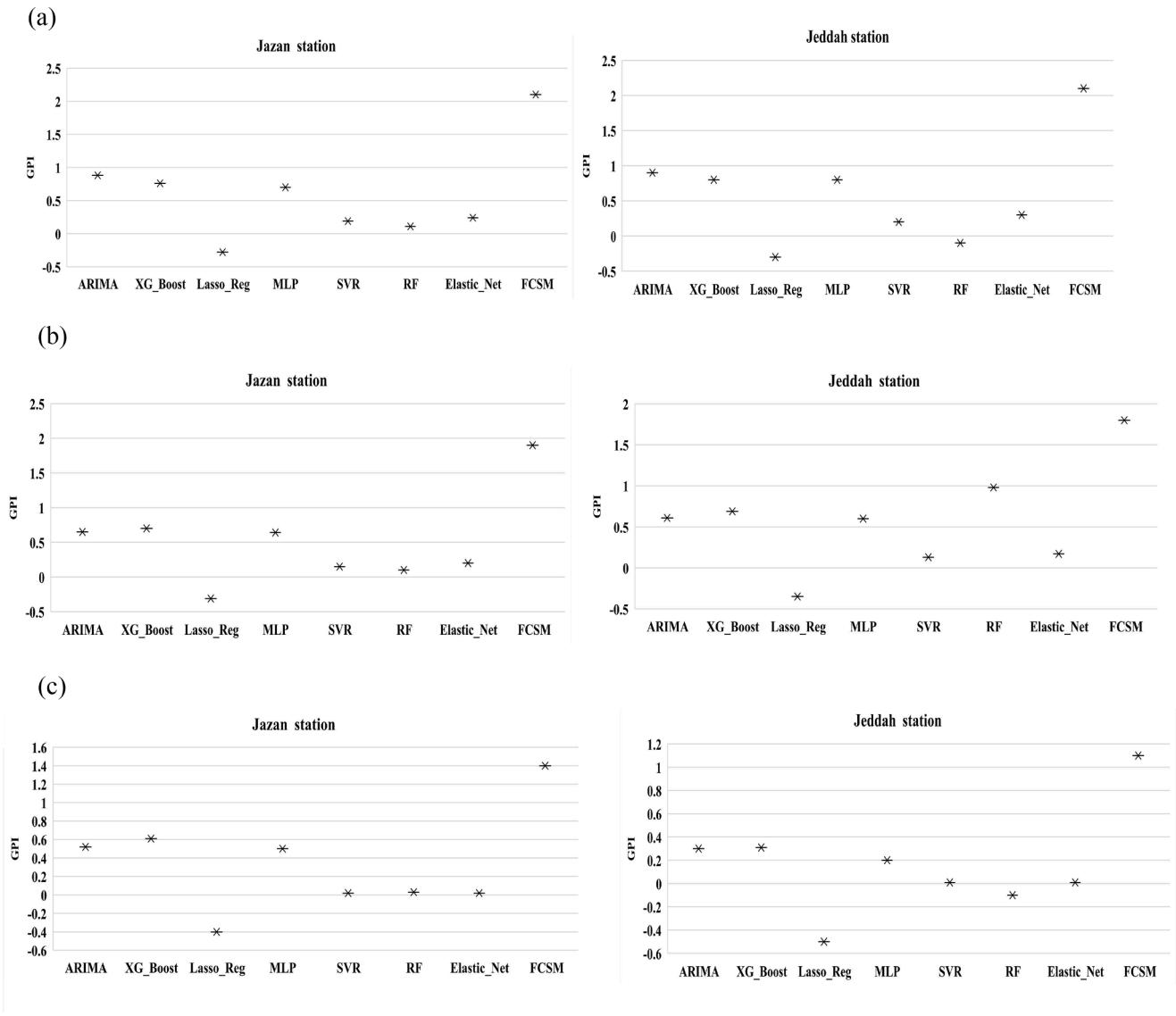


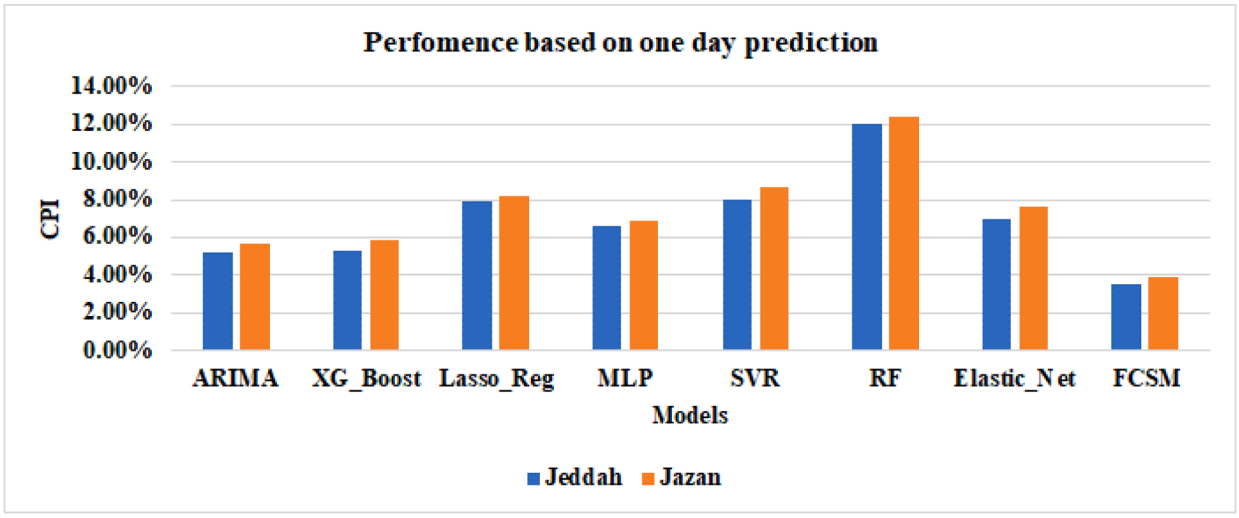
Fig. 13. Global Performance indicator (GPI) for daily, weekly, and monthly DBT_{air} prediction of all the models.

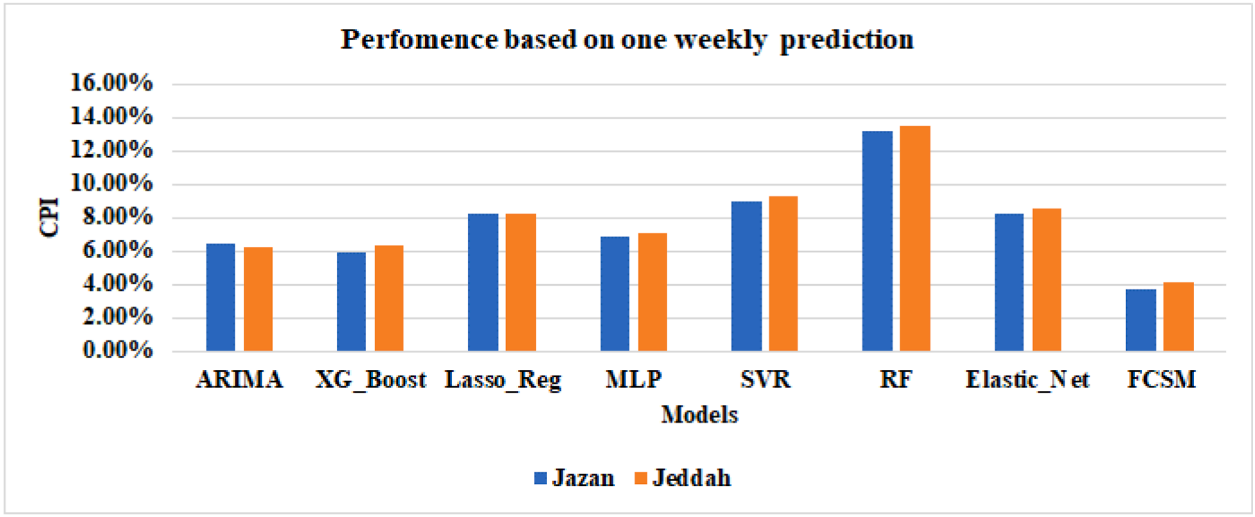
noticeable drop in performance, as shown in Tables 5 when it applied to predict air temperature one month ahead. It's worth noticing that the predictions made by the proposed model were close to the actual values. Also, it is important to mention that although predicting air temperature using different time intervals and different sources of data was a very difficult task, the proposed model was not affected by these factors.

