Nonlinear Dynamic Analysis of Meteo–rological Variables for Ha'il Region, Saudi Arabia, for the Period 1990-2022

Mohammed Abdul Majid
Student
National University of Malaysia
Faculty of Science and Technology
Malaysia

Mohd Salmi M. Noorani
Professor
National University of Malaysia
Malaysia

Fatimah Abdul Razak
Associate Professor
National University of Malaysia
Malaysia

The study applies diverse methods of chaos detection to meteorological variable data (air temperature, relative humidity, surface pressure, precipitation, and wind speed for Ha'il, Saudi Arabia) to understand the nonlinear dynamics and to classify their nature. Additionally, Random Forest Algorithm model is used to predict the precipitation and wind speed. The results obtained by classical and modern approaches are compared. All the variables are found to be chaotic based on correlation dimension, approximate entropy, and 0–1 test. The chaos decision tree algorithm diagnoses air temperature, relative humidity, and wind speed as chaotic, while precipitation and surface pressure are identified as stochastic. This shows that the classical methods are well-validated with the modern methods. Nevertheless, some of them contradict modern methods. The analysis for 32 years of data showed no precipitation for 92% of the time during the entire period based on the Random Forest algorithm.

Keywords: *Nonlinear dynamic analysis; chaos theory; climate variables; time series; chaos detection; recurrence analysis, random forest algorithm*

1. INTRODUCTION

Understanding the inherent dynamics of meteorological variables is vital for the optimum utilization of natural Renewable Energy Resources. Saudi Arabia covers the largest part of the Arabian Peninsula [1]. A vast arid land desert climate dynamics with scarce precipitation in many parts of the region. The regulation concerning greenhouse gas mitigation demands the utilization of clean and renewable energy sources in Saudi Arabia. Undoubtedly, the region is witnessing an increased power demand in the future due to population, indus–trial, and economic growth [2].

Studying dynamic variability, high fluctuations, and its relationship with climate change can be useful in weather forecasting [3]. Understanding the nonlinear behavior of meteorological variables using time series data is an important scientific area in the analysis and forecasting of spatiotemporal systems. The observations collected at fixed time intervals as a time series are a dynamic system. Nonlinear chaotic systems exhibit dy–namic behavior sensitive to initial conditions, a pheno–menon popularly known as the butterfly effect [4]. Linear methods cannot fully capture the dynamics of these systems, and studies relating to chaos theory have been extensively used in identifying and characterizing such systems. The work on nonlinear dynamics in weather forecasting has revolutionized the long-held belief that the climate is only random [4]. In climatic research, applying nonlinear dynamic theory can imp–rove the understanding of the behavior of meteoro–logical variables and their prediction [5-8]. The diff–erence between stochastic and chaotic systems is that the former yields multiple outcomes to one input, which is studied under the probability distribution theory. The chaotic system is sensitive to initial conditions, and its solution reveals a strange attractor [9]. Because of the developments in nonlinear methods and advancements in computer technology, the detection of chaos in dynamic systems has become a reality.

Some of the signatures of chaos are the nonlinear inter-dependence, order of degrees of freedom, comp–lexity, deterministic nature, and a small perturbation resulting in a significant future behavior [10]. Advan–cements in the analysis of time series data [11,12] and the development of computer-based forecasting algo–rithms [13,14] have greatly contributed to chaos theory. The field has witnessed significant theoretical and methodological developments and extensive applica–tions in interdisciplinary areas [10]. The chaos detection methods are classified as classical or modern. Insights into the behavior of time-varying systems and their modeling are essential in climate and weather fore–casting and natural disasters such as flooding, tornadoes, and earthquakes [15]. The methods such as recurrence plot (RP) [16], recurrence quantification analysis (RQA) [17], Chaos decision tree algorithm (CDTA) [18], and 0-1 test for chaos are relatively modern.

There are several classical methods; among them, four methods, such as Correlation Dimension (CD), Lyapunov exponent (LE), Approximate entropy (AE), and Phase Space Reconstruction (PSR), are taken into account. The CD is widely used for variables such as temperature [19-21], wind speed [5,20,22], and preci–pitation [19,20,23,24]. Lyapunov exponent is another popular method that estimates the chaotic nature of the system. A positive value of divergence of trajectories (λ) indicates chaos [22, 25-26]. That also applies to rainfall [27] and

Received: March 2023, Accepted: April 2023
Correspondence to: Mohammed Abdul Majid, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia,
E-mail: abdulmajid_2000@yahoo.com
doi: 10.5937/fme2302231M

many meteorological variables [19,28-30]. Approximate entropy is a quantification measure of variable complexity. It is used to classify systems in several disciplines, including earth sciences [31], PSR provides visualization of trajectories [23] and is an appropriate tool for qualitative analysis of system dynamics [32].

Chaos theory deals with systems evolution, which shows sensitivity to initial conditions. Its hallmark is a departure from stability to an unstable state or from order to chaos, such as a mechanical system. A single spherical pendulum in motion has decipherable behavioral patterns, and thus it is deterministic, whereas two-body systems such as double pendulums show chaotic behavior. Therefore, its future state is virtually unpredictable [33]. The existence of a strange attractor is qualitative proof of chaotic dynamics in a system, and it is quantified by methods such as CD, LE, and entropy. A promising outcome of the theory is a new research area called nonlinear science. Weather is a classic example of a chaotic system. The predictability of weather is only possible for a short time.

Chaos is ubiquitous in nature, and engineering systems are no exception. It has been applied in the field of mechanical engineering. In structural dynamics, it has been used to analyze complex structures such as bridges, buildings, or airplane wings as they benefit from chaos theory [34]. The behavior of these structures can be modeled using nonlinear equations, which can lead to chaos under certain conditions. Engineers can predict and avoid potential failures or catastrophic events by analyzing chaotic behavior in these systems. Robotics is another area where chaos theory can also be used to design and control robotic systems [35].

By analyzing the chaotic dynamics of robotic systems, engineers can design better control algorithms to stabilize the system and reduce vibration or mechanical noise. Turbulent flows, such as those in fluid dynamics, are difficult to predict and control. However, chaos theory can help explain the complex behavior of turbulent flows and offer insights into how to model and control them. Chaos theory can be applied in the manufacturing process [36] to optimize production efficiency and reduce waste. By analyzing the chaotic behavior of manufacturing processes, engineers can identify areas where variability can be reduced, leading to higher-quality products and reduced costs. These are just a few examples of how chaos theory can be applied in mechanical engineering. The following are some of the direct applications of chaos theory to mechanical engineering problems:

The application of chaos theory for predicting wind speed [37] has important implications for the design of wind turbines. Chaos analysis of a vehicle semi-active suspension system with uncertainties [38]. This study applies Chaos theory was applied to analyze the dynamic behavior of a vehicle's semi-active suspension system under various uncertainties [38]. These studies demonstrate the broad applicability of chaos theory in mechanical engineering, ranging from wind energy to vehicle suspension systems to gear tooth design.

To our knowledge, no comprehensive study on meteorological parameters is reported in the literature. Moreover, a combination of classical methods with newly developed CDTA methods [18], recurrence ana–lysis, and 0-1 tests for chaos in conjunction with classical methods in the Saudi Arabian context is used for the first time.

The recent data set for 32 years has been used to understand the dynamics of these variables and their relationship with climate change.

The Random Forest algorithm (RFA) is a wide–spread robust, supervised computer algorithm. It is ver–satile and can be used in both classification and reg–ression problems. The rainfall in the Arabian Peninsula is scarce, and studies relating to it could contribute to understanding its dynamics and managing water resour–ces in the region. Therefore, more consideration is given to the precipitation. Wind speed prediction is important from a renewable energy resources assess–ment and grid integrity point of view. Saudi Arabia is a signee of the GHG reduction protocol and has prio–ritized utilizing wind and solar sources to generate clean energy.

