Journal of AL-Rafidain University College for Sciences (2023); Issue 54; 418 - 433

AL- Rafidain University College

PISSN: (1681-6870); EISSN: (2790-2293)

Journal of AL-Rafidain University College for Sciences

Available online at: https://www.jrucs.iq

JRUCS

Journal of AL-Rafidain University College for Sciences

Statistical Modeling of Temperatures in Iraq Under a Fuzzy Environment

Bashar K. Ali	Mehdi W. Naserallah
bashar.alhallaq@mizan.edu.iq	mehdi.wahab@uokerbala.edu.iq
Division of Health and Vital Statistics-Department of Planning and Human Resources - Babylon Health Directorate - Ministry of Health & Environment, Babylon, Iraq.	Department of Statistics - Faculty of Administration and Economics - Kerbala University, Kerbala, Iraq

Article Information

Article History:
Received: December, 14, 2022
Accepted: February, 28, 2023
Available Online: December, 31, 2023

Keywords:
Fuzzy information system, Estimation, Temperatures, Fitting data, Linear model, Logarithmic model, Inverse model, Quadratic model, Cubic model, Compound model, S model Power model, Exponential model, Logistic model.

Correspondence:
Bashar K. Ali
bashar.alhallaq@mizan.edu.iq

Abstract

In this article, we modeling a set of data which represent the temperatures per day in the governorates of Iraq, which were taken from the General Authority for Meteorology and Seismic Monitoring for the year (2021) under variety statistical models, namely, linear, logarithmic, inverse, quadratic, cubic, compound , S, power, Growth , exponential, and logistic model by using classical principle and fuzzy principle by building a fuzzy information system under vary values of Alfa-cuts to generate membership values to the set of Temperatures to obtain a classical set that takes into account the inaccuracy in data collection , significance of the models was testing by the probabilistic value Sig. to reach to the best model that represents the data of Temperatures , we are compare among them by using mean square error (MSE). We are concluded that the use of the principle of fuzziness in the fitting of the models led to an increase in the accuracy of these models, and the mean squares error (MSE) for all the models that have been fitted is reduced on whether the data are traditional. We are also note that the best model in representing the Temperatures data with is the is power model to having it the lowest (MSE) among all the models, followed by the, Compound, Exponential Growth models at all α-cut coefficients (0.2, 0.4, 0.7), and that the rest models are not suitable for data on the number of Temperatures in Iraq, and we also note that the best model that achieved a fit for the data was at the α-cut = 0.7 (MSE= 0.32). We are notice that increase one unit of time led to increase temperature with (0.437) degree, That indicate after two years, the average daily temperature in Iraq will be (33.45).

https://doi.org/10.55562/jrucs.v54i1.610

1. Introduction

In applied statistics, a model, formula, or collection of formulas or models are created using previous information and experiences to give a more accurate and realistic portrayal of an occurrence in the actual world. The goal of the statistician is to produce the most precise conclusions and predictions regarding that phenomena. Insofar as it is easy to comprehend and use, the statistical model, which is a mathematical representation of a situation in the actual world, aims to condense and make sense of the data in order to be as realistic as possible. The researcher's objective is frequently to develop a model using data that has been gathered from the real world while retaining as much of the original data as possible. Our ability to perceive the world and think deeply about it is still limited, and as a result, we find ourselves in the face of uncertainty (inaccuracy) everywhere and at all times that results from lack of information. As a result, I believe that human knowledge is becoming increasingly important in today's society. Human knowledge is gained from our experiences of the world in which we live and the use of our thinking abilities to create a mass of information (observations). Fuzzy ideas, like "long," "far more," "little," etc., infiltrate our perceptions of the real world because they lack a categorically defined boundary. They are both partially true and partially untrue. These are what are known as fuzzy or ambiguous concepts (Fuzzy), which the human brain uses but which computers cannot. The more closely a person examines a problem in the actual world, the more confused they feel about how to solve it (Zadeh, 1973, 28). Fuzzy logic is one of the mathematical tools and contemporary methodologies that has proven to be extremely effective at solving issues on a big scale in a variety of application areas. As a result, this logic has expanded to touch the majority of contemporary technological elements. By assuming that humans do not represent classes of things (such as the class of bald men or the class of numbers much greater than 50) as entirely separate but rather as groups in which they may have different degrees of affiliation, it is one of the most useful ways of simulating human experience in a realistic manner. The fuzzy logic theory, which is a solution to the issue of representing approximate information by focusing on deduction through expressions, came to fill major gaps in the traditional logic (crisp) when inferring in uncertain and inaccurate circumstances, as there are many phenomena dealing with inaccurate and not clearly defined information. Linguistically, By providing the degree of belonging (membership) to any element in the set inside the real field [0,1], this degree denotes the degree of belonging of the element to the fuzzy subset. In this research, we discuss the average daily temperatures in the Iraqi governorates for the year 2021, which were obtained from the General Authority for Meteorology and Seismic Monitoring., namely, linear, logarithmic, inverse, quadratic, cubic, compound , S, power, Growth , exponential, and logistic model for data that representing the Temperatures in Iraq by using classical principle and fuzzy principle by building a fuzzy information system under vary values of Alfa-cuts to generate membership values to the set of Temperatures to obtain a classical set that takes into account the inaccuracy in data collection.

2. Statistical Models:[20]

Response variable (the dependent variable) and independent variable are two different sorts of variables that may be included in the statistical model. It is the study's goal variable and the model's output that the researcher wishes to analyze. The model may include one, two, or many explanatory variables that are measured or chosen by the researcher. They act as the model's inputs and aid in illuminating the relationship between changes in the independent variables and the response variable, or dependent variable. This is a critical step in the modeling process, and one of the key components of a good model is how well it captures the phenomenon that is being studied. Linear models and nonlinear models are the two categories into which statistical models fall :

2.1. Linear models: [4][11]

A continuous response variable is described by a linear model as a function of one or more predictor variables. They can help you comprehend and forecast the behavior of complicated systems as well as analyze experimental, financial, and biological data. Utilizing the statistical method of linear regression, a linear model is created. A dependent variable y (also known as the response) and one or more independent variables Xi are related in the model (called the predictors). The Simple Linear Regression model is used when there is only one independent variable, and the Multiple Linear Regression model is used when there are two or more independent variables.