For further analysis, Nash–Sutcliffe Index *(ENS)* the Kling– Gupta Efficiency *(KGE)*, and the Absolute Percentage Bias *(APB,%)* were used to show intuitively the prediction performance of the proposed model. Based on performance results in Tables 5, the proposed outperform the benchmark models for two stations for weekly, daily, and monthly prediction. At Jeddah station, for the proposed model FCSM yielded *ENS* =0.983 and KGE 0.982 compared to *ENS* =0.924 and KGE 0.915 for ARIMA model for daily temperature prediction. Similarly, for the Absolute Percentage Bias (APB%) error, which was calculated during the testing phase. The proposed FCSM model produced a lower value of APB than the benchmark models. It achieved less than 3% for two stations for daily, weekly, and monthly prediction. For example, the proposed model's lowest value of APB produced 2% for daily temperature prediction, while the RF model for both stations obtained the highest value of APB= 6.57%.

For all two stations and different time intervals, the best *ENS,KGE* and *APB* were found for the Jazan station. Most models at Jazan station showed excellent performance with daily prediction. However, the performances of most models dropped when they dealt with the monthly prediction problem. The results indicate that empirical Fourier decomposition (MEFD) can significantly improve air temperature prediction based on multiscale increment entropy (MSIE), coupled with FCSM. The proposed algorithm could reduce the nonstationary nature of time series data.

We also used a forecasting error metric to evaluate the proposed model. Fig. 10 compares the proposed FCSM model against the benchmark models in terms of forecasting error using box plots for two stations using three-time intervals. The First quartile, represented by 25th percentile, and the third quartile, represented by 75th percentile, were designated using the lower and upper lines of the boxplot. The central line denoted the median (50th percentile). Two horizontal lines marked the smallest and largest non-outliers. From the obtained results in Fig. 10, the smaller value FE with a smaller quartile range refers to better prediction results. It can be noticed that the proposed MEFD-MSIE-FCNN yielded overall better performance for daily temperature prediction for the two stations, followed by the other models ARIMA, MLP, and XG_Boost. In addition, Fig. 10 show the results for weekly and monthly prediction. Based on the results, the proposed FCSM model achieved