This study aims to apply diverse methods to meteo–rological variable data to understand the nonlinear dynamics involved and classify their nature. Further–more, the study also attempts to predict two important variables, precipitation, and wind speed, as they are important renewable energy sources. The meteoro–logical variables are daily air temperature, relative humidity, surface pressure, precipitation, and wind speed. These variables are important sources of a rene–wable energy assessment, such as humidity [39], wind speed, and rainfall [40,43,44- 49]. The proposed work has practical applications in weather and atmospheric science [50] and multi-disciplinary areas such as engineering, mathematics, and the stock market [51,52].

2. SITE AND DATA DESCRIPTION

The daily time series data between 1990 and 2022 (11964 data points) was obtained for the Ha'il region from the POWER project's hourly, v2.2.8 version MERRA-2 (Modern-Era Retrospective analysis for Research and Applications) [53]. MERRA-2 uses the resolution ½° latitude by ⅝° longitude with an average for 0.5 x 0.625-degree latitude/longitude region equi–valent to 968.37 meters for data derivation. The latitude, longitude, and altitude of the Ha'il region are 27.634602□N, 41.723572□E, and 991 meters, respecti–vely. The area of the Ha'il region is 103,887 km². The air temperature and relative humidity are measured at 2 meters, and wind speed at 10 meters above the ground level. Its climate is arid and hot, and rainfall is scarce.

3. METHODOLOGY

The schematic diagram map of the classification of chaos detection methods is illustrated in Figure 1. A description of the methods is given below.

3.1 Classical Methods

Dimension analysis and dimensionality of a time series estimate the number of variables responsible for a sys–tem's dynamic evolution and the extent of their vari–ations. A non-negative integer value of the dimension implies the existence of a chaotic process, whereas a stochastic dynamic system produces an infinity. Grass–berger-Procaccia algorithm [11] is among the popular algorithms used to compute CD of time-sequenced data in weather systems. CD can be estimated either by equation (1) based on the box-counting method or by Grassberger and Procaccia using equation (2) [11].

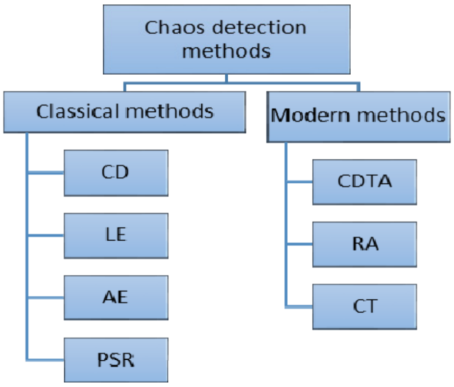


Figure 1. Methodological approach for the detection of the chaos in meteorological variables.

$$\text{Correlation dimension} := \lim_{\varepsilon \to 0} \frac{\ln \sum_{i=1}^{N_l(\varepsilon)} P_i^2}{\ln \varepsilon} \tag{1}$$

Correlation integral:

$$C(\varepsilon) := \lim_{N_l \to \infty} \frac{1}{N^2} \left\{ \left(x_i, x_j\right), \left\| x_i - x_j \right\| < \varepsilon \right\} \tag{2}$$

where N_l is the length of the analyzed signal, ε is the threshold distance, x_i, and x_j are the adjacent trajectories, and P_i is the associated probability.

However, LE measures the average exponential divergence of two adjacent trajectories (x_i *and* x_j). The largest exponent ($\lambda > 0$) is computed using equation (3).

$$\lambda_1 = \frac{1}{t_M - t_0} \sum_{k=1}^{M} \log_2 \frac{L'(t_k)}{L(t_{k-1})} \tag{3}$$

where M indicates the replacement steps $L(t_{k-1})$: Euclidean distance between two points within the vicinity of a set threshold and $L'(t_k)$is the length of $L(t_{k-1})t_k$ at k^{th}-time.

A high value of AE implies higher nonlinearity [54]. If 'X' $\in \{x_1, x_2, x_3 \dots, x_n\}$ random where '$p(x)$' is its associated probability, then entropy is defined as in equation (4):

$$H(X) = -\sum_{x \in X} p(x) \log p(x) \tag{4}$$

Unlike the first three methods (CD, LE, and AE), PSR is a qualitative visual tool representing a system's evo–lution, and its coordinates describe the state completely at any given time [55]. It is used for nonlinear systems and is based on the delay-embedding method. The delay-embedding technique converts a univariate time series to a multivariate system. The false nearest neighbor and mutual information estimate the embed–ding dimension and time delay, respectively. PSR plot captures the nonlinearity and unfolds the dynamics of the system. The complexity of many dynamical systems can be classified using PSR, including hydrologic systems [32].

3.2 Modern Methods

One of the modern methods CDTA used is developed in MATLAB [32]. It can be used for time series classifi–cation in diverse fields in real and simulated systems with noisy data. The Recurrence Plot and the recurrence quantification analysis are emerging data analysis techniques [52]. The 0-1 test for chaos is a relatively new method employed to detect chaos in time series data [14]. The classical methods of detecting chaos in time series are the correlation dimension, Lyapunov exponent, approximate entropy, and phase space recon–struction, including the 0-1 test for chaos. A Random Forest algorithm predicts rain/no rain as a predictive classification problem and wind speed prediction as a regression problem.

One relatively recent advanced data analysis tech–nique is the Recurrence Plot (RP) [16]. A recurrence plot is used to identify the nature of dynamical systems [55]. Multi-dimensional systems are projected using delay-embedding using a one-dimensional time series. The embedding unfolds complex dynamics [11, 56-57]. The Mutual Information algorithm is used to estimate an appropriate time delay τ [58]. The value for which the mutual information goes to zero is an optimum time delay. The recurrence analysis is the RP and RQ, which has been extensively applied in several disciplines, including weather sciences [8, 59].

To determine the chaotic nature of a deterministic dynamical system, the maximal LE 'λ' is estimated. If $\lambda > 0$, then it implies chaos, a proof of the divergence of trajectories [60]. This can be done directly from a time series without using the PSR method [14]. One draw–back of the classical methods is their high sensitivity to measurement noise [18]. The 0-1 test for chaos for univariate data uses equation (5), whose solution is given in equation (6) and is not sensitive to noise.

$$\begin{aligned} p(n+1) &= p(n) + \varphi(n) \cos cn \\ q(n+1) &+ q(n) + \varphi(n) \sin cn \end{aligned} \tag{5}$$

$$\begin{aligned} p_c(n) &= \sum_{j=1}^{n} \varphi(j) \cos jc \\ q_c(n) &= \sum_{j=1}^{n} \varphi(j) \sin jc \end{aligned} \tag{6}$$

where 'c' $\in (0, 2\pi)$, $\eta_n \in [-0.5, 0.5]$, and σ denote the level of noise. Two more tests are the z_1 test by Gottwald-Melbourne [14] and the practical test by Bensaida [61].

4. RANDOM FOREST FOR PRECIPITATION AND WIND SPEED PREDICTION

RFA is an extension of bagging for decision trees that can be used for classification or regression. In predictive analytics, this algorithm is considered a very effective learning model in data science. It is an ensemble of simple trees. Individual trees in the algorithm are developed, and randomized samples with their features are selected [62]. The accuracy is the ratio of the correct classification to its total number [63]. The letters T, P, F, and N are used for true, positive, false, and negative in TP, TN, FP, and FN, respectively. Classification Accuracy (CA) is given by equation (7):

$$CA = \frac{TP + TN}{N_s} \quad (7)$$

where, N_s is the total number of samples and misclassi–fication error = (1-classification accuracy). Cohen's Kappa Statistic (CKS) is given by equation (8):

$$CKS = \frac{\frac{TP}{N_s} + \frac{TN}{N_s} - A}{1 - A} \quad (8)$$

The value A can be determined using equation (9):

$$A = \left(\frac{FN+TP}{N_s}\right)\left(\frac{FP+TP}{N_s}\right) + \left(\frac{FP+TN}{N_s}\right)\left(\frac{FN+TN}{N_s}\right) \quad (9)$$

$$\text{Precision (positive predictive value)} = \frac{TP}{TP+FP} \quad (10)$$

$$\text{TP Rate (Sensitivity)} = \frac{TP}{TP+FN} \quad (11)$$

$$\text{FP Rate (1-specificity)} = \frac{FP}{FP+TN} \quad (12)$$

The system's performance is measured using equations (7) – (12). Air temperature, surface pressure, precipitation, and humidity are input to predict wind speed. The precipitation is predicted using the rest of the variables as inputs. Due to the chaotic nature of the wind speed, the prediction horizon is short, and any accurate long-term prediction is almost impossible.