$y = a_0 + \sum biXi + ei$ (1)

Or

$Y_i = a_0 + b_1 X_{i1} + b_2 X_{i2} + \cdots . + b_K X_{ip} + e_i$; i=1,2,….., n (2)

Where b stands for the estimated linear parameters, ei is the random error, and we assume that it has a normal distribution with a mean of zero and a constant variance. The errors are also independent of one another, thus we may mathematically define them as follows:

$e \sim N(0, \sigma^2{}_I)$, $cov(e_i, e_j) = 0$ (3)

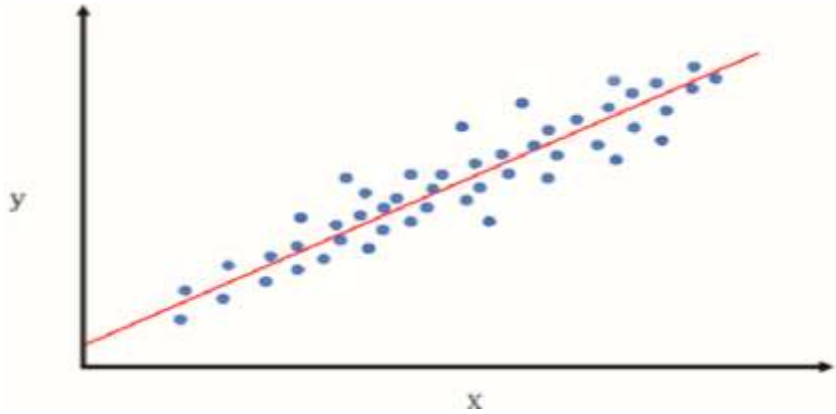

Figure (1): Liners Model curve

2.2. Nonlinear Models: [18]

Nonlinear models are of high importance, despite the scarcity of studies related to them compared to linear models, but they have wide applications in applied and natural studies. There are some phenomena that are subject to a general nonlinear trend, meaning that the data of these phenomena in the first place take different nonlinear forms, unlike the phenomena that take the general linear trend. The description of these phenomena is relatively complex using general linear trend methods, It necessitates looking for different kinds of curves in order to resolve this issue.

In statistics, nonlinear regression is a type of regression analysis in which observational data are represented by a function that depends on one or more independent variables and is a nonlinear combination of the model parameters. In order to fit the data, a method of sequential approximations is used.

The following are some examples of nonlinear models:

2.2.1. Logarithmic Model: [13]

For the purpose of predicting y-values that fall inside (interpolate) or outside (extrapolate) the plotted values, the logarithmic regression equation, often known as the Ln regression model (extrapolate). Regressions using logarithms are frequently used to model environmental data. A logarithmic model's general equation is:

$y = \beta_0 + \sum \beta i ln(Xi) + \epsilon i$ (4)

It is extremely common to logarithmically convert the variables in a regression model when there is a non-linear connection between the independent and dependent variables. The linear model is maintained even when the effective connection is non-linear by employing the logarithm of one or more variables rather than their un-logged form.. By making logarithmic adjustments, a highly skewed value can be transformed into one that is more like to normal. (In reality, there is a distribution called the log-normal distribution, which is described as having a skewed untransformed scale but a regularly distributed logarithm.

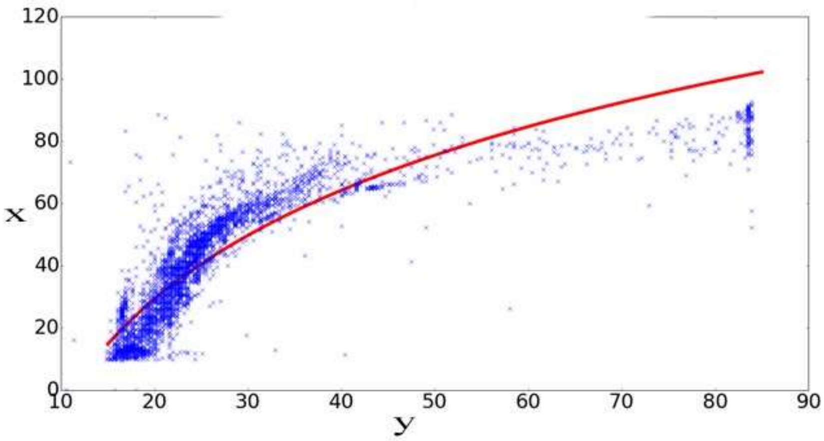

Figure (2): Logarithmic Model curve

2.2.2. Inverse Model: [6]

Inverse regression's core idea is to restrict the distribution of the input signal. The most prevalent presumption that is made is the linear design condition mfor any $a \in R^n$, there exist $c_0 \in R$ and $a \in R^d$ such that,

$$E\left(a^T \varphi(t) \backslash B^T \varphi(t)\right) = c_0 + c^T B^T \varphi(t) \tag{5}$$

It can be shown that the aforementioned condition is satisfied when the repressors are elliptically distributed or when the level curves of the relevant probability density function are ellipsoids. Li and Wang (2007) present a more in-depth study of the assumption's implications, which are not yet fully understood. The basic theorem for first order moment-based inverse regression algorithms is now ready to be stated. Should the system be supplied by:

$$y(t) = f\left(B^T \varphi(t), e(t)\right), \quad B \in R^{n_b \mathrm{x} d} \tag{6}$$

$$E(\varphi(t) \backslash y(t)) - E\left(\varphi(t)\right) \in \Re \quad (Cov(\varphi(t)B) \tag{7}$$

Where the operator $\Re$ retunes the column space of its argument.

The repressors are transformed to have a zero mean and an identical covariance matrix in the following on the assumption that they are standardized.

$$\varsigma(t) =^{\Delta} \ \ulcorner^{-1}\left(\varphi(t)\right) - E\left(\varphi(t)\right) \ , \tag{8}$$

where we've specified $\ulcorner =^{\Delta} \ Cov\left(\varphi(t)\right)^{\frac{1}{2}}$ (7) may now be expressed as:

$$E(\varsigma(t) \backslash y(t)) \ \in \Re \ \ (\ulcorner B) \tag{9}$$

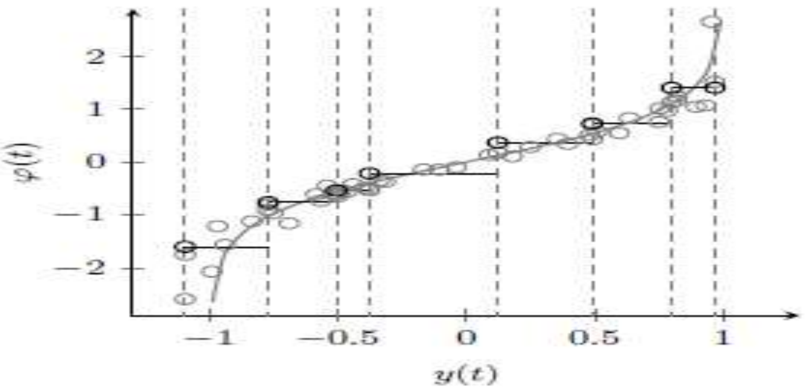

Figure (3): Inverse Regression Model curve

2.2.3. Quadratic Model: [9]

The second-degree quadratic equation curve is also known as the parabolic trend equation because it resembles a parabola when it is opened from the top, bottom, right, or left as seen in the figure. The equation's generic formula is:

