Vol. 1 No. 2, December 2021

FUAM Journal of Pure and Applied Science

Available online at
www.fuamjpas.org.ng

An official Publication of
College of Science
Joseph Sarwuan Tarka University,
Makurdi.

FUAM Journal of Pure and Applied Science
© College of Science, Joseph Sarwuan Tarka University, Makurdi
https://fuamjpas.org.ng/
Vol.1. No. 2. December, 2021

Modeling the Chance Dynamics of Evapotranspiration with Some Climatic Variables as Covariates

*I.O. Agada, S.I. Aondoakaa and E.J. Eweh

Department of Physics, Joseph Sarwuan Tarka University, Makurdi. Benue State, Nigeria.
*Correspondence E-mail: gadexx@yahoo.com

Received: 16/08/2021 Accepted: 17/10/2021 Published online: 25/10/2021

Abstract

Over the years, authors have focused on various methods of computing evapotranspiration undermining the stochastic nature of its distribution and the effect of atmospheric variables (covariates). This work employed the binary logistic regression model in modeling and predicting the chance dynamic of evapotranspiration using maximum and minimum relative-humidity, maximum and minimum air temperature, solar irradiance and wind speed as covariates. The result shows that; the model was able to classify correctly 89.4% of high evapotranspiration, 91.5% of low evapotranspiration and an overall 90.4% correct classification. The chance of high evapotranspiration occurring in Kano is higher than low evapotranspiration for a unit rise in any of the covariates except minimum relative humidity. Findings from this study clearly show that logistic regression model can predict evapotranspiration very efficiently.

Keywords: Penman-Monteith's, Evapotranspiration, Wind Speed, Binary Logistic regression.

Introduction

Evapotranspiration is the combination of two processes, namely; evaporation and transpiration. Evaporation and transpiration occur simultaneously, and there is no easy way of distinguishing one process from the other. Evaporation is the loss of water from open bodies of water, such as reservoirs and lakes, bare soil, wetlands and snow cover while transpiration is the loss from living-plant surfaces [1]. Evapotranspiration plays a significant role in maintaining water balance of the terrestrial ecosystem. One of the major challenges that researchers in meteorology and climatology are facing all over the world, is the development of accurate prediction models. Thus, accurate modeling and predicting the chance dynamics of evapotranspiration is essential for well-organized irrigation management, environmental assessment, water resources management, ecosystem modelers, crop production and solar energy system [2]. Many factors such as weather, the crop, the environment and management affect the rate of water loss to the atmosphere by evapotranspiration. The weather factors include climatic variables such as maximum and minimum relative-humidity, maximum and minimum air temperature, solar irradiance and wind speed. Evapotranspiration is usually computed from weather data, as it is difficult or expensive to obtain accurate field measurements. The rate of evapotranspiration is determined using different approaches. These approaches are generally classified into direct and indirect measurement [3].

The direct measurements include the use of lysimeters and atmometers. The indirect measurement involves the use of empirical models such as Temperature-based models, radiation-based models and a combination approach based on Penman model in estimating evapotranspiration using meteorological data [4]. Many researchers have recommended different methods in calculating reference evapotranspiration [5-10] but the Penman-Monteith's equation has been recommend as the best estimator of Reference Evapotranspiration *(ET_o)* if all meteorological data is available [11]. In this work, the binary logistic regression was used in modeling the chance dynamics of evapotranspiration using climatic variables as covariate in Kano. Logistic Regression is a method used when dependent variables are binary, tertiary, ternary and quaternary [12].

The logistic regression is preferred over the simple linear regression because the dependent variable to be predicted is a discrete value, while it is continuous in the linear regression analysis. Also, there is no precondition in the logistic regression regarding the distribution of independent variables. Depending on the kind of a dependent variable, there are three main methods of logistic regression analysis namely: binary logistic regression, ordinal logistic regression and nominal logistic regression [12]. The binary logistic regression model is adopted for this study because it is a regression model that examines the relationship between discrete or continuous (independent) variables and those which have binary result variables (dependent variables). The independent variables are the covariates while evapotranspiration is the dependent variable in this work. In general, the logistic regression aims at estimating parameters according to the logistic model that is formed. Literatures on binary logistic regression with applications to diverse aspect of life are vast [13-17]. This study differs from previous research, because this study did not stop at estimating evapotranspiration in Kano. In furtherance, this study developed a model for predicting the chance dynamic of evapotranspiration using some climatic variables as covariates in Kano.

Kano is located between latitude 11.7574°N and 8.6601°E, and it is the commercial nerve center of northern Nigerian and the third-largest city in Nigeria after Lagos and Ibadan. The principal inhabitants of the city are the Hausa people. Kano is 488 meters above sea level. The city lies to the north of the Jos Plateau, in the Sudanian Savanna region that stretches across the south of the Sahel. The city lies near where the Kano and Challawa rivers flowing from the southwest converge to form the Hadejia River, which eventually flows into Lake Chad to the east. [18]. The temperature of Kano ranges between 15.8°C and 33°C, but during the harmattan, it falls down to as low as 10°C. Kano has two seasonal periods which consist of four to five months of wet season (May - September) and a long dry season lasting from October to April [19].