5. RESULTS AND DISCUSSIONS

The site-dependent long-term statistics of the variables are listed in Table A, Appendix A. The coefficient of variation values for all variables is less than 1, indi–cating a low variance except for precipitation. The negative values of kurtosis for pressure and air tempe–rature imply a light-tailed distribution that lacks outliers or extreme values. The kurtosis values of all the variables are in the range [-2,2] except for the higher precipitation. The high kurtosis value for the precipi–tation indicates an outlier of 35.04 mm/day, whereas the number of zeroes is 92%. Kurtosis values of air temperature and surface pressure are negative, implying that the distribution is platykurtic. It indicates the pre–sence of less extreme data in tails, as shown in Figures A(a) and A(d). The histogram Figure 2(a) shows the distribution of the values and related frequencies.

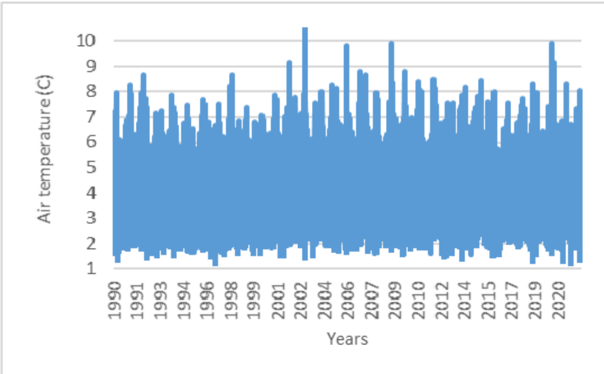
The value of skewness of air temperature, Table A, is negative, which implies a negatively skewed distributi–on. This can be observed from the histogram, Figure A(a), with a fatter tail on the left. However, the skew–ness values of all the other variables are positive, which indicates that the right tail is longer, as can be seen from the histograms, except for the surface pressure. The surface pressure has a bin at the extreme left, which is an outlier.

Mann-Kendall non-parametric trend test assessed trends in the time series data. The test failed to reject the null hypothesis at the alpha significance level of 0.05. Table 1 indicates the absence of a general trend in the data statistically, and the same can be observed in the time series plots, Figure 2 (a-e). Their statistics are provided in Table A, and further analysis is done using chaos theory. The absence of a trend in each data time series leads to an assumption that it is independent, identically distributed, and uncorrelated.

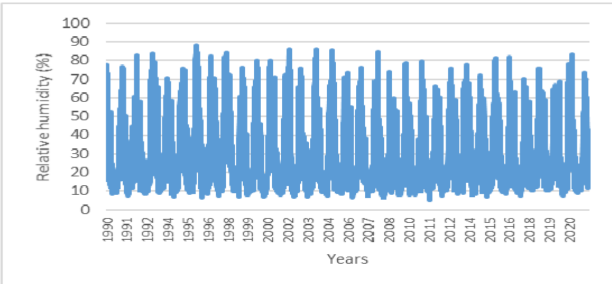
Table 1. Mann-Kendall non-parametric trend analysis test

Variable	p-value	Trend	α = 0.05
Air temperature (°C)	0.3223	No	p > α
Relative humidity (%)	0.6767	No	p > α
Precipitation (mm/day)	0.2160	No	p > α
Surface pressure (kPa)	0.0763	No	p > α
Wind speed (m/s)	0.8046	No	p > α

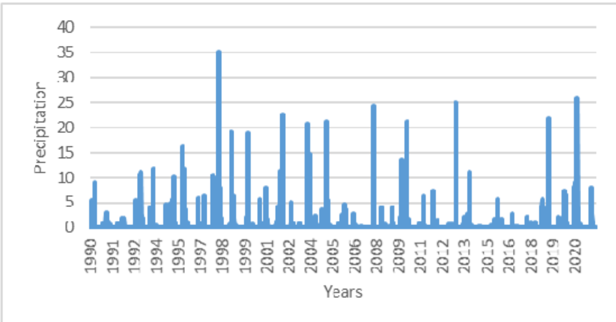
The annual maximum temperature of >30 □ from May to September and the lowest relative humidity of <20% are observed. Precipitation data show over 90% zeros indicating no rains, as is typical in the Arabian Peninsula. The precipitation's mean and corresponding standard deviation values are 0.084 and 0.961. It indicates a high variation; about 68 % of the data is within the range of one standard deviation of the mean, and it is a skewed distribution. The prediction results are provided in section 4.4.



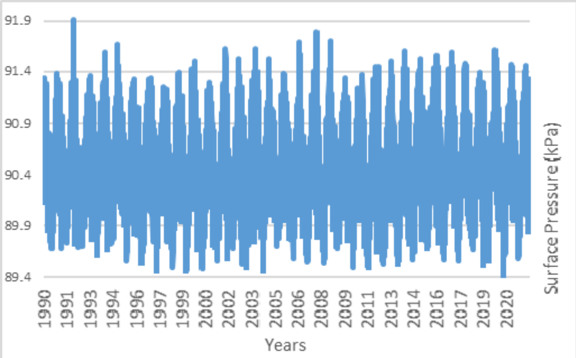
a) Time series plot of air temperature (C).



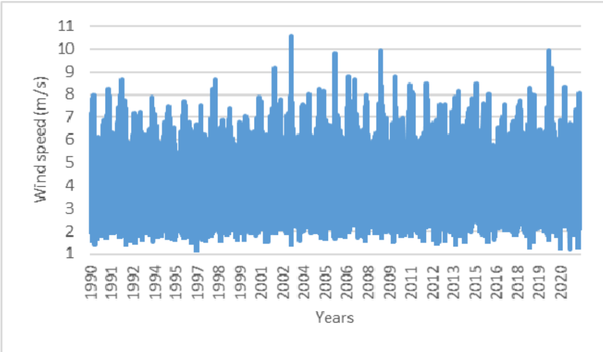
(b) Time series plot of relative humidity (%).



(c) Time series plot of precipitation (mm/day)



(d) Time series plot of surface pressure (kPa).



(e) Time series plot of wind speed (m/s).

Figure 2. (a-e). Time series data plots of the meteorological variables for the Ha'il region of Saudi Arabia (1990-2022).

The frequency distribution plots of the variables are provided, in Figure A, in the appendix.

6. CLASSICAL METHOD-BASED ANALYSIS

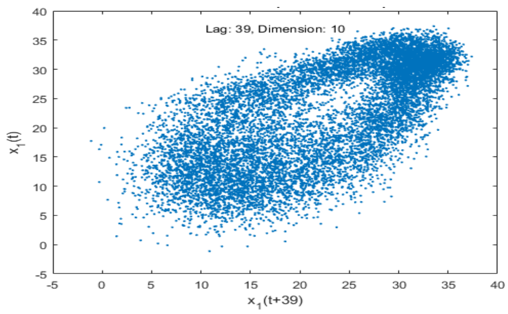
CD values range from 0.04 to 2.29, indicating low-dimensional chaos. The highest value is observed for surface pressure and the lowest for precipitation. These values confirm the result obtained by the CDTA, except that precipitation and surface pressure data show random behavior. A non-integer value indicates chaotic behavior in time series, suggesting that the dynamics are governed by two to three variables [10].