$$y = a_0 + bX + b_2X^2 + \varepsilon \qquad (10)$$

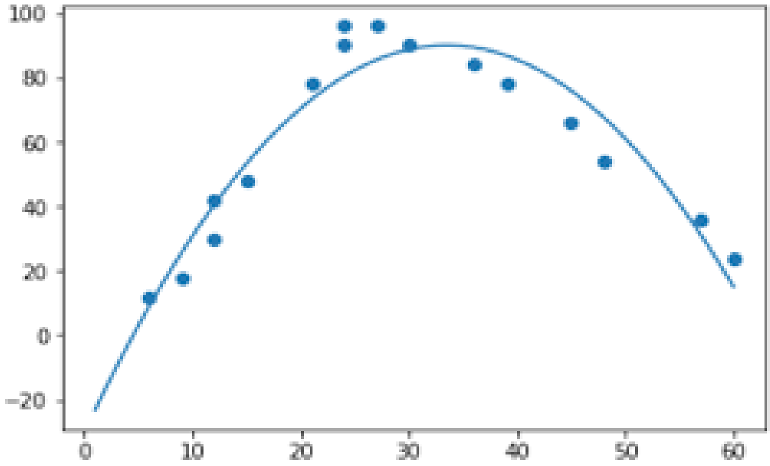

Figure (4): Quadratic Regression Model curve

2.2.4. Cubic Model: [16]

Equation of the third-degree curve (cubic Equation), The equation's general formula is as follows:

$$y = a_0 + b_1X + b_2X^2 + b_3X^3 + \varepsilon \qquad (11)$$

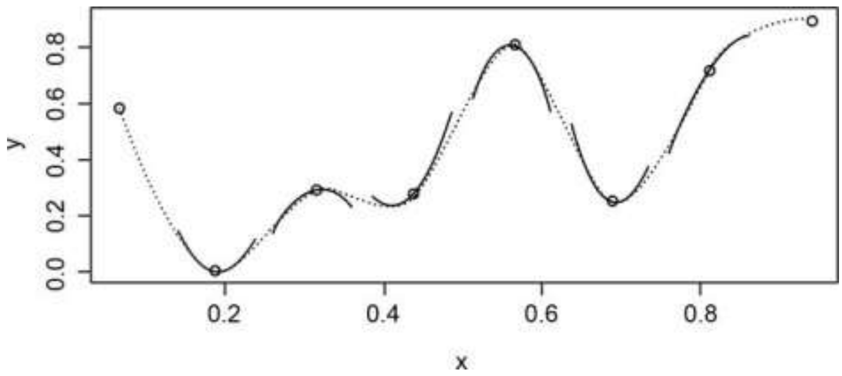

Figure (5): Cubic Regression Model curve

2.2.5. Compound Model: [14]

Where variation only occurs in the Y or X direction, one would minimize the sum of squared distances along the vertical or horizontal axis for the OLS on Y and X separately to obtain the best regression line for each case. Naturally, one would wish to find a regression line where Y and X are equally random. $Y = \alpha 0 + \beta_1 X$ that will cut down on both directions of fluctuation. As shown in Figure (1) below, this may be done by minimizing the weighted average of the squared vertical and horizontal distances as follows:

$$SS_y = y\sum_{i=1}^{n}(Y_i - \hat{Y}_i)2 + (1-y)\sum_{i=1}^{n}(X_i - \hat{X}_i)2$$

$$= y\sum_{i=1}^{n}(Y - a_0 + \sum biXi)^2 + (1-y)\sum_{i=1}^{n}(X_i - \frac{Y_i - b0}{\sum bi})^2 \quad ; \ 0 \leq y \leq 1 \qquad (12)$$

The OLS on Y or X is obtained at the two extreme values of y1 and y0, respectively. The regression's least squares estimators may be obtained for each Parameters by solving $\frac{\partial SS_r}{\partial a0} = 0$ and $\frac{\partial SS_r}{\partial \sum bi} = 0$ simultaneously.

Let $S_{xx} = \sum_{i=1}^{n}(Xi - \overline{X})^2$, $S_{YY} = \sum_{i=1}^{n}(Yi - \overline{Y})^2$, and $S_{xy} = \sum_{i=1}^{n}(Xi - \overline{X})\,(Yi - \overline{Y})$ that the generated estimators from the compound regression model $\hat{b}0$ and $\hat{b}i$ would be content**:**

$$\hat{a}_0 = \overline{Y} - \sum bi\overline{X} \tag{13}$$

$$\hat{b}i = \frac{y}{1-y}\sum bi^4 S_{xx} - \frac{y}{1-y}\sum bi^3 S_{xy} - S_{YY} \tag{14}$$

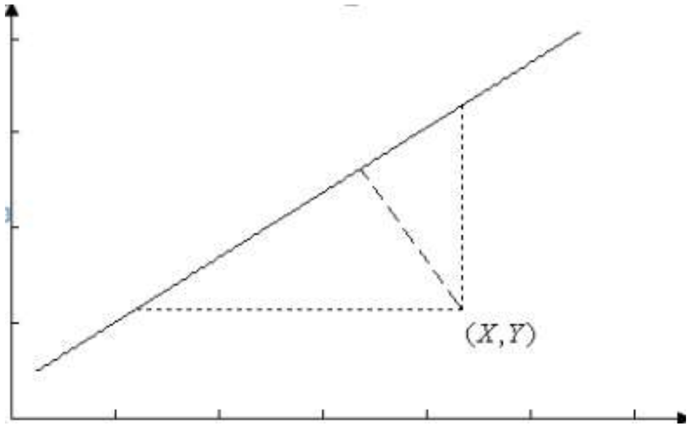

Figure (6): Compound Regression Model curve

2.2.6. S Model: [12]

This model can be easily recreated S regression by using the logistic function as following :

$$\mathrm{f(x)} = \frac{\mathrm{L}}{1+e^{k(x-x_0)}} \tag{15}$$

Where e is the natural logarithmic base , x_0 the value of the sigmoid midpoint, the curves maximum value, k the steepness of the curve.

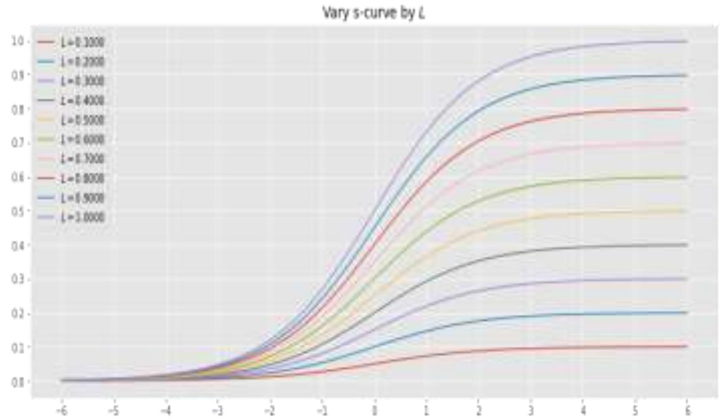

Figure (7): S-shape Regression Model curve

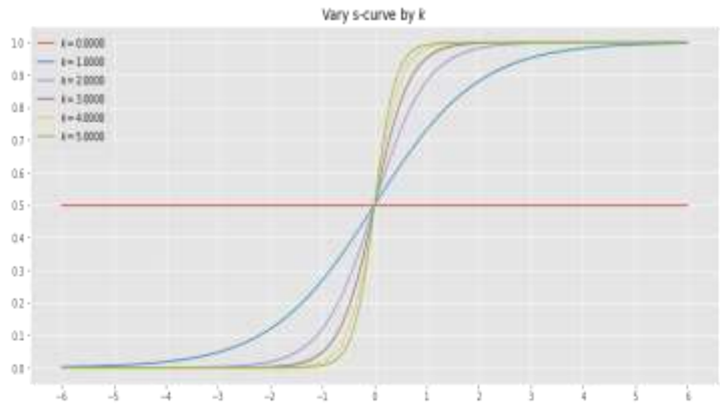

Figure (8): S-shape Regression Model curve

2.2.7. Power Model: [2]

The power regression model is another non-linear regression model and it is based on the equation:

$$y = a_{\mathbf{0}}\boldsymbol{X}^{b} \tag{16}$$

When we take the natural log of both sides of the equation, we get the following equation:

$$\ln(y) = \ln(b_{\mathbf{0}}) + b\ln(x) \tag{17}$$

This equation contains a linear regression model's basic structure, and I've included an error component, represented by the letter e:

$$y' = a'\mathbf{0} + bx' + e \tag{18}$$

A log-log regression model is one that has the equation ln y = b ln x + δ . Given that this equation is true, we have:

$$\begin{aligned} y &= \boldsymbol{e}^{\boldsymbol{lny}} = \boldsymbol{e}^{b\ln(x)+\delta} = \boldsymbol{e}^{b\ln(x)}\boldsymbol{e}^{\delta} \\ &= \left(\boldsymbol{e}^{\ln(x)}\right)^{b}\boldsymbol{e}^{\delta} = \boldsymbol{e}^{\delta}X^{b} \end{aligned} \tag{19}$$

Any such model may therefore be written as a power regression model of type. y = $a_{\mathbf{0}}$ X^{b} by setting $a_{\mathbf{0}}$ = $\boldsymbol{e}^{\delta}$.

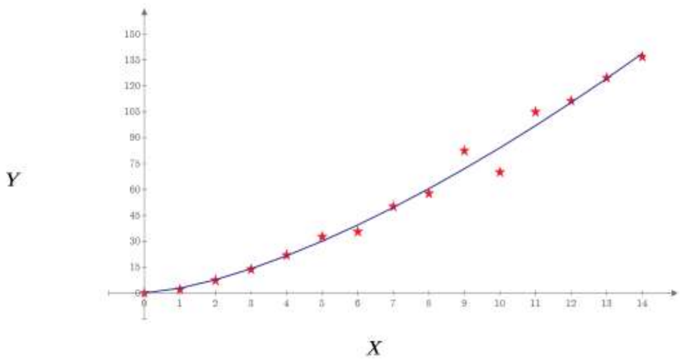

Figure (10): Power Regression Model curve

2.2.8. Exponential Model: [17]

For variables that can develop at steady rates over a particular time period, the phenomena data can occasionally take the shape of an exponential function. The semi-logarithmic equation is used to investigate a situation where the trend increases or decreases by a fixed annual percentage and the trend is exponential if the trend of the data is shown as a straight line. This type of equation is used to measure trends with constant annual rates of change. The general equation for the exponential trend is as follows:

$$y = b_{\mathbf{0}}b^{X} \tag{20}$$

where b has to be positive. We have an exponential growth model for b > 1. We have an exponential decay model when 0 b 1.

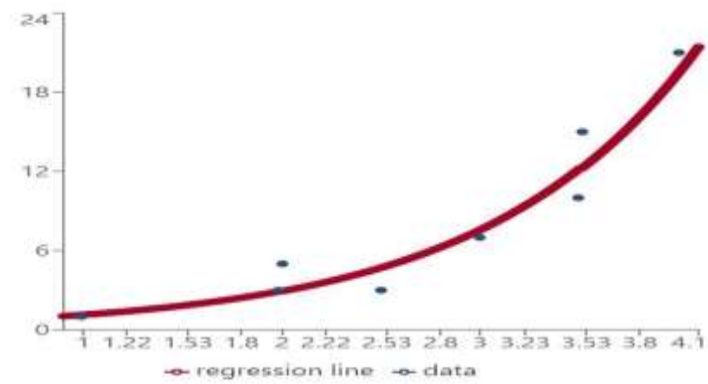

Figure (11): Exponential Regression Model curve

2.2.9. Logistic Model: [3]

When the dependent variable in the regression model is of the descriptive (qualitative) type and accepts two values in numerical form (0,1), such as (success 1, failure 0), (recovered from disease 1, no cure from disease 0), (access 1, no Access 0), etc., it is referred to as a Binary Logistic Regression (BLR) model. In logistic regression, the goal is to determine whether the phenomena under study will occur or not rather than to quantify the connection between the independent factors (measuring the change produced by the independent variables in the dependent variable). The dependent variable must be a binary variable that fits the Bernoulli distribution and take the values (1) with probability P (the likelihood that the response will occur) and the value (0) with two probability q=1-P. This is the basic concept of logistic regression (the probability of the response not occurring). In a linear regression if the independent and dependent variables have continuous values, the model that connects the variables is called:

$$Y = \beta_0 + \beta_1 X_i + e \tag{21}$$

Since Y is a continuous variable represented by a variable, and since the average of Y's observed (actual) values is E(Y/X) and e=Y-Yˆ, Hence, equation (19) may be expressed as follows:

$$E(Y/X) = a_0 + b_1 X_i \tag{22}$$

As is common knowledge, these models' right-hand sides in regression take values between (-) and (+), however when the binary dependent variable takes values between 0 and 1, linear regression is inappropriate because $E(Y/X) = P(Y = 1) = P'$ because the number on the right side can only fall between 0 and 1. The model is therefore irrelevant from the perspective of regression. To solve this problem, the dependent variable's natural logarithm is input, and because 0≤P≤1, the ratio P/(1-P) is positive between (∞,0). The regression model may be written as, 0≤P/(1-P)≤ ∞ in the case of a single independent variable, which is as follows.

$$\ln\left(\frac{P}{1-P}\right) = a_0 + b_1 X_i \tag{23}$$

The model looks like this if we have more than one independent variable:

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \sum_{i=1}^{k} \beta_j X_{ij} \quad ; j = 1,2,\ldots..k \quad , \quad i = 1,2,\ldots..,n \tag{24}$$

The function (24) may be expressed as follows by taking the inverse of the natural logarithm (Exp) of the function.:

$$P = \frac{1}{1+\exp(-\beta_0 + \sum_{i=1}^{k} \beta_j X_{ij})} \tag{25}$$

This model is also known as the logistic regression model or the logit model, and the transformation is denoted by the symbol ln (P/(1-P)). The logit transformation or the logarithm odds ratio are both used in the continuous logistic function, which takes variables. (0,1), where y tends to zero as the right side approaches (-∞), and y tends to one as the right side approaches (+), and the logistic function is the same when the right side is equal to 1. As a result, the logistic regression model is a logarithmic transformation of the linear regression by turning it into a logistic function. As a result, it will follow the traits of the logistic distribution, which confines the possibilities to the range of (1, 0), hence the name logistic regression.