Penman-Monteith's (FAO-56) Model.
The FAO has adopted the Penman-Monteith's equation (FAO56-PM) as the standard technique in computing reference ET. For a given day, the reference evapotranspiration rate is stated as [11]:

$$ET_O = \frac{0.408\varDelta(R_n - G) + \gamma\frac{900}{T+273}U_2(e_s - e_a)}{\varDelta + \gamma(1+0.34U_2)}\ (mm\ day^{-1}) \qquad (1)$$

where ET_O , T , γ , $\varDelta$, R_n, G , e_s, e_a, and U_2 are reference evapotranspiration rate $(mm\ day^{-1})$, daily mean air temperature, psychometric constant (kPa °C^{-1}), slope of the vapor pressure curve (kPa °C^{-1}), net radiation at the crop surface (MJ m^{-2} d^{-1}), soil heat flux density (MJ m^{-2} d^{-1}), saturation vapor pressure (kPa), actual vapor pressure (kPa) and wind speed at 2m above the ground.
The psychrometric constant can be computed using equation (2):

$$\gamma = 0.000665P \qquad (2)$$

$$P = 101.3\left[\frac{293-0.0065Z}{293}\right]^{5.26} \qquad (3)$$

where P = the atmospheric pressure (kPa) and Z = elevation above sea level (m).

The slope of the vapor pressure curve is given as:

$$\varDelta = \frac{4098\left[0.610exp\left(\frac{17.27*T_{mean}}{T_{mean}+23\ .3}\right)\right]}{(T_{mean}+237.3)^2} \qquad (4)$$

$$T_{mean} = \frac{T_{max.}+T_{min.}}{2} \qquad (5)$$

where T_{mean} = average daily mean air temperature

The net radiation at the crop surface can be computed using equation (6):

$$R_n = R_{ns} - R_T \qquad (6)$$

Given: $R_{ns} = (1-a)R_s$

(7)

where

$$R_T = \sigma[\frac{(T_{max}+273.16)^4 + (T_{min}+273.16)^4}{2}](0.34 - 0.14\sqrt{e_a})[1.35\frac{R_s}{R_{so}} - 0.35]$$

(8)

$$e_a = \frac{e_{(T_{min})}[\frac{RH_{max}}{100}] + e_{(T_{max})}[\frac{RH_{min}}{100}]}{2}.$$

$$e_{(T_{max})} = 0.6108\ \exp(\frac{17.27T_{max}}{T_{max}+237.3})$$

$$e_{(T_{min})} = 0.6108\ \exp(\frac{17.27T_{min}}{T_{min}+237.3}) \qquad (9)$$

where R_{ns} = net solar radiation (MJ m^{-2} day^{-1}), a = albedo (0.3), (R_T) = net terrestrial (long wave) radiation, σ = Stefan-Boltzmann constant [4.903x10^{-9}MJ K^{-4} m^{-2}day^{-1}], R_s = incoming solar radiation (MJm^{-2} day^{-1}), (e_a) = actual vapor pressure, $e_{(T_{min})}$ and $e_{(T_{max})}$ = daily saturation vapour pressure at minimum and maximum temperature, and RH_{max}, RH_{min} are maximum and minimum relative humidity.

The clear-sky radiation R_{so} is given by:

$$R_{so} = (0.75 + 2E10-5Z)R_a$$

(10)

where

$$R_a = \frac{24\times3600}{\pi}G_{sc}\partial r(\cos\varphi\cos\delta\sin\omega_s + \frac{\pi\omega_s}{180}\sin\varphi\sin\delta)$$

(11)

$$\partial r = 1 + 0.033\cos[\frac{2\pi}{365}j] \qquad (12)$$

$$\delta = 23.45 \sin(360 \frac{284 + j}{365}) \quad (13)$$

$$\omega_s = \arccos[-\tan(\varphi)\tan(\delta)] \quad (14)$$

where (R_a) = daily extraterrestrial radiation, G_{sc} = solar constant (1367w/m²) , (ω_s) = sun angle

(δ) = solar declination, φ = latitude of a particular location, (∂r) = inverse relative distance Earth-Sun and j = number of the day in the year between 1 (1 January) and 365 or 366 (31 December)

The wind speed measured at heights other than 2 m can be estimated according to equation (15):

$$U_2 = U_h \frac{4.87}{ln(67.8h - 5.42)} \quad (15)$$

where, U_2 , U_h and h are wind speed 2 m above the ground surface (ms^{-1}), measured wind speed 2 m above the ground surface (ms^{-1}) and height of the measurement above the ground surface (m) respectively.

Logistic Regression

Logistic regression analysis studies the relationship between a categorical dependent variable and a set of independent variables. Generally, the response variable is binary (such as high or low, success or failure, presence or absence, etc) in logistic regression. Binary Logistic regression analysis is a statistical technique that examines the influence of various factors on a dichotomous outcome by estimating the probability of the event's occurrence [20]. Suppose the numerical values of 0 and 1 are assigned to the two outcomes of a binary variable, π is the proportion of observations with an outcome of '1' and '$1-\pi$'is the probability of a outcome of '0'. Mathematically, the logit transformation is written as:

$$logit(y) = ln\left(\frac{\pi}{1-\pi}\right) \quad (16)$$

where the ratio $\frac{\pi}{1-\pi}$ is called the odds and the logit is the logarithm of the *odds*, or just *log odds*.