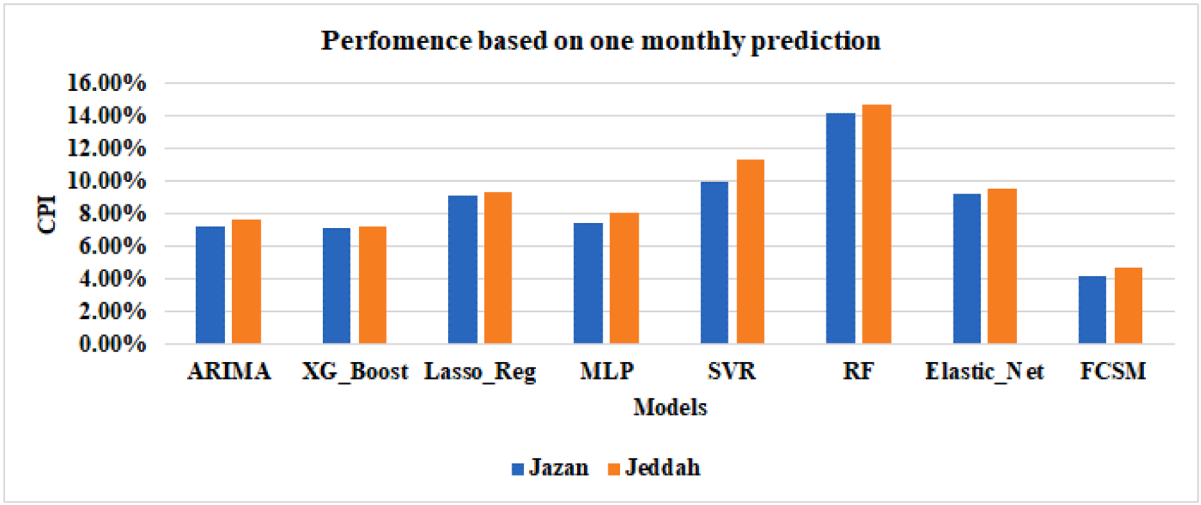


Fig. 14. CPI-based results for all models.

great performance and did not show any significant drop for both stations. However, the other models showed low performance for monthly prediction.

Fig. 11 shows the Taylor diagrams of all models. The diagram shows how the proposed model is closest to the actual values during the testing phase by computing the angular location of the inverse cosine of the correlation coefficient. In Figs. 11a-11c, we can observe that the correlation coefficient (r) on the x-axis and the standard deviation (SD) on the polar axis are employed to show the best fit of the model to the predictor. For all prediction time intervals, the proposed FCSM model

gained the highest *r* value, showing the closest prediction to the actual air temperature data for two stations. Again, the data generated by other models differed significantly from that generated by the proposed model. With weekly and monthly air temperature prediction in Fig. 11 (b-c), the proposed FCSM outperformed all models and generated the closest data to the predictor followed by ARIMA, MLP, and XG_Boost.

We also employed Legate's and Maccabee's (LM) metric for further analysis in this paper. Fig. 12 shows the performance of all models using Legate's and MacCabe's for daily, weekly, and monthly prediction. The proposed model obtained LM 0.98 and 0.97 at Jazan and Jeddah stations, respectively, for daily prediction, which were the lowest values compared to other models. The results in Fig. 12 demonstrated the efficacy of combining MEFD, MSIE, and FCNN models in air predicting. Thus, considering the various entropies tested in the previous section, the corresponding prediction performance can be further enhanced using other features in future work.

Based on the obtained results presented so far, it can be observed that the proposed FCSM model has considerably advanced in state of the art used in daily, weekly, and monthly DBT_{air} predictions. The practical implication of the proposed model was also apparent. Environment decision-makers and professionals in the weather sector can benefit from using the proposed FCSM model as an intelligent framework for obtaining accurate predictions. It can be applied for different prediction applications such as electricity, pollution prediction, and oceanic energy production, as it successfully dealt with the non-stationary and nonlinear behaviour of time series data. In this section, we shed more light on some critical findings.