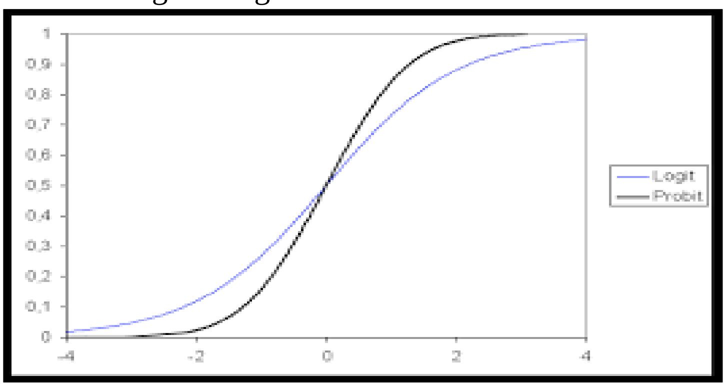

Figure (12) Logistic Regression Model curve

2.2.10. Growth Model: [22]

A statistical technique for examining change over time with longitudinal data is growth curve modeling. To examine trends over time and variations in changes over time among individuals, data obtained from people at various moments in time are analyzed. Growth curve models emphasize both individual differences and individual similarities, as expressed by the covariance structure and the mean structure, respectively. The model can also be expanded to account for differences in change across time in terms of additional factors. This entry expands on the concept of growth curve modeling before talking about unconditional growth curve models and their expansions.

An analysis of a growth curve has two levels. The Level 1 paradigm emphasizes how each person's developmental changes (i.e., the variation within individual over time). The model is presented:

$$Y_{ij}= \beta_{0j}+ \beta_{1j}\,(time)+ \beta_{1j}\,(time)^2+\ldots+ \beta_{ij}\,(time)^n+ r_{ij} \quad (26)$$

Where r_{ij} is the residual in the outcome variable for individual I at Time t, Y_{ij} is the repeatedly measured outcome variable for individual I at Time t, 0 is the initial status of the outcome variable for individual i is the linear rate of change for individual I 2 and 3 are the quadratic and cubic rates of change, respectively.

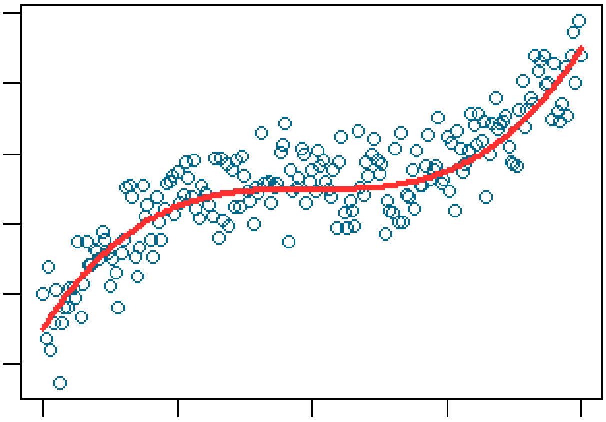

Figure (13) Growth Regression Model curve

3. Fuzzy sets:[15][7]

One of the fundamental tenets of mathematical sciences is the idea of the fuzzy set. The conventional set, or the normal concept of a set, is a set in which the elements either belong to it or do not; there is an absolute distinction between belonging to it and not; and there are very distinct and defined borders for each element that belongs to it. The element cannot be in the group at the same time as another person or object.

Let X a universal set, then the fuzzy subset $\tilde{A}$ consisting of a membership feature $\mu_{\tilde{A}}(x)$ which provide ranges of values $[0, 1]$ for each and every value of x in the fuzzy sample space X:

$$\tilde{A} = \{(x_i, \mu_{\tilde{A}}(x_i)), x \in X, i = 1,2,3, \ldots\ldots n, o < \mu_{\tilde{A}}(x) < 1\} \quad (27)$$

4. Membership function: [1][10]

The membership function concept is given the most weight in the theory of fuzzy sets, which is used to represent many kinds of fuzzy sets. The map that determines the level of health (degree of membership verification) for each element's affiliation with the fuzzy set, or, to put it another way, a function that creates values between 0 and 1 to indicate the degree to which each element in the entire group belongs to the fuzzy set, A range between zero and one is a necessary condition for this function. Its value is non-negative.

5. Fuzzy Numbers:

Fuzzy numbers are used to express uncertainty, which commonly takes the shape of triangles, trapezoids, or other forms. Uncertainty, which frequently takes the form of triangles, trapezoids, or other forms, is described by fuzzy numbers.:

1. Normalized and convex
2. The semi-continuous _a belonging function starts at the top.
3. The level that is particular to each $\alpha \in [0,1]$
4. Based on the true number set R

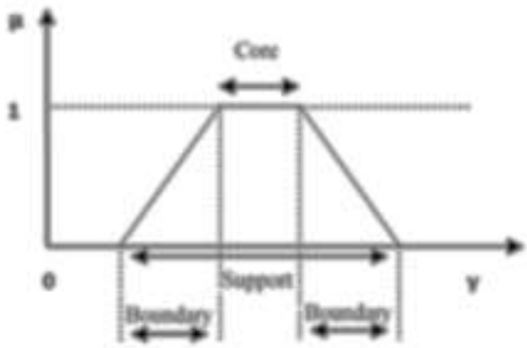

Figure (14): the fuzzy number

5.1. Fuzzy Triangle Number [5][8]

The triangle formed by the interval [a 1,a 3] and its head at x=a2 is known as a1,a2,a3, and may be represented as follows:

$\tilde{N}=(a_1/a_2/a_3)$

Additionally, its membership feature is:

$$\mu_{\tilde{N}}(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & a_1 \le x \le a_2 \\ \frac{a_3-x}{a_3-a_2} & a_2 \le x \le a_3 \\ 0 & o.w \end{cases} \quad (28)$$

5.2. Trapezoidal Fuzzy Number: [5][8]

The triangle with its head at [a2,a 3] and the interval [a 1,a 4] is also known as a 1,a 2,a 3,a 3, and may be represented as $\tilde{M} = (a_1/a_2, a_3/a_4)$

Additionally, its membership feature is:

$$\mu_{\tilde{M}}(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & a_1 \le x \le a_2 \\ 1 & a_2 \le x \le a_3 \\ \frac{a_4-x}{a_4-a_3} & a_3 \le x \le a_4 \\ 0 & o.w \end{cases} \quad (29)$$

6. Fuzzy sample space:[21]

The fuzzy areas $\tilde{x} = (\tilde{x}_1, \dots\dots\dots, \tilde{x}_n)$ from $= (X_1, \dots\dots\dots, X_n)$. ie. Is the set of fuzyy sets for X with membership functions has Borel Measure, and achieved a orthogonally constraint $\sum_{\tilde{x}\in X} \mu_{\tilde{x}}(x) = 1$ and called fuzzy information system (FIS).