The odds of an event occurring is defined as the ratio of the probability that the event will occur divided by the probability that the event will not occur. Since logistic regression calculates the odds, the impact of independent variables is usually explained in terms of odds relating 'y' (mean of the response variable) and X (explanatory variable) through the equation $y = \alpha + \beta x$. Mathematically, the natural log odds as a linear function of the explanatory variable are given as:

$$logit(y) = natural \log(odds) = ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta x \quad (17)$$

The parameters of the logistic regression are α and β. Equation (17) is the simple form of the logistic model. Taking the antilog of Equation (17) on both sides, one derives an equation to predict the probability of the occurrence of the outcome of interest as follows:

$$\pi = Probability\ (y = outcome\ of\ interest/X = x, a\ specific\ value\ of\ X)$$

$$\pi = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} = \frac{1}{1+e^{-(\alpha+\beta x)}} \quad (18)$$

Extending the logic of the simple logistic regression of equation 18 to multiple predictors, a complex logistic regression can be constructed as:

$$logit(y) = ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 \quad 19)$$

Hence,

$$\pi = Probability\ (y = outcome\ of\ interest/X = x, a\ specific\ value\ of\ X)$$

$$\pi = \frac{e^{\alpha+\beta_1 X_1+\beta_2 X_2}}{1+e^{\alpha+\beta_1 X_1+\beta_2 X_2}} = \frac{1}{1+e^{-(\alpha+\beta_1 X_1+\beta_2 X_2)}} \quad (20)$$

where π is the probability of the event, α is the Y intercept, βs are regression coefficients, and Xs are a set of predictors. The 'α and βs' are usually estimated using the maximum likelihood method, which is preferred over the weighted least squares approach [17]. The maximum likelihood method usually uses the modified Newton-Raphson method [21]. [22] stated that the probability density function of y_i is given as:

$$P(Y|\pi) = \prod_{i=1}^{n} \pi^{y_i} (1-\pi)^{1-y_i} \quad (21)$$

This gives rise to the log-likelihood:

$$L(y_1, y_2, \ldots y_n, \beta_1, \beta_2, \ldots \beta_n) = L = \prod_{i=1}^{n} P(Y|\pi) = \prod_{i=1}^{n} P(Y|\pi)(1-\pi)^{1-y_i}$$

$$lnL = \sum_{i=1}^{n}[ln\ {}_i^{y_i} + \ln(1-\pi)^{1-y_i}] = \sum_{i=1}^{n}[y_i ln\pi_i + (1-y_i)\ln(1-\pi)]$$

$$= \sum_{i=1}^{n}\left[ln\left(\frac{\pi_i}{1-\pi_i}\right)\right] + \sum_{i=1}^{n}[ln(1-\pi_i)] \quad (22)$$

Since,

$$\pi_i = \frac{e^{x_i\beta_i}}{1+e^{x_i\beta_i}}, 1-\pi_i = \frac{1}{1+e^{x_i\beta_i}} \text{ and } \frac{\pi_i}{1-\pi_i} = e^{x_i\beta_i}$$

Therefore;

$$lnL = \sum_{i=1}^{n}\left[-y_i \ln(1-e^{1-x_i\beta_i} + (1-y_i)\ln\left(1-\frac{1}{1-e^{-x_i\beta_i}}\right)\right] \quad (23)$$

The gradient of the likelihood with respect to estimated coefficient is:

$$\frac{\partial lnL}{\partial \beta} = \sum_{i=1}^{n}(y_i - \pi_i)x_i = 0 \quad (24)$$

Equation (24) is a nonlinear equation, thus requires a unique method of solution in the case of logistic regression. There are two procedures obtainable for testing the significance of one or more independent variables in a logistic regression. They are the likelihood ratio tests and Wald tests. The likelihood-ratio test statistic is given as [23]:

$$-2log\left(\frac{L_0}{L_1}\right) = -2[\log(L_0) - log(L_1)] = -2(L_0 - L_1) \quad (25)$$

where L_0 is likelihood of the fitted model and L_1 is likelihood of saturated model.

The log transformation of the likelihood functions (equation 25) yields a chi-squared statistic. The chi-squared statistic (χ^2) is given as [17]:

$$\chi^2 = \sum_{i=1}^{n} r_i^2 \quad (26)$$

Where $r_i = \frac{y_i - \hat{y}_i}{\sqrt{\hat{y}(1-\hat{y})}}$, $\hat{y}(1-\hat{y})$ is the standard deviation of the residuals and $\hat{y} = \frac{e^{\alpha+\beta_i X_{i1}+\ldots\ldots+\beta_k X_{ik}}}{1+e^{\alpha+\beta_i X_{i1}+\ldots\ldots+\beta_k X_{ik}}}$

This statistic follows a χ^2 distribution with n− (k+1) degrees of freedom, so that p-values can be calculated.