A positive value of LE represents chaos, as it is strong evidence of the chaotic nature of the system. The λ values are positive for three variables and negative for surface pressure and precipitation time-series data (Table 2). The precipitation and pressure are classified as stochastic by the CDTA. The value of λ is approximately 1.65 and 1.41 for air temperature and relative humidity, respectively. The inverse of λ is related to the systems' predictability of less than a day; however, the inverse of λ is associated with wind speed for one day. These values indicate the chaotic nature of the system and hence a rapid divergence of trajectories. A limited unpredictability horizon of chaotic systems is evidence of unstable behavior. The CD was used to detect chaos in the precipitation time series, and its value is 0.04 (Table 2) [27]. The calculated AE values lie in the range of [0.18, 2.10], as listed in Table 2. The lowest value corresponds to precipitation, whereas the highest value is related to wind speed. AE is related to inherent irregularity and lack of predictability in time series data.

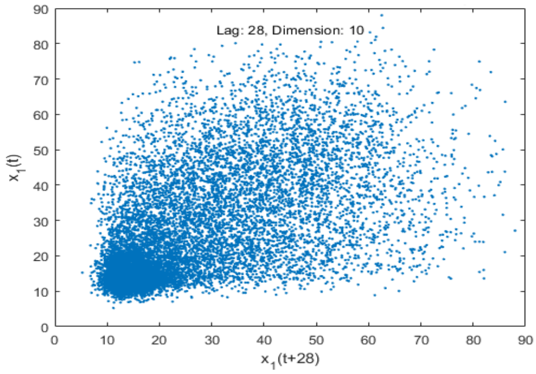
Table 2. The values of CD, LE, and AE.

Variables	CD	Largest LE	AE
Air temperature (°C)	1.75	1.65	0.73
Relative humidity (%)	1.37	1.41	1.18
Precipitation (mm/day)	0.04	-0.31	0.16
Surface pressure (kPa)	2.29	-0.63	1.39
Wind speed (m/s)	1.78	0.96	2.10

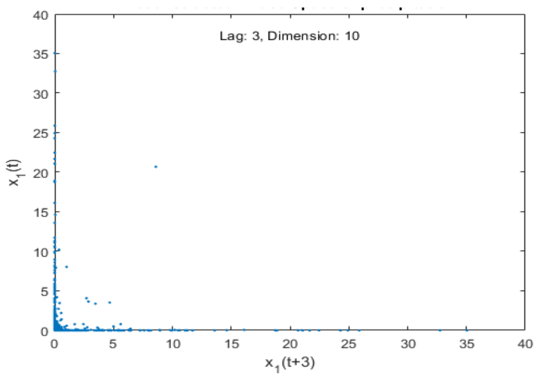
PSR of air temperature, relative humidity, precipi–tation, surface pressure, and wind speed (Figure 3a–e) are plotted using the estimated delays of 39, 28, 4, 3, 3, and embedding dimensions as 10, 10, 1, 10, and 10; respectively. The projection of a higher-dimensional time series into a phase space plot shows that the under–lying dynamics appear to be dimensionally low chaotic. The plots reveal the presence of attractor-like structures in their complex dynamics. This complexity can also be attributed to high and low points. They also show low dimensional complexity as CD values are less than 4. The scattering of the attractors and the presence of outliers can be visually detected [7].



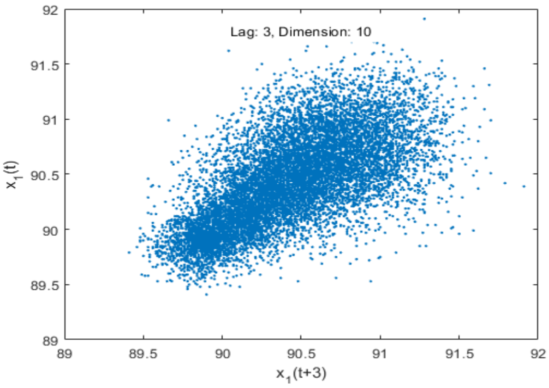
(a) Phase space reconstruction plot of air temperature.



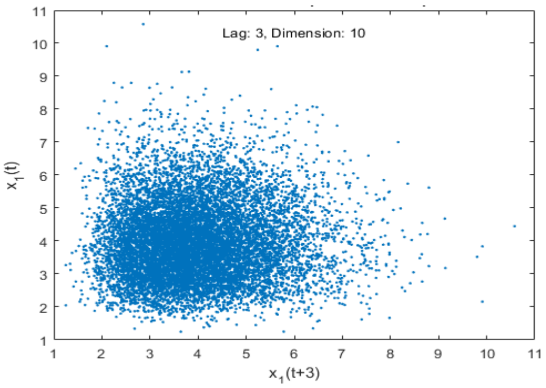
(b) Phase space reconstruction plot of relative humidity.



(c) Phase space reconstruction plot of precipitation.



(d) Phase space reconstruction plot of surface pressure.



(e) Phase space reconstruction plot of wind speed.

Figure 3 (a-e). Phase space plots based on daily mean time series data.

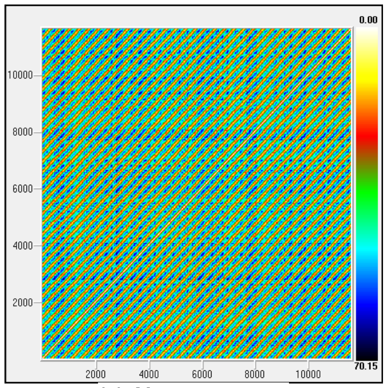
7. MODERN METHOD-BASED ANALYSIS

RP is generated based on the calculation of three parameters: embedding dimension, time delay, and the threshold value. False nearest neighbor provides the value of the embedding dimension, and it is chosen when points in a neighborhood drop to zero. If the embedding dimension is 'n = 4', then it implies that a model of 4 state variables can be developed. It is used to calculate an optimal embedding dimension [63]. Mutual information is used to estimate the time delay value [64]. However, various methods are employed to calculate a threshold value [26].

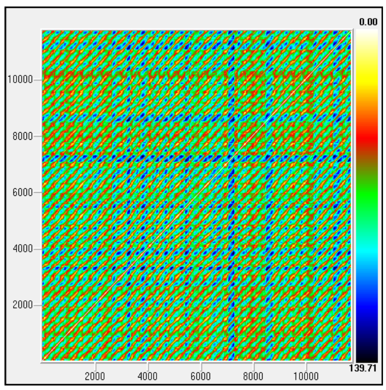
In this study, the method of 5-6% of the maximal space diameter is chosen for the threshold value, whe–reas the embedding dimension and time delay values are used to reconstruct PSR.

RPs are studied based on the presence of dots, lines, and homogeneous and non-homogeneous structures. Figures 4 (a–e) shows a variety of such structures. The structural patterns of parallel lines in Figures 4(a) and 4(b) for air temperature and relative humidity are more prominent, indicating that the trajectories will re-visit the region of attraction. RQA values are listed in Table 3, which shows a high recurrence rate for precipitation followed by humidity and temperature. The system is classified as either quasi-periodic or low-dimensional chaotic, as proven by their quantification using RQA and the methods of chaos detection. The surface pressure (Figure 4(d)) is considered random, as it has similar colors and is uniformly distributed, unlike other plots. However, the presence of parallel and vertical lines makes it difficult to classify it as completely random. LE describes the states as dynamical systems evolve [26]. The vertical lines can also be easily obser–ved, implying some either invariant states or very slow. No definite trend is observed in all the time series due to the absence of fading in the lower left and upper left corners, as shown in Figure 4(a–e). This finding corro–borates with the result obtained by Mann Kendall's trend analysis presented in Table 1. Higher values of determinism (DET) and laminarity (LAM) are also indicators of the presence of deterministic structure, except for the wind speed, which has lower values. The wind speed exhibits relatively higher complexity due to the smaller values of the RQA. The numerical values of the RQA are listed in Table 3.