7. Applied side

The data representing temperatures per day in the governorates of Iraq, which were taken from the general authority for meteorology and seismic monitoring for the year (2021), we was account the inaccuracy in measurement of data, the principle of fuzziness was used by generating a fuzzy system using Trapezoidal membership function at three α-cut coefficients (0.2 , 0.4, 0.7, 0.9), and using eleventh statistical models namely linear, logarithmic, inverse, quadratic, cube, compound, power, S, Growth, exponential, logistic, The results were as follows:

Table (1): The Models fitting for traditional data

Rank	Model	MSE
1	Power	2.14
2	Compound	2.18
2	Growth	2.18
2	Exponential	2.18
2	Logistic	5.18
3	S	9.22
4	Cubic	13.19
5	Quadratic	21.32
6	Logarithmic	54.17
7	Liner	66.88
8	Inverse	80.79

n=329, Sig< 0.001

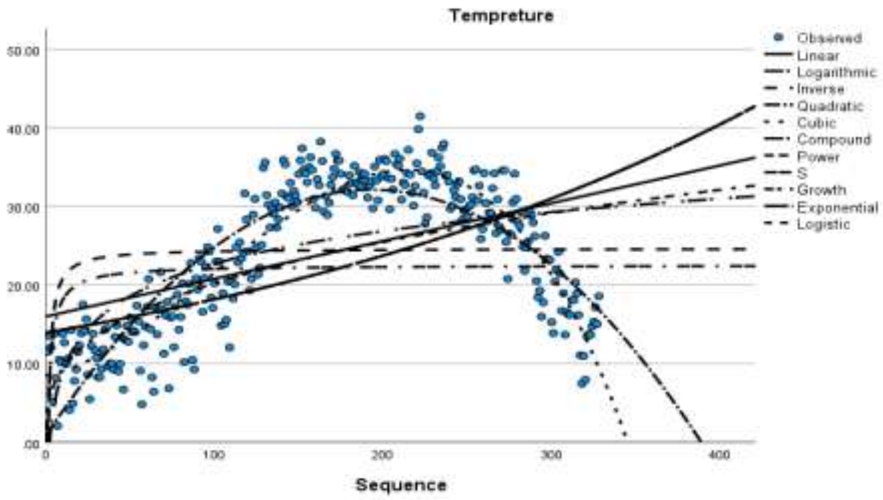

Figure (11): Fitted models curves under traditional data

Table (2) The Models fitting under α=0.2

Rank	Model	MSE
1	Power	1.34
2	Compound	1.21
3	Growth	1.22
4	Exponential	2.38
5	Logarithmic	4.11
6	Logistic	4.78
7	S	8.58
8	Cubic	12.5
9	Quadratic	18.75
10	Liner	56.77
11	Inverse	76.44

n=297, Sig< 0.001

From Table (2) and Figure (11), showed that all models are significant in the fitting of temperatures traditional data in Iraq, and the (Power) model is the most suitable for data with least mean square error (1.34), followed by (Compound) with MSE(1.21) , Growth(1.22) , Exponential(2.38) , Logarithmic (4.11), Logistic (4.78) , S(8.58) , Cubic(12.5) , Quadratic (18.75), Linear (56.77) , and finally Inverse (76.44).

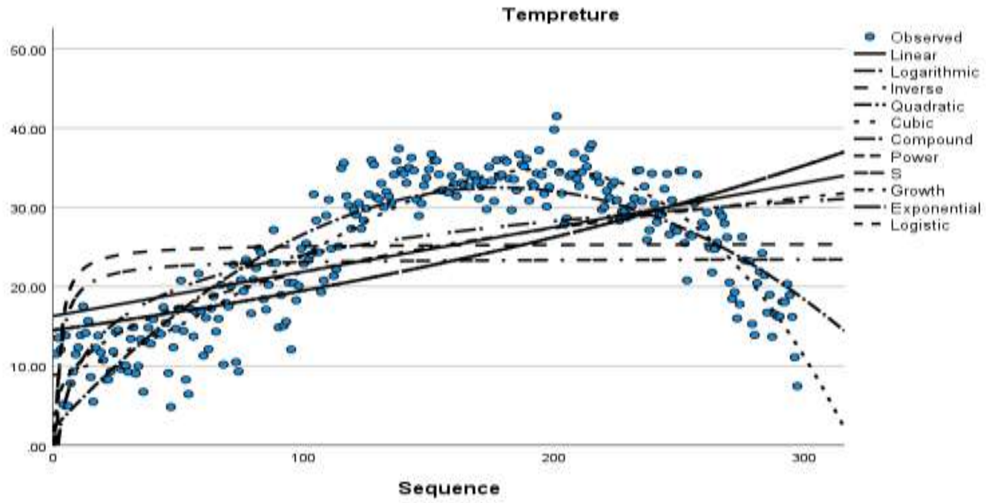

Figure (12): Fitted Models curves under α=0.2

From Table (3) and Figure (12), showed that all models are significant in the fitting of temperatures traditional data in Iraq, and the (Power) model is the most suitable for data with least mean square error (1.34), followed by (Compound) with MSE(1.21) , Growth(1.22) , Exponential(2.38) , Logarithmic (4.11), Logistic (4.78) , S(8.58) , Cubic(12.5) , Quadratic (18.75), Linear (56.77) , and finally Inverse (76.44).

Table (3): The Models fitting under α=0.4

Rank	Model	MSE
1	Power	0.66
2	Compound	0.89
3	Growth	0.98
4	Exponential	2.12
5	Logarithmic	3.78
6	Logistic	3.99
7	S	7.78
8	Cubic	11.23
9	Quadratic	16.44
10	Liner	44.89
11	Inverse	66.28

n=278, Sig< 0.001

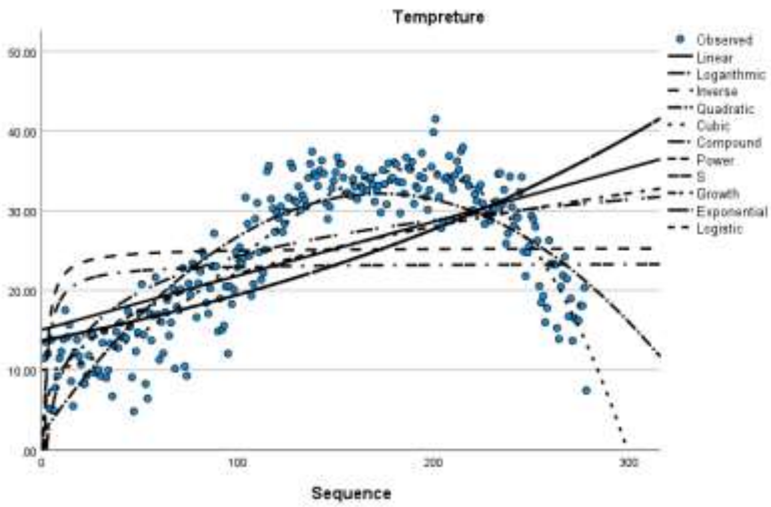

Figure (12): Fitted Models curves under α=0.4

Table (4): The Models fitting under α=0.7

Rank	Model	MSE
1	Power	0.32
2	Compound	0.55
3	Growth	0.67
4	Exponential	1.55
5	Logarithmic	2.21
6	Logistic	2.56
7	S	5.67
8	Cubic	8.52
9	Quadratic	15.78
10	Liner	35.11
11	Inverse	55.09

