FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 8 No. 3, June, (Special Issue) 2024, pp 67 - 81
DOI: https://doi.org/10.33003/fjs-2024-0803-2468

STATISTICAL ANALYSIS OF ET EXPLANATORY VARIABLES IN CONFLUENCE STATE, LOKOJA, SAVANNAH REGION OF NIGERIA

***[1]Yahaya, O. M., [2]Ogolo, E. O., [3]Agele, S. O. and [4]Ajayi, V. O.**

[1]Department of Physics, Federal University, Lokoja
[2]Department of Physics, Federal University of Technology, Akure
[3]Department of Crop Science, Federal University of Technology Akure
[4]Department of Meteorology and Climate Science, Federal University of Technology Akure

*Corresponding authors' email: otori.yahaya@fulokoja.edu.ng

ABSTRACT

This observational study investigated trends of Lokoja climate variables relating to evapotranspiration from 1989 to 2019. Studying evapotranspiration in the savannah region of Nigeria is critical for understanding the local climate, water availability, and ecosystem dynamics, and can inform sustainable development and adaptation strategies in the face of climate change. The FAO-56 Penman-Monteith model (P-M) was used to estimate the ET from the ERA5 reanalysis monthly mean temperature, precipitation, wind speed and atmospheric pressure data of Lokoja. Statistically significant explanatory variables were determined using multiple regression analysis, and multicollinearity and heteroscedasticity tests were conducted on the results. Clustered column charts used to visualize the performance of the model revealed that increase in precipitation did not automatically translate to increase in ET. Linear regression of ET against temperature revealed that temperature explains approximately 29% of the variability in ET. At 95% confidence level and 251 degree of freedom, the R^2 (0.98) with standard error 0.11 indicate that the statistical analysis of the ET explanatory variables is robust and reliable, and the model is able to accurately predict the values of the ET. The multiple regression analysis result revealed that a mean daily increase of ET, starting from 0.54mm per day, is affected positively by average wind speed (u) of about 0.31mm/day for every unit increase in wind speed (u) at two meters height, 0.27mm/day for a unit increase in net radiation (Rn), 0.98mm/day for a unit increase in vapour pressure deficit (VPD), but a decrease of 0.01mm/day for every unit increase in relative humidity (RH). If water is never a limiting factor, the statistically significant explanatory variables for potential evapotranspiration are vapour pressure deficit, wind speed, net radiation and relative humidity, temperature being the main driving factor for all. The developed equation would help to improve ET prediction, inform water management policies, and enhance agricultural practices in the Savannah region.

Keywords: Climate, Evapotranspiration, Models, Multicollinearity, Heteroscedasticity

INTRODUCTION

Evapotranspiration, the process by which water is transferred from the Earth's surface to the atmosphere through evaporation and plant transpiration, is a key component of the hydrological cycle and plays a crucial role in the water balance and agricultural productivity of a region. It is influenced by a variety of factors, including climatic, topographic, and land cover characteristics. Understanding the relationships between these factors and evapotranspiration can help improve water resource management and forecasting in a specific region like Lokoja. The amount of water vapour in the atmosphere plays an important role in weather forecasting (Mundo-Molina, 2015). It is an established fact that climate change impacts on soil water balance lead to changes in soil evaporation and plant transpiration. The Intergovernmental Panel on Climate Change (IPCC) final report shows climate impacts are already more widespread and severe than expected (IPCC, 2023). The IPCC report warned that between 32-132million people might be driven into extreme poverty in the next decade by climate change. Globally, 2023 was the hottest year on record while some scientists are already warning that there is a strong chance 2024 could beat that record.

The amount of evapotranspiration (ET) per day is controlled by a number of weather parameters like atmospheric pressure, temperature, wind speed, humidity, solar radiation, and so on (Banik et al., 2012). With the temperature increase and precipitation fluctuations, the water availability for crop production and rearing of animals are likely to be affected. Increase in temperature (global warming) results in a corresponding increase in evapotranspiration, consequently increasing the crop water requirement. The importance of evapotranspiration in water management, agriculture and weather forecast has been recently emphasized by various researchers. It drives the Earth climate system at various scales as an important component of the water cycle and energy balance (Ashaolu et al., 2018).

Methods provided for estimating ET are based on one or more measured climate variables (Yates & Strzepek, 1994). The reliability of these methods varies from one climatic system to another. The methods are classified into four categories, which are temperature-based, radiation-based, combination-type equations and pan evaporation-based. Overall, the combination-type Penman-Monteith equation is recommended by Food and Agricultural Organization (FAO) for most climatic conditions as it is considered to be most physical and reliable method and the sole standard to verify other empirical methods (Satish, 2018). It has strong fundamental physical principles, including energy balances to precisely calculate the ET. P-M method, however, is a complex equation which requires rigorous reasoning process and detailed data for four meteorological parameters; air temperature, relative humidity, wind speed, and net radiation. Such input data may not be available at many places especially in developing countries like ours though they can be estimated from available equations involving more stressful conversion processes.

Lokoja, a city located in central Nigeria, experiences a tropical savanna climate with distinct wet and dry seasons. The city lies at the confluence of the Niger and Benue Rivers, which also influences its climate and geography (Animashaun et al., 2020). The high humidity levels in Lokoja, especially during the wet season, promote evapotranspiration. The abundance of water bodies in and around Lokoja, such as the rivers and nearby reservoirs, also contribute to increased evapotranspiration. The topography of Lokoja is mostly flat, with the city located in a river valley. This flat terrain allows for efficient drainage of excess water, reducing the risk of waterlogging and promoting evapotranspiration. The vegetation cover in Lokoja, which includes forests, grasslands, and agricultural crops, also plays a role in evapotranspiration. Plants release water vapor through transpiration, which, combined with evaporation from the soil and surface water bodies, contributes to overall evapotranspiration rates in the region. The unique climatic and geographical characteristics of Lokoja, such as its high humidity levels, abundant water bodies, flat terrain, and diverse vegetation cover, all work together to create favorable conditions for evapotranspiration in the city and its surrounding areas.