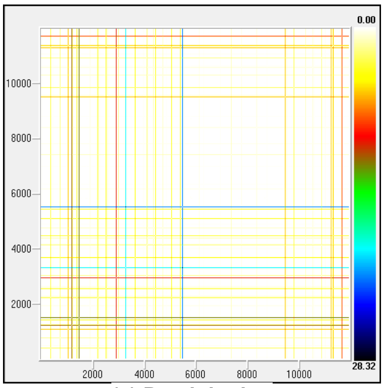
A difference in the RQA values of surface pressure, air temperature, wind speed, and relative humidity for recurrence rate (RR), DET, and LAM is listed in Table 3. The highest variation is observed in entropy values and the mean length (LMEAN). Relative humidity shows the highest LMEAN; which implies low complexity dynamics [16]. The wind speed has the smallest RQA values (RR, DET, LAM), indicating high complexity compared to the other variables. The magnitude of these values is related to their complexity in their underlying dynamics. The presence of only scattered points is an indication of a random system. Random systems imply a lack of predictability and a higher degree of irregularity [7, 65]. The RP of wind speed, air temperature, and surface pressure contain structures, as shown in Figure 4(a–e), which are somewhat similar to the periodic sinusoidal wave in spite of their chaotic nature. Based on visual textures and the higher RQA values of DET and LAM, these systems appear to be deterministic and have a possibility of short-term prediction of their future states.



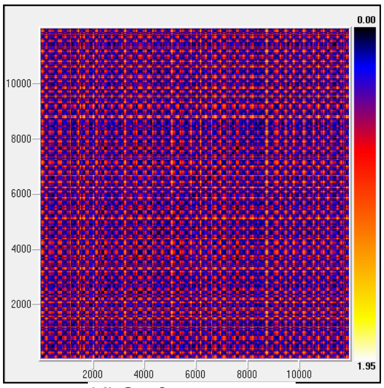
(a) Air temperature



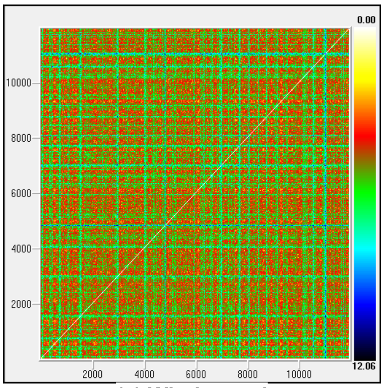
(b) Relative humidity



(c) Precipitation



(d) Surface pressure



(e) Wind speed

Figure 4. (a-e). The recurrence plot of the time series data for all the meteorological variables showing time versus time, [26], using visual recurrence analysis, Eugene 2003.

Table 3. Comparison of RQA values of pressure, tempera–ture, wind speed, humidity, and precipitation for the Ha'il region of Saudi Arabia

RQA	Air temperature (□)	Relative humidity (%)	Precipitation (mm/day)	Surface pressure (kPa)	Wind speed (m/s)
%RR	0.291	0.360	0.854	0.087	0.058
%DET	0.949	0.886	0.987	0.376	0.123
ENT	2.560	2.215	3.291	0.731	0.260
LAM	0.969	0.928	0.804	0.503	0.162
LMEAN	6.968	5.828	13.70	0.503	2.070

The nature of the time series data is further exa–mined and validated using CDTA [18]. The algorithm classified all variables as chaotic except the surface pressure and precipitation as stochastic (Table 4). The stochastic nature of these times series data are further tested using the multiple surrogate time series, and no difference was observed in the analysis.

Table 4. Characterization of time series data (daily mean values) using Chaos Tree Decision Algorithm, [40].

Time series	Classification	Permutation entropy	K-value
Air Temperature (□)	Chaotic	4.99	0.772
Relative humidity (%)	Chaotic	6.76	0.995
Precipitation (mm/day)	Stochastic	--	--
Surface Pressure (kPa)	Stochastic	--	--
Wind speed (m/s)	Chaotic	6.88	0.998

A test for LE-based chaos is used to detect the chaotic nature of the meteorological variables [61]. The test revealed negative high λ values indicating the absence of chaos at a 5% confidence level, which means the data is stochastic. The result of this test contradicts the result obtained from the CDTA in which the air temperature, relative humidity, and wind speed data are classified as chaotic. However, the results obtained for precipitation and surface pressure data validate the same results by both algorithms.

The 0 – 1 test for chaos concluded a chaotic nature of all the variables with λ values in the range of [0.9964, 0.9983]. A value of 0 implies non-chaotic data [14]. The oversampling of data can lead to smaller values of λ despite being chaotic. The data has no oversampling, as the plots do not look smooth (Figures 2(a–e)).

8. RFA-BASED PRECIPITATION AND WIND SPEED PREDICTION

The RFA determines the occurrence of "Precipitation" or "No precipitation". A train-test ratio of 80:20 (%/%) is used. The performance of the RFA model is obtained using the confusion matrix. The confusion matrix produced TP, TN, FP & FN values. Cohen's kappa statistic reported 10686 cases of no rain, and 820 of rain, a measure of how balanced the data set is. Table 5 shows the model's high performance with a classification accuracy of 92.96% and Cohen's Kappa Statistic of 0.325. The confusion matrix of the model presented in Table 6 shows that there are 162 misclassified instances. It also shows that the occurrence of precipitation could be better classified as a result of the bias of the dataset. Strikingly, 45 instances of precipitation were correctly predicted, while 127 needed to be corrected. Thus, the RF model can be considered reliable in terms of "No precipitation" and unreliable in terms of "Precipitation", [52]. However, there were two challenges, firstly, the issue of non-stationarity and secondly, nonlinearity that are inherent in precipitation data. Suggestions to improve the performance of RFA models for precipitation data may include the use of multiple output classes, such as "No precipitation", "Moderate precipitation", and "Heavy precipitation", based on the precipitation [67]. The classification accuracy is the ratio of the correct classification to the total number of classifications by RFA model [68].

In addition to predicting rainfall, the study also sought to investigate the performance of the RFA in forecasting wind speed. A train-test ratio of 70:30 (%/%) is used. The performance of the model is obtained using perfor–mance measures for regression.

Table 7 shows the poor performance of the model for wind speed with a correlation coefficient of 30%, MAE of 89.5%, and RMSE of 1.122.

Table 5. Performance of RFA model for precipitation

Performance measure	Model
Classification accuracy (%)	92.96
Misclassification error (%)	7.04
Cohen's Kappa statistic	0.325
Precision	0.914
TP rate	0.930
FP rate	0.684

Table 6. Confusion matrix for RFA model for precipitation

Actual class	Predicted class		Totals
	No precipitation	precipitation	
No Precipitation	2094	35	2129*
Precipitation	127	45	172#

*Correctly classified instances = 2139
#Incorrectly classified instances = 162

Table 7 Performance of RFA regression model

Performance measure	Model
Correlation coefficient	30%
MAE	89.5%
RMSE	1.122

A deeper understanding of the variation in global climate on a long scale provides important avenues to understand their impact on humans and in forecasting climate in the future [69-70].

The CD, LE, PSR, AE, RP, RQA, and tests for chaos are employed to investigate the nature of the dynamics of the meteorological variables. Daily mean values of climatic variables, such as air temperature, relative humidity, surface pressure, precipitation, and wind speed for the Ha'il region from 1990 to 2022, are used. The results showed positive values of CD, LE, and AE. The chaotic nature is detected in air temperature, relative humidity, and wind speed, whereas surface pressure and precipitation showed stochastic nature. The 0–1 test revealed negative LEs for all variables. This implies the stochastic nature, which contradicts the test obtained in the CDTA. Relatively smaller values of RQA for all variables, particularly for wind speed and its corresponding RP magnitudes shows a higher dimension complexity compared to the values of the other variables.