n=267, Sig< 0.001

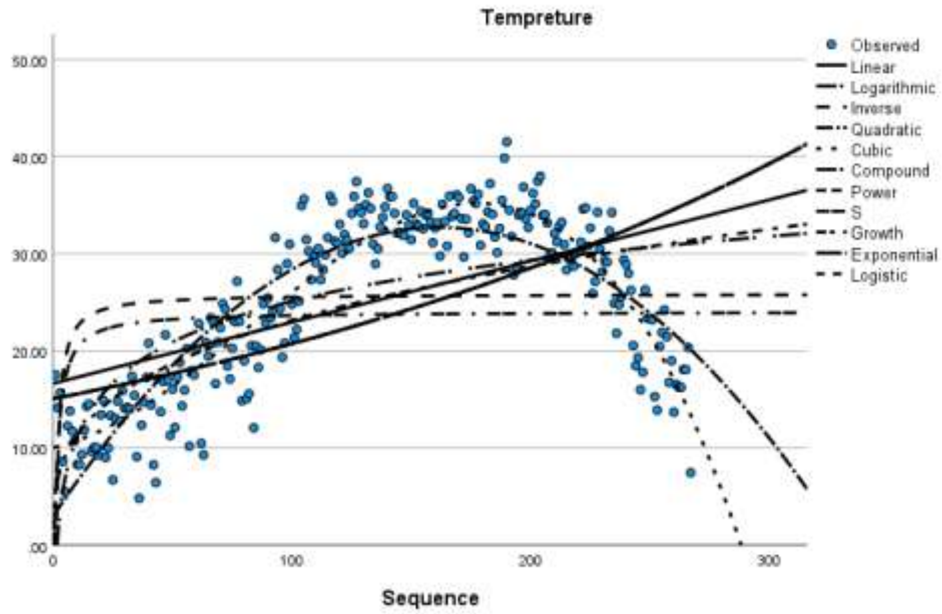

Figure (14): Fitted Models curves under α=0.7

According to these results we will estimate the Power Model with □=0.7 as following:

Table (5): ANOVA table for Power model

	Sum of Squares	df	Mean Square	F	Sig.
Regression	47.137	1	47.137	494.710	0.000
Residual	25.250	265	0.095		
Total	72.386	266			

Table (6): Coefficients table for Power model

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Ln(Case Sequence)	0.437	0.020	0.807	22.242	0.000
(Constant)	2.935	0.271		10.821	0.000

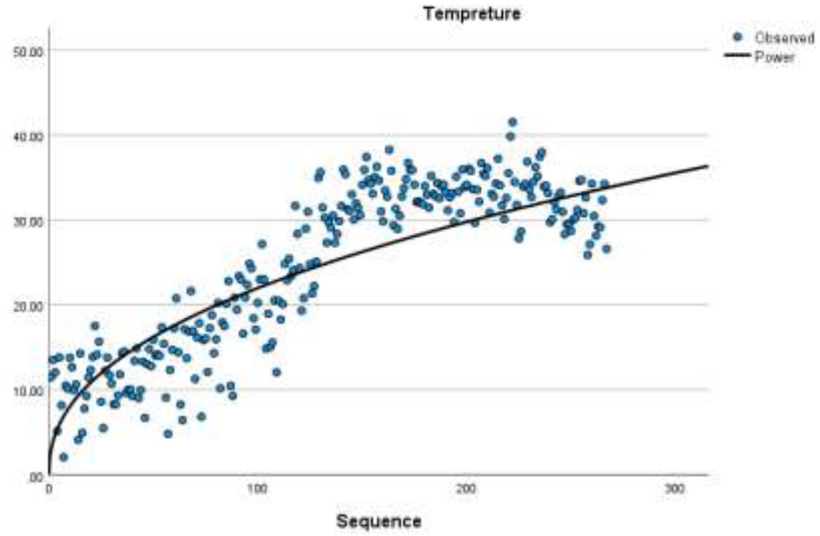

Figure (15): Power Model fitting and observes curves

8. Results & Conclusions

The tables (1), (2), (3) and (4) showed that significance of all models which used to fitting the numbers of data that represent the temperatures per day in the governorates of Iraq, and increase the value of the cut in the fuzzy set led to the fewer observations that have a degree of membership or equal to the □-cut, and the mean of the squares of error (MSE) for all the models that have been fitted is reduced. Under fuzzy sets we also note that the best model in representing the data is the Power model, which recorded the lowest (MSE) among all the models, followed by the Compound, growth ,exponential models, and that the rest models are not suitable for data on the numbers of data that represent the temperatures per day in the governorates of Iraq, and we also note that the best model that achieved a fit for the data was at the □-cut = 0.7 (MSE= 0.32).

References

[1] Abboudi, Emad Hazem, Al – Mashhadani, Athraa Kamel, Salman, Muna Shaker, (2020), "Estimation of the fuzzy reliability function using two-parameter exponential distribution as prior distribution", Periodicals of Engineering and Natural Sciences, Vol. 8, No. 2, June 2020, pp.613-625.

[2] Aberson, C. L. (2019). Applied power analysis for the behavioral sciences, second edition. Routledge.

[3] Allison, P. (2012). "Logistic Regression for Rare Events", website post at http://statisticalhorizons.com/logistic-regression-for-rare-events.

[4] Blake, K. R., & Gangestad, S. (2020). "On attenuated interactions, measurement error, and statistical power: guidelines for social and personality psychologists". Personality and Social Psychology Bulletin, 46(12), 1702-1711.

[5] Buhlmann, Peter ; Van de Geer, Sara,(2011), Statistics for High-Dimensional Data , Methods, Theory and Applications, Springer, Heidelberg Dordrecht London New York.

[6] Chaira , Tamalika, (2019), Fuzzy Set and Its Extension -The Intuitionistic Fuzzy Set, John Wiley & Sons, Inc.

[7] Christian Lyzell , Martin Enqvist, (2012), " Inverse Regression for the Wiener Class of Systems", Proceedings of the 16th IFAC Symposium on System Identification, p. 476-481.