There is a lack of adequate studies that have investigated the influence of climate change on evapotranspiration in Lokoja region. With climate change expected to impact precipitation patterns and temperatures in the region, it is important to understand how these changes will affect evapotranspiration rates. Localized evapotranspiration models are important in Lokoja, Nigeria due to the specific climate and environmental conditions in the region. The region experiences a tropical savanna climate with distinct wet and dry seasons, and accurate evapotranspiration models are essential for water resource management, agricultural planning, and climate change studies in the area. Localized evapotranspiration models can provide valuable information on the water balance in Lokoja, helping to determine water availability for irrigation, urban development, and other purposes. Additionally, these models can be used to assess the impact of climate change on evapotranspiration rates in the region, which can inform adaptation strategies and mitigation efforts. In this study, we aim to conduct a statistical analysis of evapotranspiration explanatory variables in the confluence region of Lokoja. By analyzing data on meteorological variables such as temperature, humidity, wind speed, and solar radiation, we hope to identify the key weather factors driving evapotranspiration in this region. Our analysis involves a combination of statistical techniques, including regression and spatial analysis, to determine the relationships between evapotranspiration and its explanatory variables. By gaining a better understanding of these relationships, we can develop more accurate models for predicting evapotranspiration rates in the confluence region of Lokoja, ultimately improving water resource management and agricultural planning in the area. Regression analysis is a statistical method used to examine the relationship between a dependent variable and one or more independent variables (Montgomery et al., 2012). It helps to identify and quantify the strength of the relationship between variables by calculating the correlation coefficient and determining the regression equation. Spatial analysis, on the other hand, is a technique that examines the spatial relationships and patterns of data. This includes analyzing how data is distributed across space and identifying spatial trends and patterns. In the context of understanding the relationship between evapotranspiration (ET) and meteorological variables, regression analysis can be used to determine the impact of variables such as temperature, humidity, wind speed, and solar radiation on ET rates. By conducting regression analysis, we can quantify the influence of these meteorological variables on ET and understand how changes in these variables impact ET rates. Spatial analysis can complement regression analysis by examining how the distribution of meteorological variables across space influences ET rates. By mapping the spatial distribution of meteorological variables and ET rates, we can identify spatial patterns and trends that may impact ET. This allows for a more comprehensive understanding of the relationship between meteorological variables and ET rates, ultimately providing valuable insights for water resource management and agricultural practices in the region.

MATERIALS AND METHODS

Weather Characteristics of the Area

Lokoja is the capital of Kogi State, situated in the tropical wet and dry savanna climate zone in the North-central Nigeria. It is a confluence town (where Rivers Niger and Benue meet) and lies on latitude 7°49'N of the equator, Longitude 6°44'E of the Meridian, elevation 53m above sea level and experiences average atmospheric pressure of about 100kPa. It has a mean annual rainfall of about 1150mm and mean monthly temperature close to 30°C, hot all year-round, accessed from *https://www.fulokoja.edu.ng>aboutus*. The area witnesses dry season from late October to February while rains begin in March and pick in June to September. Its characteristic damp weather is believed to be caused by high sensible temperature occasioned by high humidity. The area experiences flood between September and October annually, the October 2022 flood being the worst in history.

Data Collection

The average monthly weather data for thirty-one (31) years (1989 – 2019) used are ERA-5 reanalysis data from Copernicus data hub for Lokoja (average elevation 53m above sea level), Kogi State of Nigeria. The weather data include daily near surface minimum and maximum temperatures, wind speed, humidity and atmospheric pressure. The mean temperature and other unavailable data were estimated using appropriate equations. The estimated data include heat flux, net radiation, and actual and saturated vapour pressures.

The ET evaluation and statistical analysis are carried out using the monthly average weather data from 1989 – 2009. The resulting model was calibrated and validated using 2010 – 2019 weather data.

Model Selection

Penman-Monteith (P-M) model

The Food and Agricultural (FAO-56) Penman-Monteith ET equation is a semi-empirical standard model based on a combination of energy balance and aerodynamic variables. It is recommended by Food and Agricultural Organization (FAO) for most climatic conditions as it is considered to be most physical and reliable method and the sole standard to verify other empirical methods. It has strong fundamental physical principles, including energy balances to precisely calculate the ET.

The model is given as;

$$ET_o = \frac{0.408\Delta(Rn-G) + \Upsilon\frac{900}{Ta+273}u_2(e_s - e_a)}{\Delta + \Upsilon(1+0.34u_2)} \quad (1)$$

Where

ET_o = reference evapotranspiration rate (mm/day);

T_a = mean air temperature (°C);

u_2 = wind speed (m/s) at 2m height;
R_n = the net radiation (MJ/m^{2}day);
G = soil sensible heat flux (MJ/m^{2}day);
e_s = mean saturation vapour pressure (kPa);
e_a = mean ambient vapour pressure (kPa);
Δ = the slope of the saturated vapour pressure-temperature curve (kPa/^{o}C);
ϒ = the psychrometric constant (kPa/^{o}C);

Some estimations and conversions

Wind speed, u.

The wind speed was originally measured at 10m height above the ground level, this is converted to a value at the required 2m height using equation 2, calculated in an Excel spreadsheet.

$$u_2 = u_h \frac{4.87}{\ln(67.8h - 5.42)} \quad (2)$$

Where, u_2 = wind speed at 2m height above the ground surface (m/s);
u_h = measured wind speed at h meters (10m) height above the ground surface (m/s);
h = height of the wind speed measurement above the ground surface (m).

Atmospheric Pressure, P (kPa)

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

Where z is the elevation of the area above sea level (m). For the NIMET site at Lokoja, z is 62.4m.

Psychrometric constant, ϒ

$$ϒ = 0.000665P \quad (4)$$

Where, ϒ = psychrometric constant (kPa/^{o}C);
P = atmospheric pressure, kPa.

Mean Saturation Vapour Pressure, e_s (kPa)

$$e_{(T)} = 0.6108\exp\left[\frac{17.27\,T}{T+237.3}\right] \quad (5)$$

Where T is the air temperature (^{o}C).