The maximum temperature reached was over 30 ^{o}C from May to September, whereas the lowest relative humidity of (<20%) was observed during the same period. No definite trends are observed as per Kendall's non-parametric trend analysis test. No precipitation is observed throughout the year except in January and April. The analysis showed that during 32 years, there was no precipitation for 92% of the total duration.

9. CONCLUSIONS

The diverse methods (classical and modern) of chaos detection are employed to understand the nonlinear dynamics and to classify the nature of meteorological variables (air temperature, relative humidity, surface pressure, precipitation, and wind speed). Some applications of the chaos theory in mechanical engineering have also been briefly described. The proposed techniques are used to identify and characterize their inherent dynamic structures.

Furthermore, two important variables (precipitation and wind speed) are predicted using RFA model. The obtained results are validated and compared with the reported values. Three methods (CD, AE, and CT) have classified all the variables as chaotic. LE classified pressure as non-chaotic, whereas other variables are chaotic. Temperature, humidity, and wind speed are classified as chaotic by CDTA method during precipitation, and surface pressure as stochastic. These results revealed that the classical methods are very well validated with the modern methods. On the flip side, some contradict modern methods.

The analysis also carries out the predictability of the precipitation and wind speed. It is shown that wind speed predictability is highly challenging. Similar results are obtained by recurrence analysis. The annual maximum temperature of >30 □ from May to September and the lowest relative humidity of <20% are observed. According to the Mann-Kendall test, no definite trends are observed for all the variables. The analysis for 32 years of data showed no precipitation for 92% of the time during the entire period based on the Random Forest algorithm.

Based on these results, some dynamic systems are so complex that drawing inferences about their classification becomes challenging, whether a system is purely chaotic, random, or periodic. Some systems may exhibit mixed behavioral characteristics, such as partially chaotic with some elements of randomness or periodicity with chaotic nature and weak chaos in dynamic variability. Future studies can be planned to explore more robust methods for high-dimensional meteorological variables. The study further recommends increasing forecasting efficiency for the optimum utilization of natural clean energy resources using the proposed methodology.

ACKNOWLEDGEMENTS

The authors acknowledge the academic support provided by the National University of Malaysia. No financial support was sought. They acknowledge the use of MATLAB and Visual Recurrence Analysis by Eugene (2003).

REFERENCES

[1] Petersen, A.: *Arabian Peninsula: Islamic Archaeology*, Encyclopedia of Global Archaeology. Springer, Cham. https://doi.org/10.1007/978-3-030-30018-0_564, In: Smith, C. (eds), 2020.

[2] Almulhim, T., and Yousif, M. Al.: An Analysis of Renewable Energy Investments in Saudi Arabia: A Hybrid Framework Based on Leontief and Fuzzy Group Decision Support Models, *Available at SSRN 4183286*, 2022.

[3] Zhong, K., Zheng, F., Qin, H.W.C.and Xu, X.: Dynamic changes in temperature extremes and their association with atmospheric circulation patterns in the Songhua River Basin, China, Atmos. Res., Vol. 190, pp. 77-88.

[4] Lorenz, E.: Deterministic nonperiodic flow. J. Atmos. Sci., Vol. 20, No. 2, pp.130-141, 1963.

[5] Tsonis, A.A. and Elsner, J.B.: Chaos, Strange attractors, and weather, Bulletin of the American Meteorological Society, Vol. 70, No. 1, pp. 14-23, 1989.

[6] Sivakumar, B., Liong, S.Y., Liaw, C.Y. and Phoon, K.K.: Singapore rainfall behavior: chaotic? Journal of Hydrologic Engineering, Vol. 4, No. 1, 38-48, 1999.

[7] Yan, B., Chan, P. W., Li, Q., Y. He., and Shu, Z.: Dynamic analysis of meteorological time series in Hong Kong: A nonlinear perspective, Inter. J. Climat., Vol. 41, No. 10, pp. 4920-4932, 2021.

[8] Panagoulia, D., Vlahogianni, E.I.: Recurrence quantification analysis of extremes of maximum and minimum temperature patterns for different climate scenarios in the Mesochora catchment in central-western Greece, Atmos. Res., Vol. 205, pp. 33–47, 2018.

[9] Ott, E.: *Chaos in Dynamical Systems,* Cambridge University Press, 2002.

[10] Sivakumar, B.: *Chaos in hydrology: bridging determinism and stochasticity*. Springer, 2016.

[11] Grassberger, P., and Itamar P.: Characterization of strange attractors, Phys. Rev. Lett., Vol. 50, No. 5, p. 346, 1983.

[12] Paul Matthews, 0–1 test for chaos (https://www.mathworks.com/matlabcentral /fileexchange/25050-0-1-test-for-chaos), MATLAB Central File Exchange. Retrieved February 23, 2023.

[13] Takens, F.:*Detecting strange attractors in turbulence*. In: Dynamical Systems and Turbulence, Warwick 1980. Berlin, Heidelberg: Springer, pp. 366–381, 2006.

[14] Georg, A., Gottwald, and Ian Melbourne.: On the validity of the 0–1 test for chaos, Nonlinear., Vol. 22, No. 6, p. 1367, 2009.

[15] Marwan, N., Romano, M.C., Thiel M. and Kurths, J.: Recurrence plots for the analysis of complex systems, Phys. Rep., Vol. 438, No. 5-6, pp. 237-329, 2007.

[16] Eckmann, J.P., Kamphorst, S.O. and Ruelle, D.: Recurrence plots of dynamical systems, EPL Europhys. Lett., Volume 4, No. 9, pp. 973-977, 1987.

[17] Zbilut, J.P. and. WebberJr, C.L.: Embeddings and delays as derived from quantification of recurrence plots, Phys. Lett. A, Vol. 171, No. 3-4, pp. 199-203, 1992.

[18] D. Toker, F.T. Sommer, & M. A. D'Esposito.: Simple method for detecting chaos in nature, Comm. Bio., Vol. 3, No. 1, p. 11, 2020.

[19] Hense, A.: On the possible existence of a strange attractor for the southern oscillation, Beitr. Phys. Atmos., Vol. 60, No. 1, pp. 34–47, 1987.

[20] Zeng, X., Pielkex, X.: What does a low-dimensional weather attractor mean?, Phys. Lett. A, Vol. 175, No. 5, pp. 299–304, 1993.

[21] Zeng, X., Pielke, RA., Eykholt, R.: Estimating the fractal dimension and the predictability of the atmosphere, J. Atmos. Sci., Vol. 49, No. 8, pp. 649-659, 1992.

[22] Islam, S., Bras, RL., Rodriguez-Iturbe, I.: A possible explanation for low correlation dimension estimates for the atmosphere, J. Appl. Meteor., Vol. 32, No. 2, pp. 203-208, 1993.

[23] Sivakumar, B., Jayawardena, A.W. and Fernando T.M.K.G.: River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches, J. Hydro., Vol. 265, No. 1–4, pp. 225-245, 2002.

[24] Rodríguez-Iturbe, I.: Exploring complexity in the structure of rainfall. Advances in Water Resources*,* Vol. 14, No. 4, pp. 162–167, 1991.

[25] Rodriguez-Iturbe, I., Power, FB De., Sharifi, MB., Georgakakos, KP.: Chaos in rainfall, Water Resour. Res., Vol. 25, No. 7, pp. 1667-1675, 1989.

[26] Sharifi, MB., Georgakakos, KP., Rodriguez-Iturbe, I.: Evidence of deterministic chaos in the pulse of storm rainfall, J. Atmos. Sci. Vol. 47, No. 7, pp. 888-893, 1990.