The Wald statistic is the ratio of the square of the regression coefficient to the square of the standard error of the coefficient given as [24]:

$$W_j = \frac{\beta_j^2}{SE_{\beta j}^2} \quad (27)$$

Each Wald statistic is compared with a Chi-square with 1 degree of freedom [24].

Hosmer and Lemeshow proposed grouping the values of the estimated probabilities so as to find the overall goodness of fit. The Hosmer-Lemeshow statistic is given by [25]:

$$C = \sum_{j=1}^{g} \frac{\left(O_j - n_j' \bar{\pi}_j\right)^2}{n_j' \bar{\pi}_j (1-\bar{\pi}_j)} \quad (28)$$

where g denotes the number of groups, $n_j^{'}$ is the number of observation in the j[th] group, O_j is the sum of the Y values for the j[th] group and $\bar{\pi}_j$ is the average probability for the j[th] group. The Hosmer-Lemeshow statistic test is more reliable and robust than the traditional chi square test. Once the model is fitted, checking the validity of deductions drawn from statistical modeling techniques depends on the assumptions of the statistical model being satisfied [20]. The basic assumptions of logistic regression include:

i. That the outcome must be discrete, otherwise explained as, the dependent variable should be dichotomous in nature.
ii. There should be no outliers in the data, which can be assessed by converting the continuous predictors to standardized, or z scores, and remove values below -3.29 or greater than 3.29.
iii. There should be no multicollinearity among the predictors. This can be evaluated by a correlation matrix among the predictors.

Methodology

The daily maximum and minimum relative humidity, maximum and minimum air temperature, solar radiation and wind speed data were obtained from the International Institute of Tropical Agriculture (IITA**)** Ibadan, Nigeria for the period of thirty-four (34) years (1977-2010). The wind speed was estimated for 2 m above the ground surface using equation (15), since the data collected was measured 10 m above ground surface. The daily reference evapotranspiration *(ET₀)* was computed using equation (1) after computing the parameters from equation (2) to (15). The average monthly evapotranspiration *(ET)* was also computed.

Let a random variable E_k describe the daily evapotranspiration *(E)* with realization '0' if *'E'* is below average ($\bar{E}$) and '1' if the daily evapotranspiration (E) is above average ($\bar{E}$). This is termed low and high evapotranspiration respectively. Mathematically we have;

$$E_k = \begin{cases} 0, if\ E_k < \bar{E}(\text{low evapotranspiration}). \\ 1, if\ E_k \geq \bar{E}(\text{high evapotranspiration}). \end{cases} \quad (29)$$

If π is the probability of an outcome of 1 (high evapotranspiration), then $1-\pi$ is the probability of an outcome of 0 (low evapotranspiration). The SPSS package was used in carrying out the binary logistic regression analysis. The independent variables (covariates) are the daily maximum relative humidity, minimum relative humidity, maximum air temperature, minimum air temperature, solar radiation and wind speed, while the dependent variable (outcome) is either the high and low evapotranspiration.

Results and Discussion

Evapotranspiration as earlier explained, is the term used in describing the loss of water to the atmosphere through evaporation and transpiration. The maximum average monthly evapotranspiration occurs in the month of March and minimum in August, as observed in Figure 1. This implies that, more water is lost to the atmosphere in the month of March when compared to the month of August. This could also be as a result of high amount of solar radiation experience in the month of March than any other month of the year.

According to [26], the loss of 1 mm of water per day is the loss of 10 m^3 of water per hectare (10,000 litres per hectare). That is, over 10 m^3 of water per hectare is lost daily in kano yearly as observed in Figure 1. Since more water is lost from the month of January to May, farmers can design irrigation systems in order to meet peak water requirements for their crops.

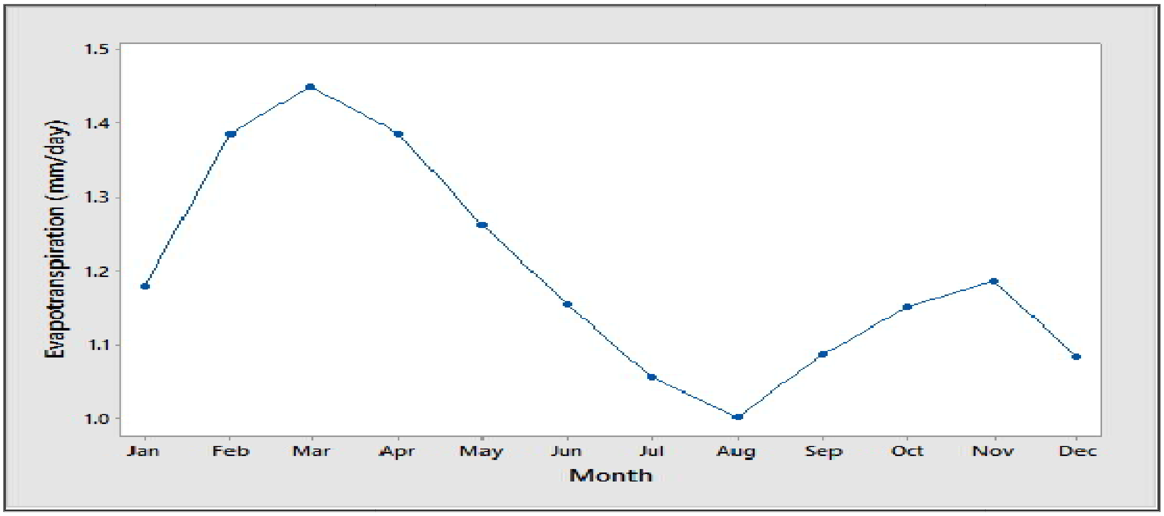

Figure 1: Average monthly Evapotranspiration for Kano State.