For accuracy, the mean saturation vapour pressure is calculated as the mean between the saturation vapour pressure at both the daily maximum and minimum air temperatures.

$$e_{(Tmax)} = 0.6108\exp\left[\frac{17.27\,Tmax}{Tmax+237.3}\right] \quad (6)$$

$$e_{(Tmin)} = 0.6108\exp\left[\frac{17.27\,Tmin}{T+237.3}\right] \quad (7)$$

Therefore,

$$e_s = \frac{eTmax + eTmin}{2} \quad (8)$$

Actual Vapour Pressure, e_a (kPa)

$$e_a = \frac{e_{(Tmin)}\left[\frac{RHmax}{100}\right] + e_{(Tmax)}\left[\frac{RHmin}{100}\right]}{2} \quad (9)$$

In the absence of minimum and maximum relative humidity but mean relative humidity (RH_{mean}), we use

$$e_a = \frac{RH(mean)}{100}\left[\frac{e_{(Tmin)} + e_{(Tmax)}}{2}\right] \quad (10)$$

Where RH is the relative humidity (%).

ET Explanatory Variables

Using the FAO Penman-Monteith model as the response variable, multiple linear regression analysis are performed to check the effects of the explanatory variables (temperature, rainfall, wind speed, relative humidity, vapour pressure deficit, atmospheric cloudiness measured by the ratio of Rs to Rso. The coefficient of determination was noted to ascertain the percentage of the variations in the ET (response variable) that can be explained by the explanatory variables individually and in general, the standard error determines the precision of the estimates, significance F and P-values point out the variables that are statistically significant.

The main goal of performing regression is to understand the data, predictions based on regression line are for average (mean) values, not the actual values; actual values will vary around the mean value. The scatter plots show how "correct" the developed model can be. The scattered residual plots indicate whether the model is appropriate or not.

Variance Inflation Factor (VIF) was performed on the explanatory variables to test for multicollinearity. The variables with multicollinearity problem (VIF greater than 5) were identified and centering or standardization of the variables was performed to reduce the effect of the multicollinearity for keeping the interpretation of coefficients of the developed equation uniform. This is also important because the parameters that were originally statistically insignificant might become significant after normalization. Thereafter, the statistically significant standardized variables were isolated and another round of multiple regression analysis performed to determine their appropriate coefficients which yielded a new regression model with its associated error Ɛ.

Variance Inflation Factor (VIF) Test for Multicollinearity

The VIF is a simple test to assess multicollinearity in a regression model by identifying correlation between independent variables (the explanatory variables) and the strength of that correlation. The VIF calculation is done on each variable in an Excel spreadsheet, using the formula;

$$\frac{1}{1-R^2} \quad (11)$$

Where R^2 is the correlation coefficient of each of the variables been considered as the response variable at a time. Each of the variables is plotted as the response variable against others as explanatory variables and in each case, the R square value is noted and used in the equation (11) above to find the VIF.

Validation of the Developed Model

A set of predicted ET values of the new model developed are obtained using the 2010 – 2019 data and compared with the values from the FAO Penman-Monteith model (taken as observed) within the same period. The Root Mean Squared Error (RSME) and Mean Absolute Error (MAE) are then calculated in Excel spreadsheet to evaluate the accuracy of the developed model.

The residual is calculated as:

$$\text{Residual} = ET_{pm} - ET_{predicted} \quad (12)$$

The error Ɛ in the developed equation is determined as the mean of the residual. This is added to the model equation to correct its coefficient.

$$ET = a + b.VPD + c.u_2 + d.R_n + e.RH + Ɛ \quad (13)$$

Where a, b, c, d and e are numerical coefficients of determination and may be location-based. Ɛ is the correction error.

Breusch-Pagan Test for Heteroscedasticity

Heteroscedacity check was performed on the model using Breusch-Pagan test. This is important in determining whether the results of the new model regression analysis can be trusted.

STEPS:

i. The residual is calculated as $ET_{observed} - ET_{predicted}$.
ii. The residual is squared
iii. The squared residual is plotted as response variable against the explanatory variables in another multiple regression.
iv. The Chi-Square test statistics is calculated on the original data using the formula;

$$\chi^2 = nR^2 \quad (14)$$

Where, n = the number of observations (n = 252 months)
R^2 = the new R squared value from the squared residual plot (0.000368).
$\chi^2 = 252 \times 0.000368$
$\chi^2 = 0.092736$

- Next step is to find the P-value associated with the calculated Chi-Square test statistic using the following command in excel;

P-value = CHISQ.DIST.RT (test statistics, degrees of freedom) (15)

Here, our degree of freedom is
df = n – 1 (16)
= 252 - 1 = 251
Therefore,

P-value = CHISQ.DIST.RT (0.092736, 251)
= 1.00

- A P-value greater than 0.05 shows there is no sufficient evidence to confirm the presence of heteroscedacity in the regression model developed.

RSME and MAE

RMSE (Root Mean Squared Error) and MAE (Mean Absolute Error) are commonly used metrics to evaluate the accuracy of a predictive model in statistics. RMSE is calculated by taking the square root of the average of the squared differences between the predicted values and the actual values. It is often used to measure the standard deviation of the errors the model makes in its predictions. MAE is calculated by taking the average of the absolute differences between the predicted values and the actual values. It is a simpler measure of the average error made by the model.

To calculate RMSE in Excel, we used the following formula:
=SQRT(SUMSQ(predicted values - actual values)/n) (17)

To calculate MAE in Excel, we used the following formula:
=SUM(ABS(predicted values - actual values)/n) (18)

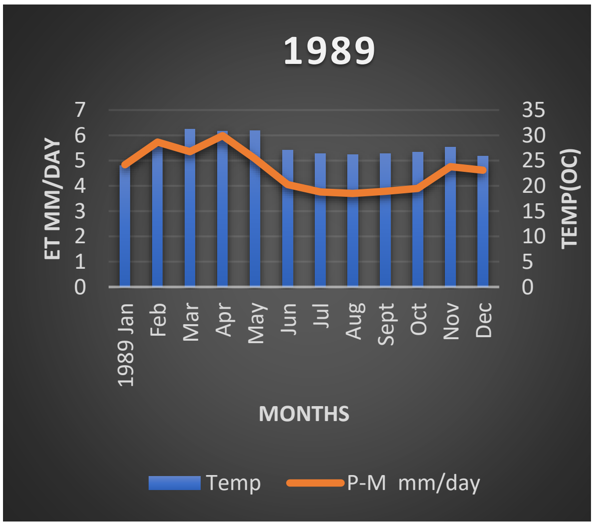
In both formulae, "predicted values" refer to the values predicted by the model, "actual values" refer to the observed values, and "n" is the total number of observations.