1. The results obtained from two regions in Saudia Arabia provided important understandings regarding regional temperature changes. By analysing the abnormal patterns, environmental influences, and geographic factor, the proposed model can be applied to other regions and cities. As each region has unique environment conditions which influence DBT_{air}, the proposed model parameters can be adjusted by training using different climate variables.
2. To make a thorough investigation, Fig. 13 illustrates the Global Performance Indicator *GPI* values. It was shown from Fig. 13 that the proposed model FCSM gained higher *GP I* values compared to other models. From Fig. 13(a-c), it can be observed that the proposed model FCSM has stable performance with two stations and higher GPI= 2.8 for daily temperature prediction. This higher value in GPI for the proposed model was because of the lower *RMSE* of the proposed model compared to other models. In addition, the proposed FCSM outperforms all other benchmark models for monthly and weekly air temperature prediction. For instance, at Jeddah and Jazan stations, it obtained *GP I* =1.8, 2.1 respectively, while the other models, ARIMA, MLP, and XGBoost, gained the second-highest values.
3. Despite the great performance of the proposed FCSM model against the benchmark models, it has some limitations that should be addressed in future research. The proposed FCSM model is required to assess its generalisation capabilities for air temperature prediction using more stations. The current study only included data from two stations collected from Saudi Arabia. The prediction capacity of the FCSM model for additional regions was not evaluated. Considering air temperature data from other areas, such as Australia and other Middle Eastern countries, may be beneficial in validating the proposal. Furthermore, another limitation of the proposed model was that it showed a slight fluctuation in prediction results when applied to monthly prediction. Investigation of more entropy and statistical features could improve the long-term prediction results.
4. For more evaluation, a combined Performance Index (CPI) was also adopted in this paper to demonstrate the prediction capability of the proposed FCSM model. CPI is calculated as a weighted sum of numerous metrics, including Kolmogorov–Smirnov test integral (V *KSI*), *RMSE* and relative frequency of exceedance (over). It was calculated as $CP\ I = (KSI + OVER + 2RMSE)\ /4$). It should be explained that the lower the CPI value, the better performance. Fig. 14 shows the performance of models based on *CPI* metric at two stations. Based on the results of the CPI bar chart, it can be noticed that the proposed FCSM model yielded the lowest *CPI* values for two stations and three prediction time intervals. For instance, at Jazan station and Jeddah station, the proposed model obtained the lowest *CPI*=3.5%, 3.9%, ≈ for one-day prediction. In addition, it also yielded the lowest CPI for weekly and monthly predictions. According to the results of the preceding GPI and CPI metrics analysis, the proposed FCSM model outperforms the other models in addressing air temperature prediction issues, and it is more suited for long-term and short-term prediction.

The model only used time-series data, yet in future independent studies, satellite data as in [11] can be used, or the ERA5 Reanalysis spatial data [52] can also be applied for DBT_{air} forecasting.

5. According to our findings the fluctuation in monthly prediction results was related to the size of the dataset. The predictive accuracy of the proposed model may fluctuate due to shorter time-series. One of future solution is that improving data preprocessing and augmentation by considering such as interpolation method that could help balance the distributions of data across time. This can ensure the proposed model is not biased for certain periods of time.
6. Although the proposed model demonstrated significant prediction results, but there are some limitations that restrict the scope of thi study. The proposed model has not been tested with noisy data could affect the model's accuracy. In addition, the proposed model combined several models that can increase the computational time and delay the response in real time applications. Furthermore, data from two locations were used to test the proposed model. This can affect the generalisation of the proposed model.
7. The proposed model was tested using data from two regions, it may not extract the full variability of time series patterns across diverse climatic conditions in different locatons. This could limit the generalization to adapt well to the new data. Thus, using data from other regions with different environment, and infrastructure are required for further evaluation. Additionaly, the proposed model accuracy was slightly lower in monthly predictions which can be addrssed in the future by incorporating more input variables and historical data.

6. Conclusions

In this paper an accurate forecasting air temperatures dry bulb (i.e., DBTair) model was designed. To predict daily, weekly, and monthly air temperatures, we have proposed a hybrid intelligent model named FCSM, where MEFD is used for data decomposition, MSIE as a feature extractor, LSTM and FCNN is implemented to predict DBTair. The proposed model was trained and tested using the data from two stations (i.e., Jazan and Jeddah) in Saudia Arabia. The proposed FCSM model shows promising improvements against comparing models where the MEFD is used to handle the non-stationarity and non-linearity issues within in the data. The MSIE is adopted as a feature extractor to capture the most significant and relevant features whereas the LSTM and FCNN is implemented as predictive models to predict multistep ahead DBTair. The results showed that the proposed FCSM model was able to capture complex patterns and nonlinear relationships in time series data as compared to the traditional models. Additionally, the FCSM reduced overfitting, and adapted well to reflections in the time series data. The high accuracy obtained by the FCSM demonstrated that it can be a reliable model to assist informed decisions for environmental, hydrological, agricultural and energy applications. The future research will focus on integrating the proposed model with Internet of Things (IoT) sensors for real-time data collection. This can improve the predictive

capabilities of model to unexpected temperature variations. In addition, the proposed model will be tested in various climatic conditions to generalise the adoptability in different geographic locations. The scope of the proposed FCSM model can be extend in other sectors such as water monitoring, environment, and renewable energy areas for accurate predictions to make better and timely decision.