Logistic regression is used to predict a categorical (usually dichotomous) variable from a set of predictor variables. In this study, the high and low evapotranspiration is our categorical variable while solar radiation, maximum and minimum air temperature, maximum and minimum relative humidity and wind speed are the covariates or predictor variables. Modeling and predicting the chance dynamics (high and low) of evapotranspiration in Kano is of great importance to crops and farmers. The Binary Logistic Regression procedure reports the Hosmer-Lemeshow goodness-of-fit statistic and uses it to determine whether the model adequately describes the data. The Hosmer-Lemeshow tests the null hypothesis of the predictions made by the model fit with the observed group memberships, as shown in Table 1. [27] stated that the Hosmer-Lemeshow tests procedure suffers from several problems, and that even Hosmer and Lemeshow have acknowledged the problems with this test. He went further to state that with large sample sizes, the test may be significant, even when the fit is good but with small sample sizes, it may not be significant, even with poor fit.

Table 1: Model summary and Hosmer-Lemeshow tests

	Model Summary			Hosmer - Lemeshow Test		
Step	-2 Log likelihood	Cox and Snell R Square	Nagelkerke R Square	Chi-square	df	Sig.
1	12760.612	0.287	0.384	137.781	8	0.00
2	9019.167	0.475	0.634	233.700	8	0.00
3	7954.910	0.519	0.693	380.644	8	0.00
4	7004.578	0.555	0.741	1508.051	8	0.00
5	6842.009	0.561	0.749	1882.979	8	0.00
6	6748.868	0.564	0.753	1974.220	8	0.00

df = degree of freedom

As the number of steps is increasing, the -2 log-likelihood is decreasing, while the R^2 (Cox and Snell and Nagelkerke) is increasing as shown in Table 1. In the linear regression model, the coefficient of determination 'R^2' summarizes the proportion of variance in the dependent variable associated with the independent variables. A larger value of R^2 indicates that more of the variation is explained by the model, to a maximum of 1. For binary logistic regression models with a categorical dependent variable, it is not possible to compute a single R^2 statistic that has all the characteristics of R^2 in the linear regression model, so these estimates are computed instead. The Cox and Snell and Nagelkerke methods are generally used in estimating the coefficient of determination in binary logistic regression. The Cox and Snell and Nagelkerke under the model summary have R^2 equal to 0.56 (56%) and 0.75 (75%) respectively after step 6 as shown in Table 1. Each step explains the addition of one climatic variable to the model. R square defines the model applicability. Thus, higher values are indication of adequate data fitness using the model.

The classification table shows the practical results of using the logistic regression model. Cells on the diagonal are correct predictions, as presented in Table 2. Of the cases used to create the model, 5965 of the 6517 data on low evapotransipiration are classified correctly, while 5107 of the 5714 data on high evapotransipiration data are classified correctly. This model was able to classify correctly 89.4% of high evapotranspiration and 91.5% of low evapotranspiration. From step to step, the improvement in classification signifies how well the model performs, as shown in the percentage correct in Table 2. Overall, 90.5% of the cases are classified correctly, as shown in Table 2.

By default, SPSS sets its threshold to 0.5 in order to make decision after classifying the subjects in the analysis.

Table 2: Classification table of the observed and predicted high and low Evapotranspiration

STEP	OBSERVED		PREDICTED EVAPOTRANSIPIRATION LOW	HIGH	Percentage Correct
1	EVAPOTRANSIPIRATION	LOW	4985	1532	76.5
		HIGH	1897	3817	66.8
	Overall Percentage				72.0
2	EVAPOTRANSIPIRATION	LOW	5604	913	86.0
		HIGH	999	4715	82.5
	Overall Percentage				84.4
3	EVAPOTRANSIPIRATION	LOW	5761	756	88.4
		HIGH	820	4894	85.6
	Overall Percentage				87.1
4	EVAPOTRANSIPIRATION	LOW	5918	599	90.8
		HIGH	653	5061	88.6
	Overall Percentage				89.8
5	EVAPOTRANSIPIRATION	LOW	5942	575	91.2
		HIGH	634	5080	88.9
	Overall Percentage				90.1
6	EVAPOTRANSIPIRATION	LOW	5965	552	91.5
		HIGH	607	5107	89.4
	Overall Percentage				90.5

If the probability of the event is greater than or equal to some threshold, then the chance of the event occurring is high. The probability of high evapotranspiration occurring is greater than low evapotranspiration since the threshold is greater than 0.5 as presented in Figure 2

.