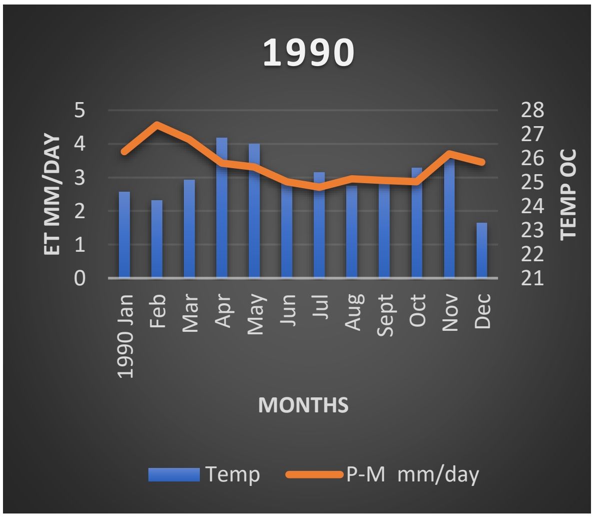
RESULTS AND DISCUSSION

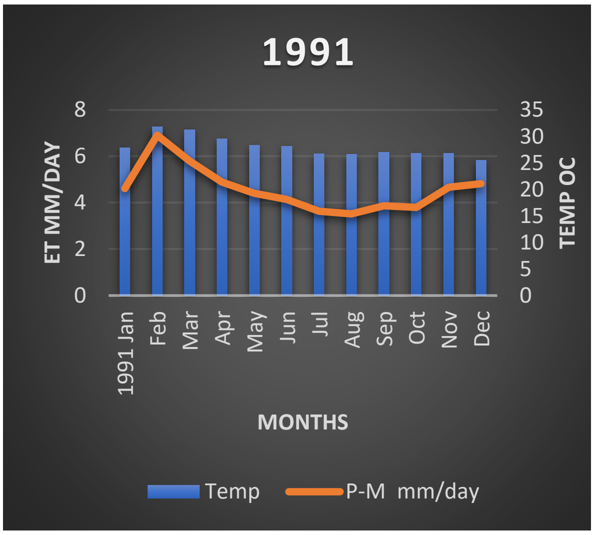
Changes in Climate Variables and ET

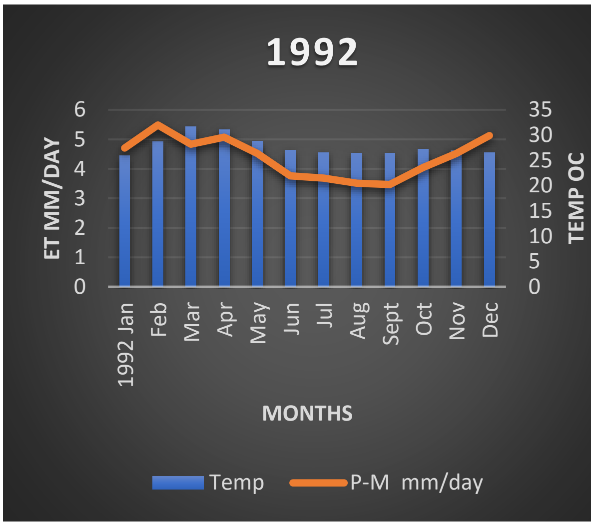
Throughout the 31 years (372 months) under investigation, it was observed that the wet season (April – September) witnessed lower mean daily evapotranspiration (ET) compared with the dry season (October – March) each year, possibly due to higher relative humidity (RH) and lower net radiation (Rn) during wet seasons, in trend with the temperature as observed from the figures. The month of August experienced the least mean daily ET, to show that increased precipitation does not automatically translate to an increase in average ET.

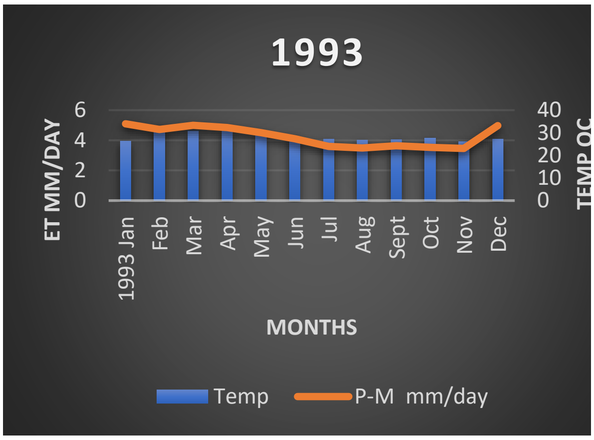
Figure 1 shows there was a decrease in average ET from 1994 to 2008 as mean temperature also remained low within this period compared to the periods before and after; the low temperature contributes to the low ET within the period. The last decade (2009 – 2019) has witnessed increasing mean air temperature, consequently leading to the observed consistent increase in ET as shown in figure 2. The lowest mean daily temperature and ET, 25ºC and 2.0mm/day respectively, were observed within the year 1994 – 1997. In contrast to the maximum mean daily temperature of about 27.5ºC and ET 4.0mm/day observed in the years from 1998 – 2008, the last decade (2009 – 2019) has shown extreme values of mean temperature and ET in the confluence state as high as 30ºC and 6.0mm/day respectively, indicating warming of the region.