[27] Jayawardena, AW., Lai, F.: Analysis and prediction of chaos in rainfall and stream flow time series, J. Hydrol., Vol. 153, No. 124, pp. 23-52, 1994.

[28] Fraedrich, K.: Estimating weather and climate predictability on attractors, J. Atmos. Sci., Vol. 44, No. 4, pp. 722-728, 1987.

[29] Fraedrich, K.: Estimating the dimensions of weather and climate attractors, J. Atmos. Sci., Vol. 43, No. 5, pp. 419-432, 1986.

[30] Kawamura, A., McKerchar, AI., Spigel, RH., Jinno, K.: Chaotic characteristics of the Southern Oscillation Index time series, J. Hydrol., Vo4, 204, No. 124, pp.168-181, 1998.

[31] Delgado-Bonal, A., Marshak, A.: Approximate entropy and sample entropy: A comprehensive tutorial. Entropy 21, no. 6: 541, 2019.

[32] Sivakumar, B., Jayawardena, A.W. and. Li, W.K.: Hydrologic complexity and classification: a simple data reconstruction approach, Hydrol. Proc., Vol. 21, No. 20, pp. 2713-2728, 2007.

[33] Kaheman, K., Bramburger, J.J., Kutz, J.N.: Saddle transport and chaos in the double pendulum. Nonlinear Dyn 111, 7199–7233, 2023. https://doi.org/10.1007/s11071-023-08232-0.

[34] Jacquelin, E., Sondipon A., J-J. Sinou, Michael I. F.: Polynomial chaos expansion in structural dynamics: Accelerating the convergence of the first two statistical moment sequences. Journal of Sound and Vibration 356: 14-154, 2015.

[35] Zang, X., Iqbal, S., Zhu, Y., Liu, X. Zhao, J., Applications of chaotic dynamics in robotics, International Journal of Advanced Robotic Systems, 13(2), p.60, 2016.

[36] Beaumariage, T., Kempf. K. The nature and origin of chaos in manufacturing systems. In Proceedings of 1994 IEEE/SEMI Advanced Semiconductor

Manufacturing Conference and Workshop (ASMC), pp. 169-174. IEEE, 1994
[37] “An, X, Dongxiang J, Minghao Z., Chao L. Short-term prediction of wind power using EMD and chaotic theory. Communications in Nonlinear Science and Numerical Simulation 17, no. 2: 1036-1042, 2012.
[38] Zhang C, Wang X, Jiang J., Xinwei J., Xu X., Liang C., Liu H., Atindana, A. V.: Robust controller design of a semi-active quasi-zero stiffness air suspension based on polynomial chaos expansion. Journal of Vibration and Control: 10775463231153706, 2023.
[39] Liu, X., Gao, H., Ward, J.E., et al.: Power genera–tion from ambient humidity using protein nano–wires, Nat., Vol. 578, No. 7796, pp. 550–554, 2020.
[40] Kumar, A. *et al.:* Power Generation with Raindrops. In: Li, X., Rashidi, M.M., Lather, R.S., Raman, R. (eds*) Emerging Trends in Mechanical and Industrial Engineering*. Lecture Notes in Mechanical Engineering. Springer, Singapore, 2023.
[41] Rašuo, B.: On solving boundary value problem in fluid mechanics by Fourier’s method: Wall Interference of Transonic Wind Tunnels. In Analysis and Simulation of Multifield Problems, Springer Berlin Heidelberg, pp. 317-322, 2003.
[42] Rašuo, B., Dinulović, M., Veg A., Grbović A., and Bengin A.: Harmonization of new wind turbine rotor blades development process: A review, Ren. Sust. Ener. Rev., Vol. 39, pp. 874-882, 2014
[43] Ocokoljić, G., Rašuo B., and Kozić M.: Supporting system interference on aerodynamic characteristics of an aircraft model in a low-speed wind tunnel, Aerosp. Sci. Technol. Vol. 64, pp. 133-146, 2017.
[44] Nuha, H., M. Mohandes, and S. Rehman. "Vertical wind speed extrapolation using regularized extreme learning machine." FME Transactions 50, no. 3: 412-421, 2022.
[45] Hussain, Faisal M., Shafiqur Rehman, and Fahad A. Al-Sulaiman. "Performance analysis of a solar chimney power plant for different geographical locations of Saudi Arabia." FME Transactions 49, no. 1: 64-71, 2021.
[46] Gupta, D., Vikas K., Ishan A., M. Vasudevan, N. Natarajan. "Short-term prediction of wind power density using convolutional LSTM network." *FME Transactions* 49, no. 3: 653-663, 2021.
[47] Mohandes, M., S. Rehman, H. Nuha, M. S. Islam, and F. H. Schulze. "Accuracy of wind speed predictability with heights using Recurrent Neural networks." *FME Transactions* 49, no. 4: 908-918, 2021.
[48] Nuha, H., M. Mohandes, and S. Rehman. "Vertical wind speed extrapolation using regularized extreme learning machine." *FME Transactions* 50, no. 3: 412-421, 2022.
[49] Sivakumar, B. "Chaos theory in hydrology: important issues and interpretations." *Journal of hydrology* 227, no. 1-4 (2000): 1-20.
[50] Zeng, X, Roger A. P, Eykholt R. Chaos theory and its applications to the atmosphere. Bulletin of the American Meteorological Society 74, no. 4: 631-644, 1993.
[51] Marwan N, Wessel N, Meyerfeldt U, Schirdewan A, Kurths J.: Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Phys. Rev. E; 66:026702-1–026702-8, 2002.
[52] Marwan, N, and Kraemer, K. H.: Trends in recurrence analysis of dynamical systems. The European Physical Journal Special Topics: 1-23, 2023.
[53] https://power.larc.nasa.gov/data-access-viewer (Retrieved in February 2023)
[54] Pincus, S.M., Goldberger, A.L.: Physiological time-series analysis: what does regularity quantify?, Am. J. Physiol. Heart Circul Physiol*.*, Vol. 266, No. 4, pp. H1643–H1656, 1994.
[55] Kirichenko, Lyudmyla, Tamara Radivilova, and Juliia Stepanenko. "Applying recurrence plots to classify time series." Computational linguistics and intelligent systems, 2021 (2021): 16-26.
[56] Packard, N.H., Crutchfield J.P., Farmer, J.D. and Shaw, R.S.: Geometry from a time series, Phys. Rev. Lett., Vol. 45, No. 9, p. 712, 1980.
[57] Abarbanel, H. D. I.: *Tools for analysing observed chaotic data. Smart Structures, Nonlinear Dyna–mics, and Control*. Prentice Hall, pp. 1-86, 1995.
[58] Fraser, A. M., and Swinney H.L.: Independent coordinates for strange attractors from mutual information, Phys. Rev. A, Vol. 33, No. 2, p. 1134, 1986.
[59] Zhenru, S., Jesson, M., Sterling, M.: Nonlinear dynamic analysis of daily rainfall variability across the UK from 1989 to 2018. J. Hydrol., Vol. 603, p. 126849, 2021.
[60] Kantz, H., & Schreiber, T.: *Nonlinear time series analysis* (Vol. 7). Cambridge university press, 2004.
[61] BenSaïda, A.: A practical test for noisy chaotic dynamics, SoftwareX, Vol. 3, pp.1-5, 2015.
[62] Yu, P.S., Yang, T.C., Chen, S.Y., Kuo, C.M. and Tseng, H.W.: Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting, J. Hydrol., Vol. *552*, pp. 92-104, 2017.
[63] Kennel, M. B., Brown R., and Abarbanel H.D.: Determining embedding dimension for phase-space reconstruction using a geometrical construc–tion, Phys. Rev. A, Vol. 45, No. 6, p. 3403, 1992.
[64] Kraskov, A., Stögbauer H., Grassberger, P. Estimating mutual information. Physical review E 69, no. 6: p. 066138, 2004.
[65] Rehman, S., Siddiqi, A. H.: Wavelet-based Hurst exponent and fractal dimensional analysis of Saudi climatic dynamics, Chao. Soliton. Fract., Vol. 40*,* No. 3, pp. 1081-1090, 2009.
[66] Ali, M., Prasad, R., Xiang, Y., and Yaseen, Z.M.: Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts, J. Hydrol., Vol. *584*, p. 124647, 2020.
[67] Pham, Q.B., Yang, T.C., Kuo, C.M., Tseng, H.W. and Yu, P.S.: Combing random forest and least

square support vector regression for improving extreme rainfall downscaling, Water, Vol. 11, No. 3, p. 451, 2019.