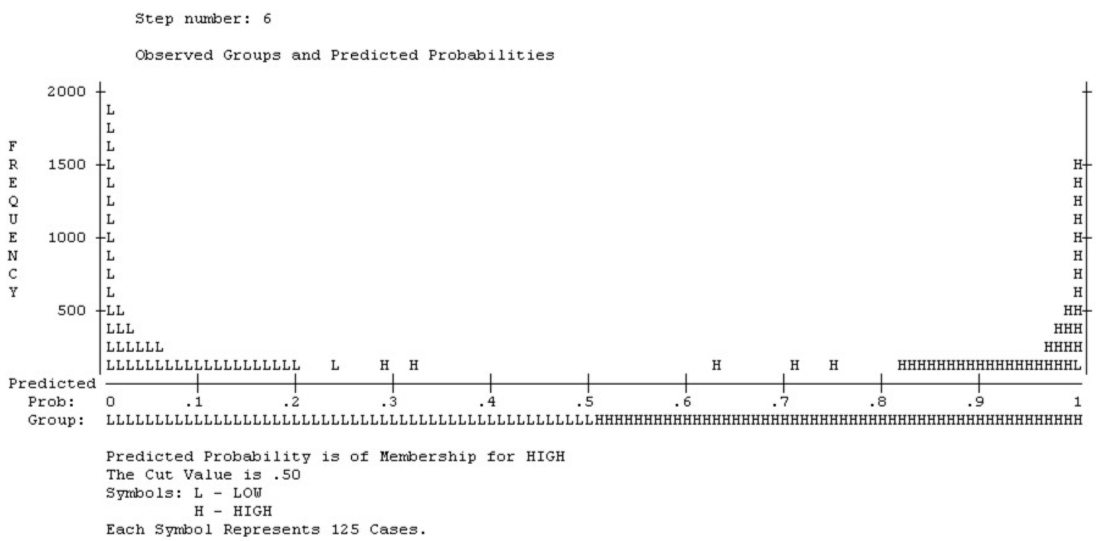

Figure 2: Frequency of observed groups and predicted probabilities of high and low evapotranspiration.

The ratio of the coefficient to its standard error, all squared, equals the Wald statistic. If the significance level of the Wald statistic is small ($p < 0.05$) then the parameter is useful to the model. From Table 3, it is observed that the significance level of the Wald statistic is less than 0.05 ($p < 0.05$), therefore the predictors and coefficient values shown in the last step can be used to make predictions. The interpretation of a logistic regression coefficient is not as straight forward as that of a linear regression coefficient. The coefficient (β) is suitable for testing the usefulness of predictors, while *Exp* (β) is easier to interpret. *Exp* (β) represents the ratio-change in the odds of the event of interest for a one-unit change in the predictor. The variables in the equation output in Table 3 shows that the regression equation in the various steps are:

1: $\ln(odds) = -5.149 + 0.276\, Solar\ \boldsymbol{radiation}.$ (30a)

2: $\ln(odds) = -19.103 + 0.339\, Max.Air\ Temp. + 0.378\, \boldsymbol{Solar}\ radiation.$ (30b)

3: $\ln(odds) - 29.071 + 0.386\, Max.Air\ Temp. + 0.297\, Min.Air\ Temp. + 0.446\, Solar\ radiation.$ (30c)

4: $\ln(odds) = -34.812 + 0.525\, Wind\ speed + 0.449\, Max.\boldsymbol{Air}\ Temp. + 0.314\, Min.Air\ Temp. + 0.495\, Solar\ \boldsymbol{radiation}.$ (30d)
(30d)

5: $\ln(odds) = -33.581 + 0.544\, Wind\ speed - 0.027\boldsymbol{Min}.rel.hum. + 0.444\, Max.Air\ Temp. + 0.319\, Min.Air\ Temp. + 0.501\, Solar\ \boldsymbol{radiation}.$ (30e)

6: $\ln(odds) = -93.916 + 0.564 Wind\ speed + 0.6\, Max.rel.hum. - 0.027Min.rel.\boldsymbol{hum}. + 0.446\, Max.Air\ Temp. + 0.326\, Min.Air\ Temp. + 0.509\, Solar\ \boldsymbol{radiation}.$ (30f)

The model in step 6 can be used to predict the odds that evapotranspiration would be high or low in Kano, because the model classify correctly 89.4% of high evapotranspiration and 91.5% of low evapotranspiration as shown in Table 2. From Table 4, the odds ratio *Exp*(β) predicts that the odds of having high evapotranspiration as a result of one unit rise in solar radiation is 1.66 times higher than the odds of having low evapotranspiration if all other variables are kept constant. The probability of high evapotranspiration occurring in Kano for one unit rise in solar radiation is 0.62(62%), while 0.38 (38%) chance for low evapotranspiration occurring if all other variables are kept constant as shown in Table 4. The chance of high evapotranspiration occurring in Kano is higher than low evapotranspiration for a unit rise in solar radiation, minimum and maximum air temperature, maximum relative humidity and wind speed expect minimum relative humidity as seen in Table 4.

Table 3: Logistic Regression Coefficients of the covariates.