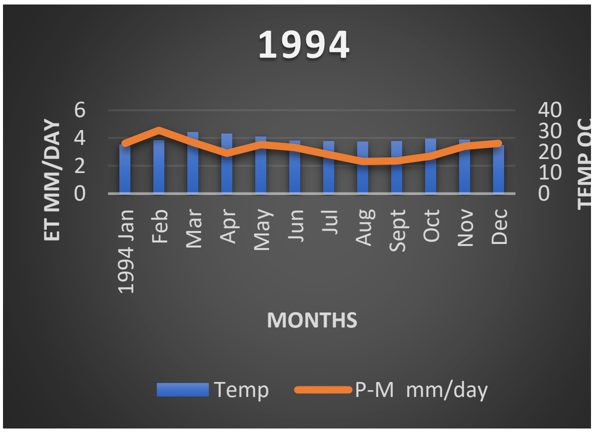


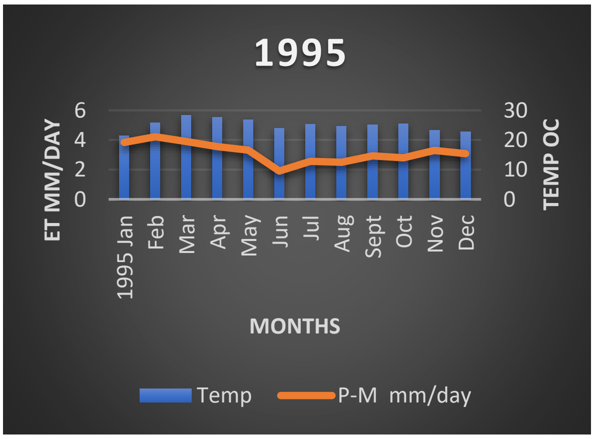


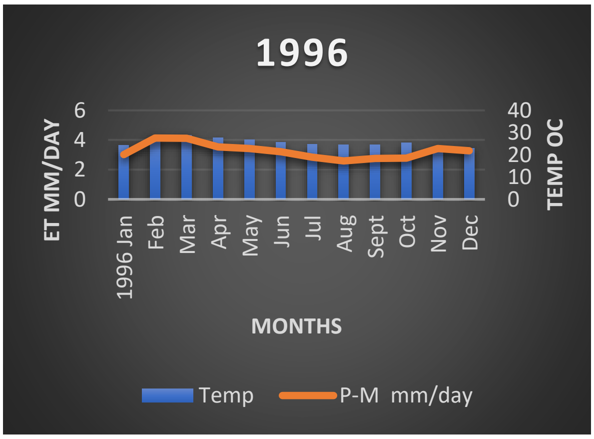

1991
ET MM/DAY
TEMP OC
1991 Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
MONTHS
Temp
P-M mm/day

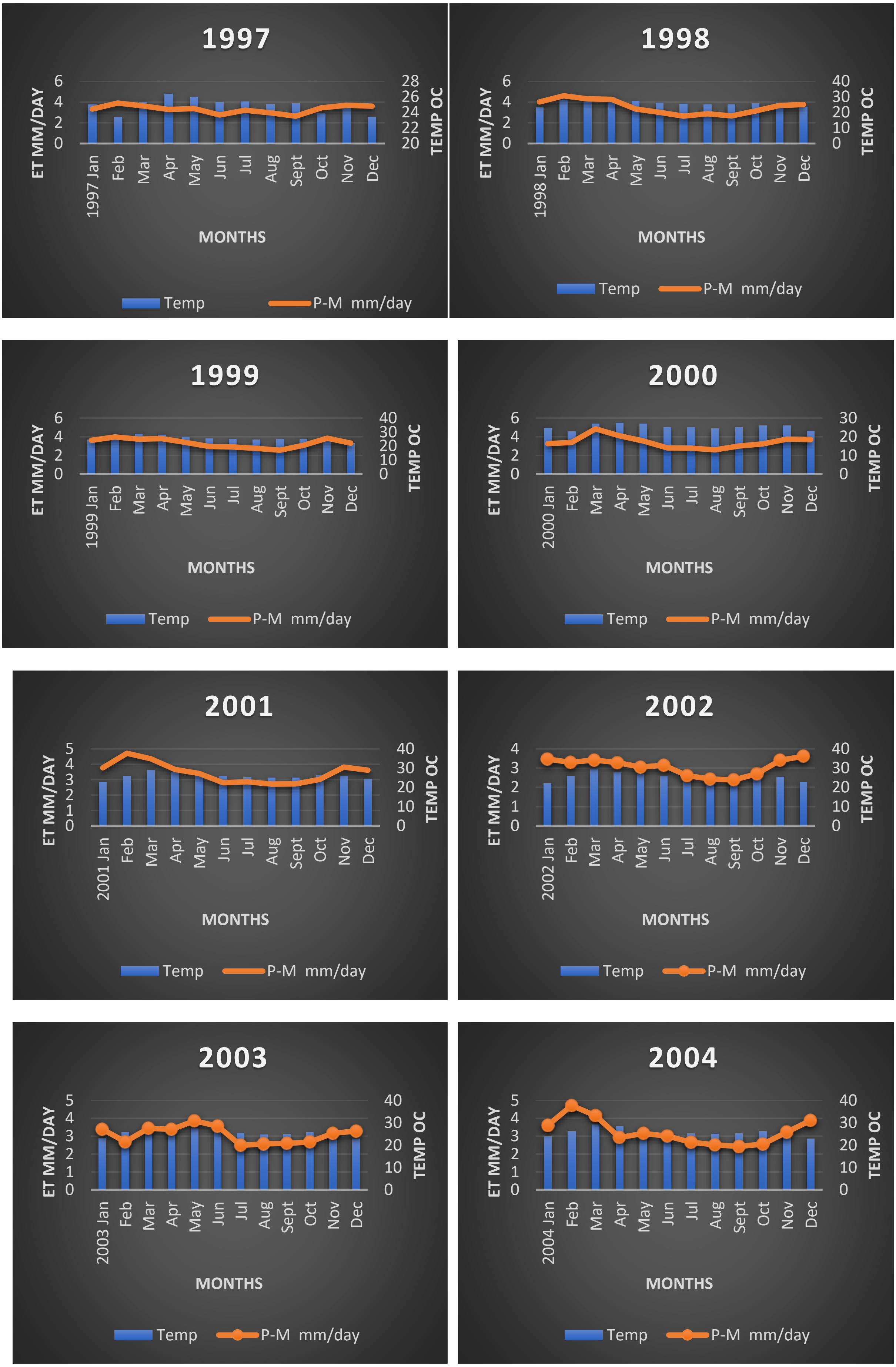

1992
ET MM/DAY
TEMP OC
1992 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

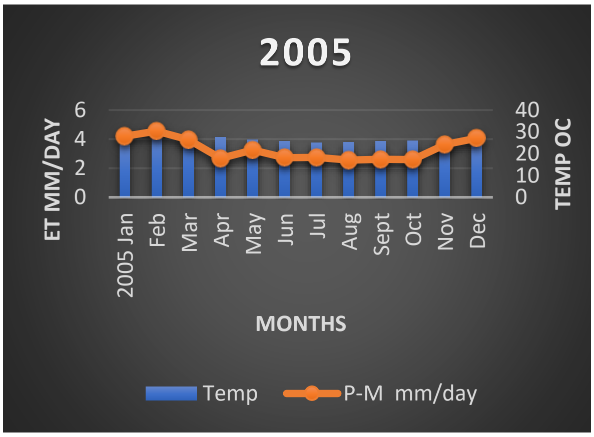

1993
ET MM/DAY
TEMP OC
1993 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

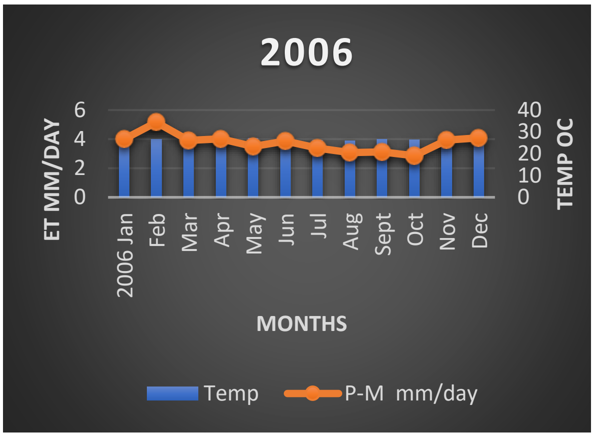

1994
ET MM/DAY
TEMP OC
1994 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

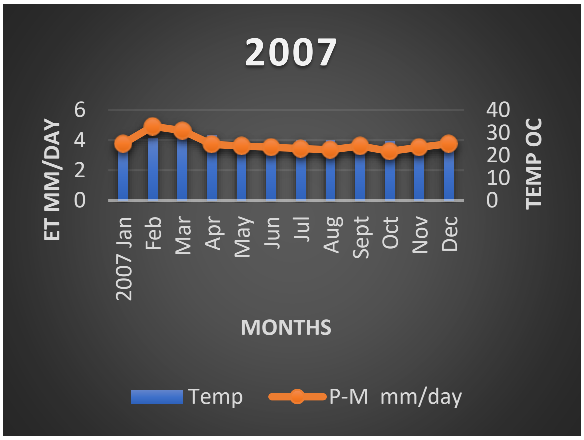

1995
ET MM/DAY
TEMP OC
1995 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

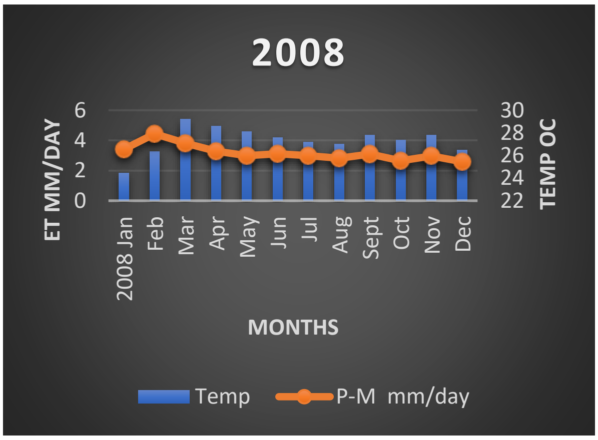

1996
ET MM/DAY
TEMP OC
1996 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