Data Availability

The authors do not have the permission to share the data.

CRediT authorship contribution statement

Mohammed Diykh: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Mumtaz Ali:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Abdulhaleem H. Labban:** Writing – review & editing, Supervision, Data curation, Conceptualization. **Ramendra Prasad:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Mehdi Jamei:** Writing – review & editing, Visualization, Methodology, Conceptualization. **Shahab Abdulla:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Aitazaz Ahsan Farooque:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are thankful to the Department of Meteorology, Saudi Arabia for providing the relevant datasets. The authors would like to thank Natural Science and Engineering Council of Canada for supporting this collaborative research.

Data availability

The authors do not have permission to share data.

References

[1] IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. R. K. Pachauri and L.A. Meyer. Geneva, Switzerland.: 151 pp.

[2] C.A. McAlpine, J. Syktus, J.G. Ryan, R.C. Deo, G.M. McKeon, H.A. McGowan, S. R. Phinn, A continent under stress: interactions, feedbacks and risks associated with impact of modified land cover on Australia's climate, Glob. Chang. Biol. 15 (9) (2009) 2206–2223.

[3] R. Prasad, D. Charan, L. Joseph, T. Nguyen-Huy, R.C. Deo, S. Singh, Daily Flood Forecasts with Intelligent Data Analytic Models: Multivariate Empirical Mode Decomposition-Based Modeling Methods (2021) 359–381.

[4] S. Li, L. Goel, P. Wang, An ensemble approach for short-term load forecasting by extreme learning machine, Appl. Energy 170 (2016) 22–29.

[5] R.M. Adnan, Z. Liang, A. Kuriqi, O. Kisi, A. Malik, B. Li, F. Mortazavizadeh, Air temperature prediction using different machine learning models, Indonesian J. Elect. Eng. Comput. Sci. 22 (1) (2021).

[6] M.S. Hanoon, A.N. Ahmed, N. Zaini, A. Razzaq, P. Kumar, M. Sherif, A. Sefelnasr, A. El-Shafie, Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia, Sci. Rep. 11 (1) (2021) 18935.

[7] J. Huang, Z. Yang, Correlation between air temperature, air pollutants, and the incidence of coronary heart disease in Liaoning Province, China: a retrospective, observational analysis, Ann. Palliat. Med. 10 (12) (2021) 12412–12419.

[8] Y. Ohashi, T. Ihara, K. Oka, Y. Takane, Y. Kikegawa, Machine learning analysis and risk prediction of weather-sensitive mortality related to cardiovascular disease during summer in Tokyo, Japan, Sci. Rep. 13 (1) (2023) 17020.

[9] L. Xu, H. Yu, X. Zhang, Y. Gan, Extreme low apparent temperature forecasting for early warning of mortality by data-driven deep learning, Journal of Geography and Cartography 6 (1) (2023).

[10] Y.R. Al-Saadi, M.S. Tapou, A.A. Badi, S. Abdulla, M. Diykh, Developing Smart Self Orienting Solar Tracker for Mobile PV Power Generation Systems, IEEe Access. 10 (2022) 79090–79099.

[11] G. Tepanosyan, S. Asmaryan, V. Muradyan, R. Avetisyan, A. Hovsepyan, A. Khlghatyan, G. Ayvazyan, F. Dell'Acqua, Machine Learning-Based Modeling of Air Temperature in the Complex Environment of Yerevan City, Armenia, Remote Sens. (Basel) 15 (11) (2023).

[12] A. Hussain, H.F. Sipra, A. Waheed, K.E. Ukhurebor, Exploring the academic perceptions of climate engineering in developing countries, Atmósfera (2024) 38.