Parameters in the Equation							
Step	Climatic Variables	β	S.E.	Wald	df	Sig.	Exp(β)
1	Solar radiation	0.276	0.005	2.54E+03	1.000	0.000	1.317
	Constant	-5.149	0.102	2.57E+03	1.000	0.000	0.006
2	Solar radiation	0.378	0.008	2.48E+03	1.000	0.000	1.459
	Max. Air Temp.	0.339	0.007	2.36E+03	1.000	0.000	1.404
	Constant	-19.103	0.347	3.03E+03	1.000	0.000	0.000
3	Solar radiation	0.446	0.009	2.48E+03	1.000	0.000	1.562
	Min. Air Temp.	0.297	0.010	873.54	1.000	0.000	1.346
	Max. Air Temp.	0.386	0.008	2.29E+03	1.000	0.000	1.471
	Constant	-29.071	0.558	2.71E+03	1.000	0.000	0.000
4	Solar radiation	0.495	0.010	2.38E+03	1.000	0.000	1.640
	Min. Air Temp.	0.314	0.011	830.153	1.000	0.000	1.369
	Max. Air Temp.	0.449	0.009	2.25E+03	1.000	0.000	1.566
	Wind speed	0.525	0.019	782.806	1.000	0.000	1.690
	Constant	-34.812	0.679	2.63E+03	1.000	0.000	0.000
5	Solar radiation	0.501	0.010	2.32E+03	1.000	0.000	1.650
	Min. Air Temp.	0.319	0.011	838.311	1.000	0.000	1.376
	Max. Air Temp.	0.444	0.010	2.14E+03	1.000	0.000	1.558
	Min. Rel. Hum.	-0.027	0.002	156.563	1.000	0.000	0.974
	Wind speed	0.544	0.019	809.439	1.000	0.000	1.723
	Constant	-33.581	0.687	2.39E+03	1.000	0.000	0.000
6	Solar radiation	0.509	0.011	2.30E+03	1.000	0.000	1.663
	Min. Air Temp.	0.326	0.011	852.801	1.000	0.000	1.385
	Max. Air Temp.	0.446	0.010	2.12E+03	1.000	0.000	1.561
	Min. Rel. Hum.	-0.027	0.002	160.845	1.000	0.000	0.973
	Max. Rel. Hum.	0.600	0.068	77.332	1.000	0.000	1.821
	Wind speed	0.564	0.019	840.032	1.000	0.000	1.757
	Constant	-93.916	6.932	183.533	1.000	0.000	0.000

Variable entered in step 1: Solar Radiation
Variable entered in step 2: Maximum Air Temperature.
Variable entered in step 3: Minimum Air Temperature.
Variable entered in step 4: Wind speed.
Variable entered in step 5: Minimum Relative Hum.
Variable entered in step 6: Maximum Relative Hum.

Table 4: Chance distribution of high and low evapotranspiration for each covariates

Covariates	π	1 − π	Exp(β)
Solar radiation	0.624	0.376	1.663
Min. Air Temp.	0.581	0.419	1.385
Max. Air Temp.	0.610	0.390	1.561
Min. Rel. Hum.	0.493	0.507	0.973
Max. Rel. Hum.	0.646	0.354	1.821
Wind speed	0.637	0.363	1.757

π = chance of high evapotranspiration for a unit rise in the associated covariate
$1-\pi$ = chance of low evapotranspiration for a unit rise in the associated covariate
Exp(β) = odds of high evapotranspiration keeping other covariates constant.

The increase in solar radiation, minimum and maximum air temperature, maximum relative humidity and wind speed tends to increase the rate of water lost to the atmosphere through evapotranspiration except minimum relative humidity. This information can be essential for well-organized irrigation management, environmental assessment, water resources management, ecosystem modelers, and crop production in Kano. This could possibly explain why [28], reported that the increasing temperature in the semi-arid region of Sokoto, Katsina, Kano, Nguru and Maiduguri may be attributed to increasing evapotranspiration, drought and desertification.

Conclusion

Our main objective is to model the chance dynamics of evapotranspiration using some climatic variables in Kano using binary logistic regression model. This model was able to classify correctly 89.4% of high evapotranspiration, 91.5% of low evapotranspiration and 90.4% overall. The chance of high evapotranspiration occurring in Kano is higher than low evapotranspiration for a unit rise in any of the covariates except minimum relative humidity. Hence, irrigated agriculture in Kano is essential to boost food production to feed the increasing population. This study provides a new insight in modeling the chance dynamics of evapotranspiration with the associated climatic variables as covariates using the binary logistic regression model.

Declaration of conflicting interests

The authors declared no potential conflicts of interest

References

[1] Chineke, T. C., Idinoba, M. E. and Ajayi, O. C. 2011.Seasonal evapotranspiration signatures under a changing landscape and ecosystem management in Nigeria: Implications for agriculture and food security. American Journal of Scientific and Industrial Research. 191:204.

[2] Lang D., Jiangkun Z., Jiaqi S., Feng L., Xing M., Wenwu W., Xuli C. and Mingfang Z. 2017. A Comparative Study of Potential Evapotranspiration Estimation by Eight Methods with FAO Penman–Monteith Method in Southwestern China Water. 9(734):1-18.

[3] Ilesanmi O.A., Oguntade, P.G. and Olufayo, A.A. 2012. Re-examination of the BMN model for estimating evapotranspiration. International Journal of Agriculture.2(6):268-272.