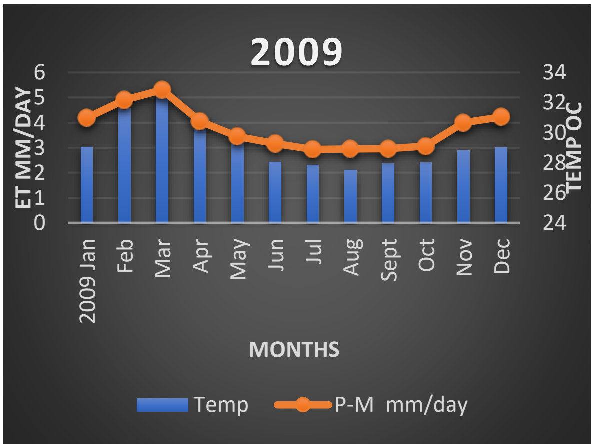

1997
ET MM/DAY
6
4
2
0
28
26
24
22
20
TEMP OC
1997 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
1998
ET MM/DAY
6
4
2
0
40
30
20
10
0
TEMP OC
1998 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
1999
ET MM/DAY
6
4
2
0
40
30
20
10
0
TEMP OC
1999 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
2000
ET MM/DAY
6
4
2
0
30
20
10
0
TEMP OC
2000 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
2001
ET MM/DAY
5
4
3
2
1
0
40
30
20
10
0
TEMP OC
2001 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
2002
ET MM/DAY
4
3
2
1
0
40
30
20
10
0
TEMP OC
2002 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
2003
ET MM/DAY
5
4
3
2
1
0
40
30
20
10
0
TEMP OC
2003 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day
2004
ET MM/DAY
5
4
3
2
1
0
40
30
20
10
0
TEMP OC
2004 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day