[68] Sanni-Anibire, M.O., Zin, R.M. and Olatunji, S.O.: Machine learning model for delay risk assessment in tall building projects, Inter. J. Const. Manag., Vol. 22, No. 11, pp. 2134-2143, 2020.

[69] Rehman, S.: Temperature and rainfall variation over Dhahran, Saudi Arabia, (1970–2006). Int. J. Climatol., Vol. 30, No. 3, pp. 445-449, 2010.

[70] Mihailović, D. T., Mimić G., and Arsenić I..: Climate predictions: The chaos and complexity in climate models, Adv. Meteorol., 2014.

NOMENCLATURE

AE	approximate entropy
°C	Degree Celsius
CA	Classification Accuracy
CD	Correlation dimension
CDTA	Chaos decision tree algorithm
CDTA	Chaos decision tree algorithm
CKS	Cohen's Kappa statistic
CT	Tests for chaos
DET	Determinism
ENT	Entropy
FN	False negatives
FP	False positives
kPa	kilopascal
LAM	Laminariy
LE	Lyapunov exponent
LMEAN	The average length of the diagonal line
m/s	Meter per second
MERRA	Modern-Era Retrospective Analysis for Research and Applications
mm/day	Millimetre per day
PSR	Phase space reconstruction
RA	Recurrence analysis
RF	Random forest
RFA	Random forest algorithm
RP	Recurrence plot
RQA	Recurrence quantification analysis
RR	Recurrence rate
TF	True positives
TP	True positives

Appendix-A

Table A. Descriptive statistics of the time series data

Meteorological variable	Mean	Standard Deviation	Coefficient of variation	Variance	Kurtosis	Minimum	Maximum	Skewness
Air temperature (C)	22.28	8.73	0.389	76.14	-1.18	-1.12	37.38	-0.290
Relative humidity (%)	27.16	15.81	0.582	250.06	0.30	5.25	88	1.037
Precipitation (mm/day)	0.08	0.92	11.09	0.84	505.28	0	35.04	20.087
Surface pressure (kPa)	90.42	0.40	0.004	0.16	-0.56	89.41	91.91	0.208
Wind speed (m/s)	3.99	1.16	0.290	1.34	0.59	1.25	10.58	0.662

Figure A (a-e):

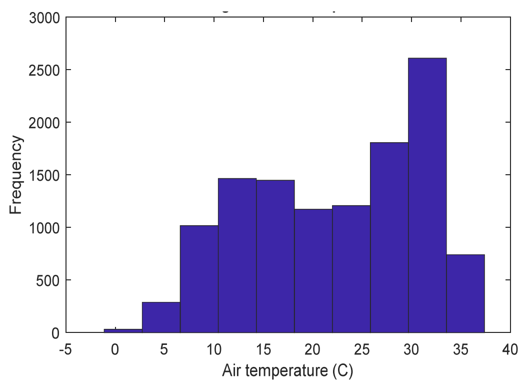


Figure A(a). Histogram of air temperature

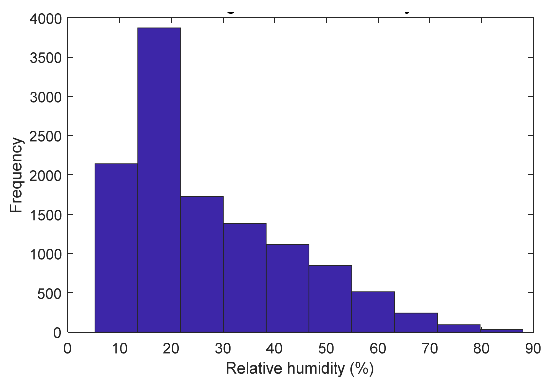


Figure A(c). Histogram of relative humidity

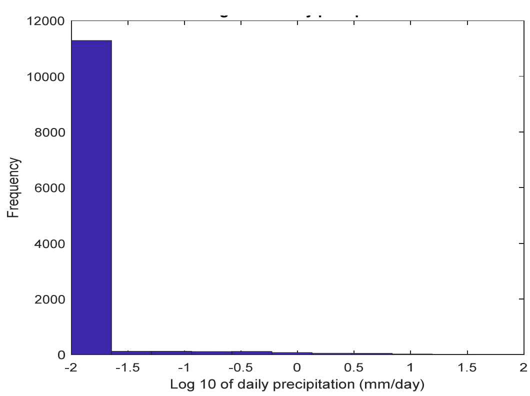


Figure A(b). Histogram of precipitation

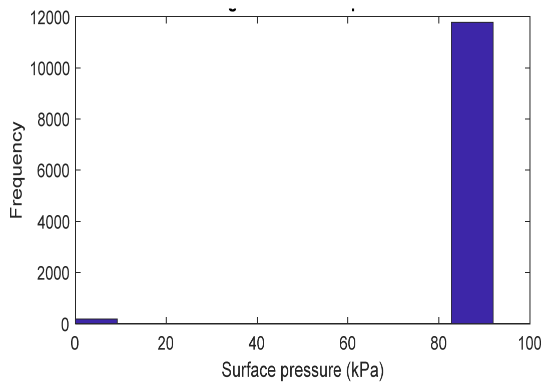


Figure A(d). Histogram of surface pressure

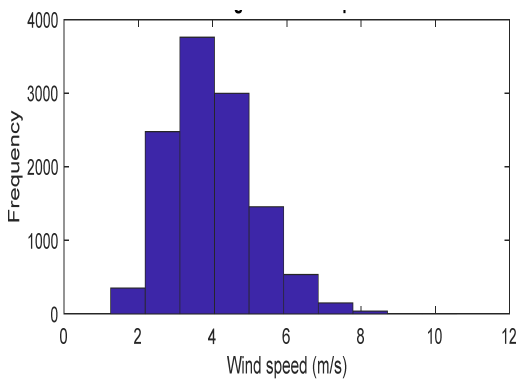


Figure A(e). Histogram of wind speed

НЕЛИНЕАРНА ДИНАМИЧКА АНАЛИЗА МЕТЕОРОЛОШКИХ ПРОМЕНЉИВИХ ЗА РЕГИОН ХАИЛ, САУДИЈСКА АРАБИЈА, ЗА ПЕРИОД 1990-2022

М.А. Мациd, М.С.М. Нурани, Ф.А. Разак

Студија примењује различите методе детекције хаоса на метеоролошке променљиве податке (тем–пература ваздуха, релативна влажност, површински притисак, падавине и брзина ветра за Хаил, Саудијска Арабија) да би се разумела нелинеарна динамика и класификовала њихова природа. Поред тога, модел алгоритма Рандом Форест се користи за предвиђање падавина и брзине ветра. Упоређени су резултати добијени класичним и савременим приступима. Утврђено је да су све променљиве хаотичне на основу димензије корелације, приближне ентропије и теста 0–1. Алгоритам дрвета одлучивања о хаосу дијагностикује температуру ваздуха, релативну влажност и брзину ветра као хаотичне, док су падавине и површински притисак идентификовани као стохастички. Ово показује да су класичне методе добро потврђене са савременим методама. Ипак, неки од њих су у супротности са савременим методама. Анализа података за 32 године показала је да није било падавина у 92% времена током читавог периода на основу алгоритма Рандом Форест.