[4] Echiegu, E.A., Ede, N.C. and Ezenne, G.I. 2016. Optimization of Blaney-Morin-Nigeria (BMN) model for estimating evapotranspiration in Enugu, Nigeria. African Journal of Agricultural Research. 11(20): 1842-1848.

[5] Thornthwaite C. W., 1948. An approach toward a rational classification of climate. Geographical Review. 38(1): 55–94.

[6] Makkink, G. F. 1957. Testing the Penman formule by means of lysimeters. Jour. Inst. Water Eng. London. 11:277-288.

[7] Turc, L. 1961. Evaluation des besoins en eau d'irrigation, evapotranspiration potentielle, formula climatique simpiifiee at mice a jour (In French) (English title: Estimation of irrigationwater requirements, potential evapon.anspiration: a simple climatic formula evolved up to date.) Ann. Agron. 12:13-49.

[8] Jensen, M. E., and liaise, H. R. 1963. Estimating evapotranspiration from solar radiation. J. Irrig. And Drain. Div., Amer. Soc. Civil Eng. Proc. 89:15-41.

[9] Jacobs, J.M.and Satti, S.R. 2001. Evaluation of Reference Evapotranspiration Methodologies and Afsirs Crop Water Use Simulation Model: Final Report; University of Florida: Gainesville, FL, USA,

[10] Alexandris S, Stricevic R, and Petkovic S. 2008. Comparative analysis of reference evapotranspiration From the surface of rainfed grass in central Serbia, calculated by six empirical methods against the Penman-Monteith formula. Eur. Water 21(22):17-28.

[11] Allen R G, Pereira L S, and Raes D. 1998. Crop evapotranspiration, Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56.

[12] Korkmaz, M., Güney, S. and Yiğîter, S.Y. 2012.The Importance of Logistic Regression Implementations in the Turkish Livestock Sector and Logistic Regression Implementations/Fields. J.Agric. Fac. HR.U.,16(2): 25-36.

[13] Bagley, S. C., White, H., and Golomb, B. A. 2001. Logistic regression in the medical literature: Standards for use and reporting, with particular attention to one medical domain. Journal of Clinical Epidemiology, 54(10): 979-985.

[14] Giancristofaro, R. A., and Salmaso, L. 2003. Model performance analysis and model validation in logistic regression. Statistica, 63(2): 375-396.

[15] Hsieh, F. Y., Bloch, D. A., and Larsen, M. D. 1998. A simple method of sample size calculation for linear and logistic regression. Statistics in Medicine,17(14): 1623-1634.

[16] Peng, C. J., Lee, K. L., and Ingersoll, G. M. 2002. An introduction to logistic regression analysis and reporting. The Journal of Educational Research,96(1): 3-14.

[17] Peng, C. J., and So, T. H. 2002. Logistic regression analysis and reporting: A primer. Understanding Statistics, 1(1): 31-70.

[18]Nwagbara, M.O. 2015. Case study: Emerging Advantages of Climate Change for Agriculture in Kano state, north western Nigeria. *American Journal of Climate Change*, 4:263-268.

[19] Ibrahim A. 2006. Efforts of the Shekarau admnistration in harnessing resources for social and economic development of Kano. A presentation to Course 28 of Command and Staff College Jaji, Research and Documentation Directorate 2.

[20] Reddy O.C. S., Habte, T. L.and Lamessa A. 2015. Binary Logistic Regression Analysis in Assessing and Identifying Factors that Influence the Use of Family Planning: The Case of Ambo Town, Ethiopia International Journal of Modern Chemistry and Applied Science, 2(2): 108-120.

[21] Hilbe J. M. 2009. Logistic Regression models. USA: Chapman and Hall Book.

[22] Ward, M. D. 2007. Maximum likelihood for social sciences strategies for analysis. Course on maximum likelihood methods for the social sciences. Cambridge University press.

[23] Hosmer D. W. and Lemeshow S. 2000. Applied Logistic Regression, Second Edition. John Wiley and Sons, Inc.

[24] Hyeoun-Ae, P. 2013. An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain . J Korean Acad Nurs. 43(2).

[25] RahmatullahImon, A.H.M. Roy,M.C. and Bhattacharjee, S.K. 2012.Prediction of Rainfall Using Logistic Regression . Pak.j.stat.oper.res. 8(3): 655-667.

[26] Harris, G. A. 2002, Irrigation: water balance scheduling, DPI Note FS0546, QDPI, Brisbane.

[27] Karl L. W. 2020, Binary Logistic Regression with SPSS. http://core.ecu.edu/psyc/ wuenschk/SPSS/SPSS-MV.htm

[28] Adefolalu, D. O. 2007. "Climate change and economic sustainability in Nigeria", Paper presented at the International Conference on Climate Change and Economic Sustainability held at Nnamdi Azikiwe University, Enugu, Nigeria.1-12.

Cite this article

Agada I.O., Aondoakaa S.I. and Eweh E.J. (2021). Modeling the Chance Dynamics of Evapotranspiration with Some Climatic Variables as Covariates. FUAM Journal of Pure and Applied Science,**1**(**2**):58-67