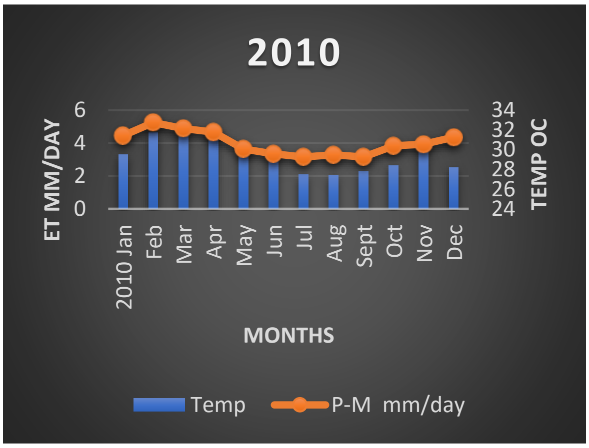

2005
ET MM/DAY
6
4
2
0
40
30
20
10
0
TEMP OC
2005 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

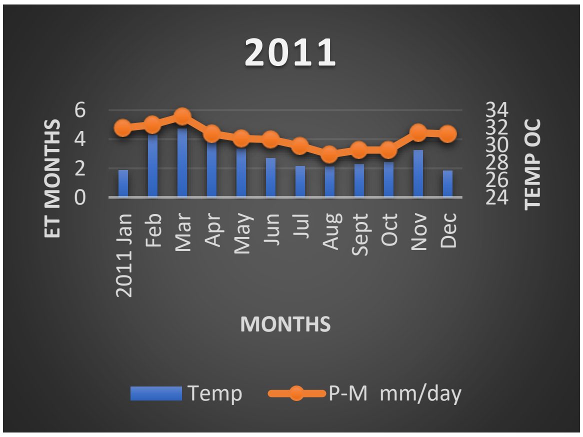

2006
ET MM/DAY
6
4
2
0
40
30
20
10
0
TEMP OC
2006 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

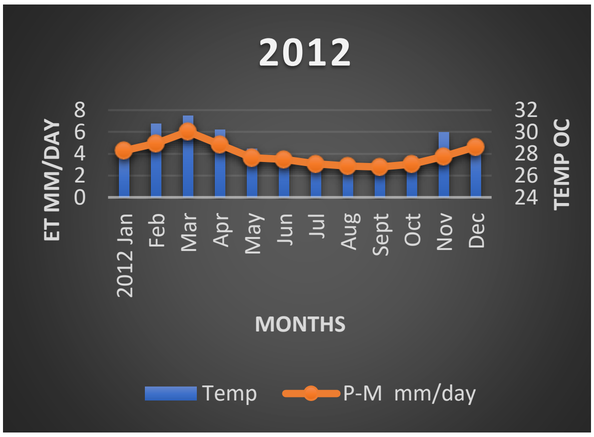

2007
ET MM/DAY
6
4
2
0
40
30
20
10
0
TEMP OC
2007 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

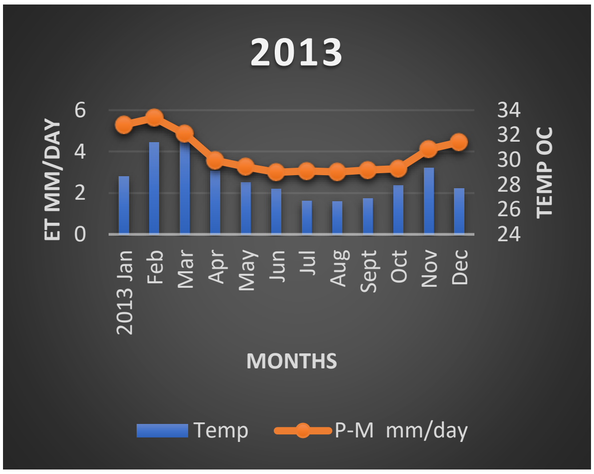

2008
ET MM/DAY
6
4
2
0
30
28
26
24
22
TEMP OC
2008 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

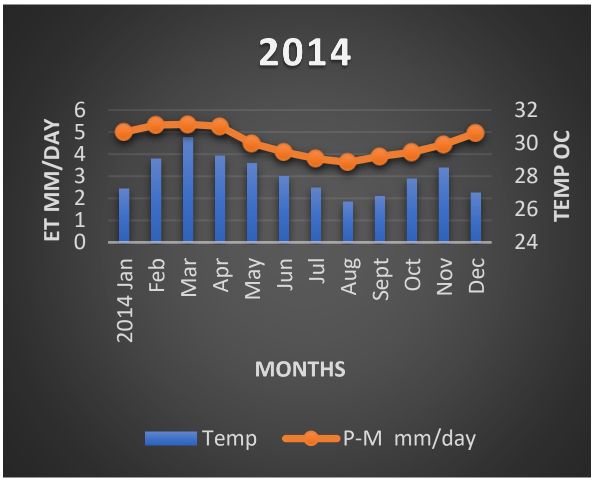

2009
ET MM/DAY
6
5
4
3
2
1
0
34
32
30
28
26
24
TEMP OC
2009 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

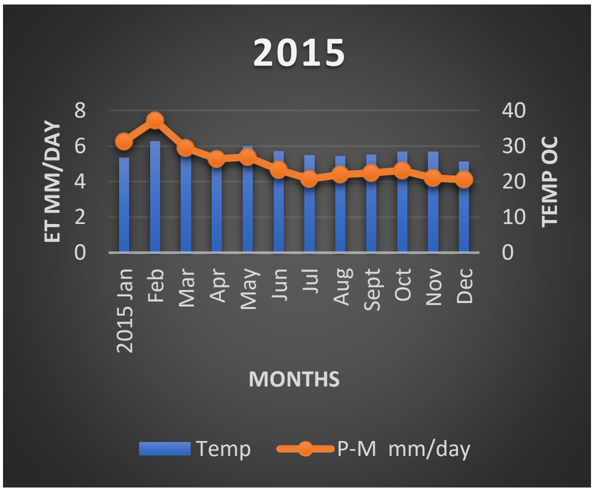

2010
ET MM/DAY
6
4
2
0
34
32
30
28
26
24
TEMP OC
2010 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

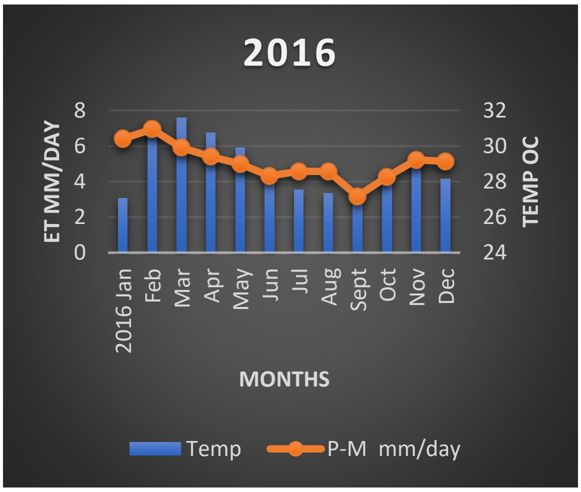

2011
ET MONTHS
6
4
2
0
34
32
30
28
26
24
TEMP OC
2011 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

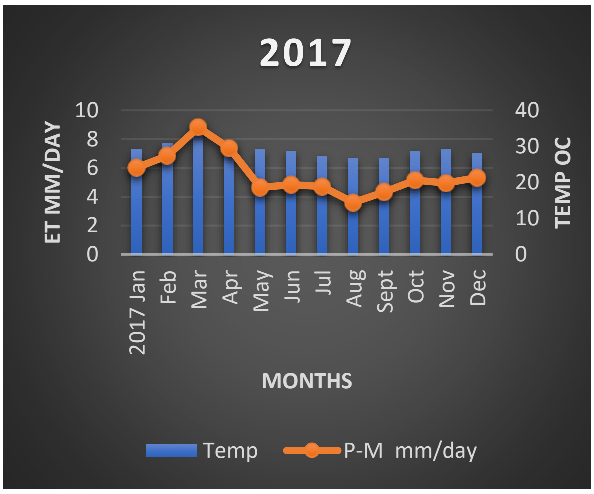

2012
ET MM/DAY
8
6
4
2
0
32
30
28
26
24
TEMP OC
2012 Jan Feb Mar Apr May Jun Jul Aug Sept Oct Nov Dec
MONTHS
Temp
P-M mm/day