[13] K.E. Ukhurebor, U.O. Aigbe, R.B. Onyancha, H. Athar, B. Okundaye, P. A. Aidonojie, B.E. Siloko, I. Hossain, H.S. Kusuma, H. Darmokoesoemo, Environmental Influence of Gas Flaring: Perspective from the Niger Delta Region of Nigeria, Geofluids 2024 (1) (2024) 1321022.

[14] K.E. Ukhurebor, K.R. Singh, V. Nayak, U.E. Gladys, Influence of the SARS-CoV-2 pandemic: a review from the climate change perspective, Environmental Science: Processes & Impacts 23 (8) (2021) 1060–1078.

[15] K.E. Ukhurebor, S.O. Azi, U.O. Aigbe, R.B. Onyancha, J.O. Emegha, Analyzing the uncertainties between reanalysis meteorological data and ground measured meteorological data, Measurement 165 (2020) 108110.

[16] B.R. Bonsal, X. Zhang, L.A. Vincent, W.D. Hogg, Characteristics of Daily and Extreme Temperatures over Canada, J. Climate 14 (2001) 1959–1976.

[17] A. Mellit, A.M. Pavan, M. Benghanem, Least squares support vector machine for short-term prediction of meteorological time series, Theor. Appl. Climatol. 111 (1-2) (2012) 297–307.

[18] M. Apaydin, M. Yumuş, A. Değirmenci, Ö. Karal, Evaluation of air temperature with machine learning regression methods using Seoul City meteorological data, Pamukkale Univ. J. Eng. Sci. 28 (5) (2022) 737–747.

[19] B. Gong, M. Langguth, Y. Ji, A. Mozaffari, S. Stadtler, K. Mache, M.G. Schultz, Temperature forecasting by deep learning methods, Geosci. Model. Dev. 15 (23) (2022) 8931–8956.

[20] J. Zhou, D. Wang, S.S. Band, E. Mirzania, T. Roshni, Atmosphere air temperature forecasting using the honey badger optimization algorithm: on the warmest and coldest areas of the world, Engineering Applications of Computational Fluid Mechanics 17 (1) (2023).

[21] T. Thi Kieu Tran, T. Lee, J.-Y. Shin, J.-S. Kim, M. Kamruzzaman, Deep Learning-Based Maximum Temperature Forecasting Assisted with Meta-Learning for Hyperparameter Optimization, Atmosphere (Basel) 11 (5) (2020).

[22] D.S. Roy, Forecasting The Air Temperature at a Weather Station Using Deep Neural Networks, Procedia Comput. Sci. 178 (2020) 38–46.

[23] J. Hou, Y. Wang, J. Zhou, Q. Tian, Prediction of hourly air temperature based on CNN–LSTM, Geomatics, Natural Hazards and Risk 13 (1) (2022) 1962–1986.

[24] M. Ali, R. Prasad, Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition, Renew. Sustain. Energy Rev. 104 (2019) 281–295.

[25] W.Y. Duan, Y. Han, L.M. Huang, B.B. Zhao, M.H. Wang, A hybrid EMD-SVR model for the short-term prediction of significant wave height, Ocean Engineering 124 (2016) 54–73.

[26] X. Li, C. Li, Improved CEEMDAN and PSO-SVR Modeling for Near-Infrared Noninvasive Glucose Detection, Comput. Math. Methods Med. 2016 (2016) 8301962.

[27] R. Prasad, R.C. Deo, Y. Li, T. Maraseni, Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition, Geoderma 330 (2018) 136–161.

[28] R. Prasad, M. Ali, P. Kwan, H. Khan, Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation, Appl. Energy 236 (2019) 778–792.

[29] W.-y. Duan, L.-m. Huang, A hybrid EMD-AR model for nonlinear and non-stationary wave forecasting, Journal of Zhejiang University Science A 17 (2016) 115–129.

[30] Q. Ouyang, W. Lu, X. Xin, Y. Zhang, W. Cheng, T. Yu, Monthly Rainfall Forecasting Using EEMD-SVR Based on Phase-Space Reconstruction, Water Resources Management 30 (7) (2016) 2311–2325.