[8] Danyaro, Kamaluddeen Usman, Jaafar, Jafreezal, Liew, Shahir, (2010), "Comparability Between Fuzzy Sets And Crisp Sets: A Semantic Web Approach", International Conference on Education and Educational Technologies - Proceedings

[9] De Barros, Laécio Carvalho, Bassanezi, Rodney Carlos, Lodwick , Weldon Alexander, (2017), "A First Course in Fuzzy Logic, Fuzzy Dynamical Systems, and

Biomathematics-Theory and Applications", Springer-Verlag Berlin Heidelberg, DOI 10.1007/978-3-662-53324-6

[10] Fujii, Toru & Konishi, Sadanori, (2006), "Nonlinear regression modeling via regularized wavelets and smoothing parameter selection," Journal of Multivariate Analysis, Elsevier, Vol. 97(9), pp. 2023-2033, October.

[11] Garg, Harish, Sharma, S.P. & Rani, Monica, (2013), "Weibull fuzzy probability distribution for analyzing the behavior of pulping unit in a paper industry" . Int. J. Industrial and Systems Engineering, Vol. 14, No. 4 , pp 395-413

[12] Hommel ,Gerhard, & Blettner, Maria, (2010), " Linear Regression Analysis ", M E D I C I N E, Deutsches Ärzteblatt International | Dtsch Arztebl Int 2010; 107(44): 776–82.

[13] J. Holton Wilson, (2003), "An Analytical Approach To Detecting Insurance Fraud Using Logistic Regression", Journal of Finance and Accountancy.

[14] Kenneth Benoit, (2011), "Linear Regression Models with Logarithmic Transformations", Methodology Institute London School of Economics, Available at: https://api.semanticscholar.org/CorpusID:15427390

[15] Ling Leng & Wei Zhu (2019), Compound Regression and Constrained Regression: Nonparametric Regression Frameworks for EIV Models, The American Statistician, Vol. 74, No. 3.

[16] Pak, Abbas; Ali, Gholam & Saraj, Mansour, (2013), "Reliability estimation in Rayleigh distribution based on fuzzy lifetime data", Int J Syst Assur Eng Manag. 5, springer , DOI 10.1007/s13198-013-0190-5

[17] Retno Tri Vulandari , Hendro Wijayanto, Afan Lathofy, (2020), "The Application of Non-linear Cubic Regression in Rice Yield Predictions", Desimal Jurnal Matematika 3(3).

[18] Richard A. Armstrong, Anthony C. Hilton, (2014), "Nonlinear Regression: Fitting an Exponential Curve", Statistical Analysis in Microbiology: Wiley.

[19] Sotirios V. Archontoulis and Fernando E. Miguez, (2014), "Nonlinear Regression Models and Applications in Agricultural Research", Agronomy Journal, 105(1).

[20] Tibshirani, Ryan, (2013), Modern regression Methods, 2nd edition, Wiley Interscience.

[21] Torabi, H. & Mirhosseini S. M., (2009), "The Most Powerful Tests for Fuzzy Hypotheses Testing with Vague Data" , Applied Mathematical Sciences, Vol. 3, No. 33, pp. 1619 – 1633.

[22] Cecilia M.S. Ma and Daniel T.L. Shek, (2014), "How to plot growth curves based on SPSS output? Illustrations based on a study on adolescent development", Int J Disabil Hum Dev 2014; 13(2): 183–190.

مجلة كلية الرافدين الجامعة للعلوم (2023)؛ العدد 54؛ 418- 433

AL- Rafidain University College

PISSN: (1681-6870); EISSN: (2790-2293)
مجلة كلية الرافدين الجامعة للعلوم

Available online at: https://www.jrucs.iq

JRUCS
Journal of AL-Rafidain University College for Sciences

النمذجة الإحصائية لدرجات الحرارة في العراق في ظل بيئة ضبابية

بشار خالد علي	مهدي وهاب نصر الله
bashar.alhallaq@mizan.edu.iq	mehdi.wahab@uokerbala.edu.iq
شعبة الإحصاء الصحي والحيوي - دائرة التخطيط والموارد البشرية - مديرية صحة بابل - وزارة الصحة والبيئة، بابل، العراق.	قسم الاحصاء - كلية الادارة والاقتصاد - جامعة كربلاء، كربلاء، العراق.

المستخلص

في هذا البحث تم ملائمة مجموعة بيانات درجات الحرارة باليوم في محافظات العراق والتي تم أخذها من الهيئة العامة للأرصاد الجوية والرصد الزلزالي لعام (2021) في ظل نماذج إحصائية متنوعة وهي: الخطي، اللوغاريتمي، المعكوس، التربيعي، التكعيبي، المركب، S، القوة، النمو، الأسي، والانموذج اللوجستي باستعمال المبدأ التقليدي والمبدأ الضبابي من خلال بناء نظام معلومات ضبابية في ظل قيم مختلفة من معاملات القطع Alfa-cuts لتوليد دوال انتماء لمجموعة درجات الحرارة للحصول على مجموعة تقليدية تأخذ في الاعتبار عدم الدقة في البيانات، تم اختبار معنوية النماذج من خلال القيمة الاحتمالية Sig. للوصول إلى أفضل نموذج يمثل بيانات درجات الحرارة، يتم المقارنة بين النماذج باستعمال معيار متوسط مربعات الخطأ (MSE) وتم التوصل الى أن استخدام مبدأ الضبابية في ناء النماذج أدى إلى زيادة دقة هذه النماذج، وأن متوسط مربعات الخطأ (MSE) لجميع النماذج التي تمت ملائمتها في حالة البيانات الضبابية اقل على ما إذا كانت البيانات تقليدية. ولاحظنا أيضًا أن أفضل انموذج في تمثيل بيانات درجات الحرارة انموذج القوى لامتلاكه أدنى (MSE) بين جميع النماذج، يليه النماذج اللوجيستية والمركبة والأسية بنفس قوة الملاءمة بنفس متوسط مربعات الخطأ (MSE) عند جميع معاملات القطع (0.2 ، 0.4 ، 0.7)، وأن النماذج التكعيبية، التربيعية، الخطية، المعكوسة ليست مناسبة للبيانات الخاصة بعدد درجات الحرارة في العراق، ونلاحظ أيضًا أن أفضل نموذج حقق ملاءمة للبيانات كان عند $\alpha\text{-cut} = 0.7$ $(MSE = 0.32)$.كما نلاحظ زيادة وحدة زمنية واحدة أدت إلى زيادة درجات الحرارة بمقدار (0.437) درجة، وهذا يشير انه بعد سنتين سيكون متوسط درجة الحرارة اليومية في العراق (33.45).

معلومات البحث

تواريخ البحث:
تاريخ تقديم البحث: 2022/12/14
تاريخ قبول البحث: 2023/2/28
تاريخ رفع البحث على الموقع: 2023/12/31

الكلمات المفتاحية:
نظام معلومات ضبابية، تقدير، درجات حرارة، ملائمة بيانات، انموذج خطي، انموذج لوغاريتمي، انموذج معكوس، انموذج تربيعي، انموذج تكعيبي، انموذج مركب، انموذج S، انموذج قوة، انموذج أسي، انموذج لوجيستي.

للمراسلة:
بشار خالد علي
bashar.alhallaq@mizan.edu.iq

https://doi.org/10.55562/jrucs.v54i1.610