[31] M.A. Colominas, G. Schlotthauer, M.E. Torres, Improved complete ensemble EMD: A suitable tool for biomedical signal processing, Biomed. Signal. Process. Control 14 (2014) 19–29.

[32] W. Zhang, Z. Qu, K. Zhang, W. Mao, Y. Ma, X. Fan, A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting, Energy Convers. Manage 136 (2017) 439–451.

[33] N. Rehman, D.P. Mandic, Multivariate empirical mode decomposition, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466 (2117) (2009) 1291–1302.

[34] N.u. Rehman, H. Aftab, Multivariate Variational Mode Decomposition, IEEE Transactions on Signal Processing 67 (23) (2019) 6039–6052.

[35] N. Ur Rehman, C. Park, N.E. Huang, D.P. Mandic, Emd Via Memd: Multivariate Noise-Aided Computation of Standard Emd, Adv. Adapt. Data Anal. 05 (02) (2013).

[36] K. He, R. Zha, J. Wu, K. Lai, Multivariate EMD-Based Modeling and Forecasting of Crude Oil Price, Sustainability. 8 (4) (2016) 387.

[37] W. Hu, B.C. Si, Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition, Geoderma 193-194 (2013) 180–188.

[38] W. Zhou, Z. Feng, Y. Xu, X. Wang, H. Lv, Empirical Fourier decomposition: An accurate signal decomposition method for nonlinear and non-stationary time series analysis, Mech. Syst. Signal. Process. 163 (2022) 108155.

[39] A. Bhattacharyya, R.B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, IEEE Trans. Biomed. Eng. 64 (9) (2017) 2003–2015.

[40] J. Gilles, Empirical wavelet transform, IEEE transactions on signal processing 61 (16) (2013) 3999–4010.
[41] R. Lafta, J. Zhang, X. Tao, Y. Li, M. Diykh, J.C.W. Lin, A structural graph-coupled advanced machine learning ensemble model for disease risk prediction in a telehealthcare environment, Big data in engineering applications (2018) 363–384.
[42] A. Mahato, K. Das, R.B. Pachori, A Multivariate Approach for Drowsiness Detection Using Empirical Fourier Decomposition, Authorea Preprints (2023).
[43] M. Costa, A.L. Goldberger, C.-K. Peng, Multiscale entropy analysis of complex physiologic time series, Phys. Rev. Lett. 89 (6) (2002) 068102.
[44] X. Liu, A. Jiang, N. Xu, J. Xue, Increment entropy as a measure of complexity for time series, Entropy 18 (1) (2016) 22.
[45] H. Mohammed, M. Diykh, Improving EEG major depression disorder classification using FBSE coupled with domain adaptation method based machine learning algorithms, Biomed. Signal. Process. Control 85 (2023) 104923.
[46] X. Wang, X. Liu, W. Pang, A. Jiang, Multiscale increment entropy: An approach for quantifying the physiological complexity of biomedical time series, Information Sciences 586 (2022) 279–293.
[47] Y. Zhang, J. Lee, M. Wainwright, M.I. Jordan, On the learnability of fully-connected neural networks, Artificial Intelligence and Statistics (2017). PMLR.
[48] S. Abdulla, M. Diykh, S. Siuly, M. Ali, An intelligent model involving multi-channels spectrum patterns based features for automatic sleep stage classification, Int. J. Med. Inform. 171 (2023) 105001.
[49] A.G. Schwing, R. Urtasun, Fully connected deep structured networks (2015) arXiv preprint arXiv:1503.02351.
[50] G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model, Artif. Intell. Rev. 53 (8) (2020) 5929–5955.
[51] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[52] D. Fister, J. Pérez-Aracil, C. Peláez-Rodríguez, J. Del Ser, S. Salcedo-Sanz, Accurate long-term air temperature prediction with Machine Learning models and data reduction techniques, Appl. Soft. Comput. 136 (2023).