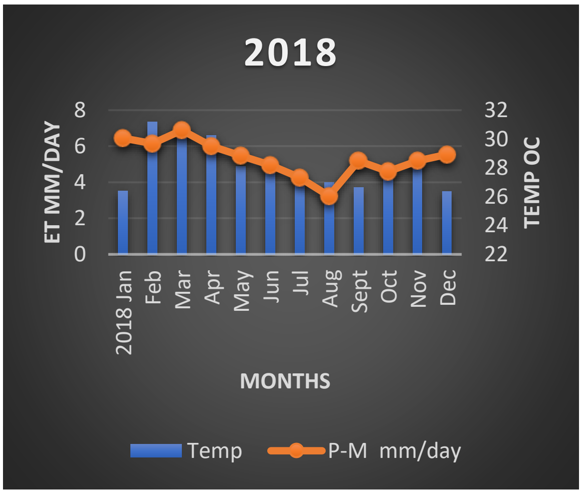

2013
ET MM/DAY
6
4
2
0
34
32
30
28
26
24
TEMP OC
2013 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day


2014
ET MM/DAY
6
5
4
3
2
1
0
32
30
28
26
24
TEMP OC
2014 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day


2015
ET MM/DAY
8
6
4
2
0
40
30
20
10
0
TEMP OC
2015 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day


2016
ET MM/DAY
8
6
4
2
0
32
30
28
26
24
TEMP OC
2016 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day


2017
ET MM/DAY
10
8
6
4
2
0
40
30
20
10
0
TEMP OC
2017 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day


2018
ET MM/DAY
8
6
4
2
0
32
30
28
26
24
22
TEMP OC
2018 Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sept
Oct
Nov
Dec
MONTHS
Temp
P-M mm/day

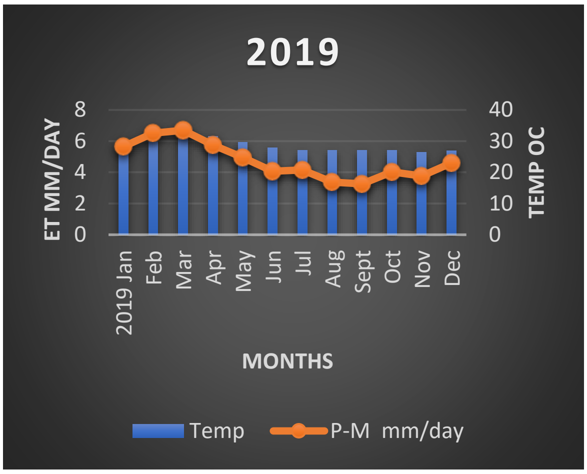


Figure 1: ET Trend with Temperature (1989 – 2019)

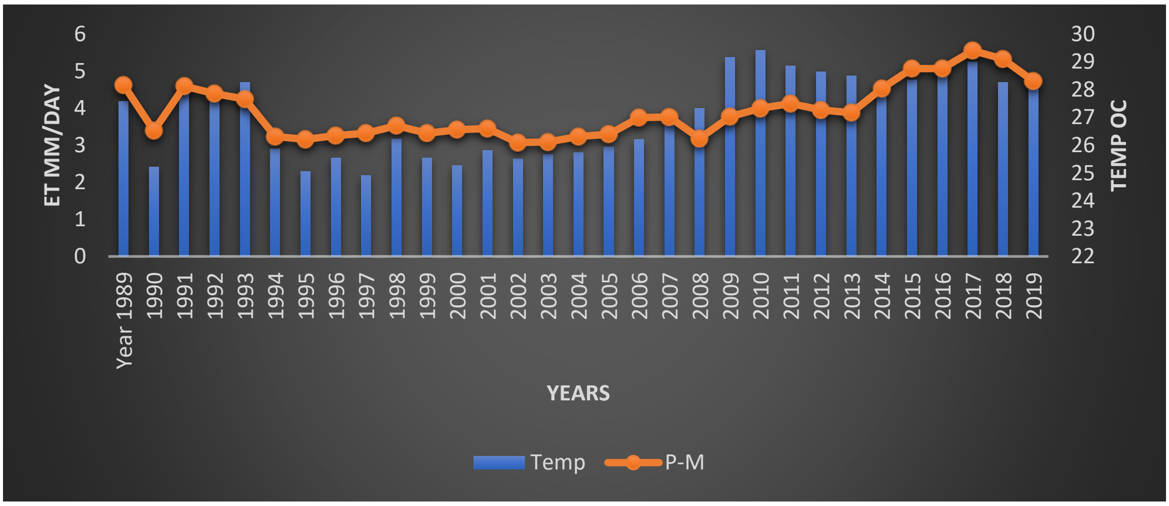


Figure 2: Annual Average Temperature with ET Trend (1989 – 2019)

Models Analysis

Result of the ET evaluation reveals mean daily ET value of 3.91mm/day for Penman-Monteith (P-M) model for the confluence region. Temperature alone is not enough to explain the variability in ET, as shown in Table 1. This reveals that only about 29% (R-square value of 0.29) of the variability in ET can be explained by temperature, according to P-M model, with intercept of about 0.15mm/day for a unit increase in temperature.

Table 1: Results of ET Model Scatter Plot with Temperature

Model	ET (mm/day)	Intercept with Temp.	R^2
Penman-Monteith (P-M)	3.91	0.145	0.29

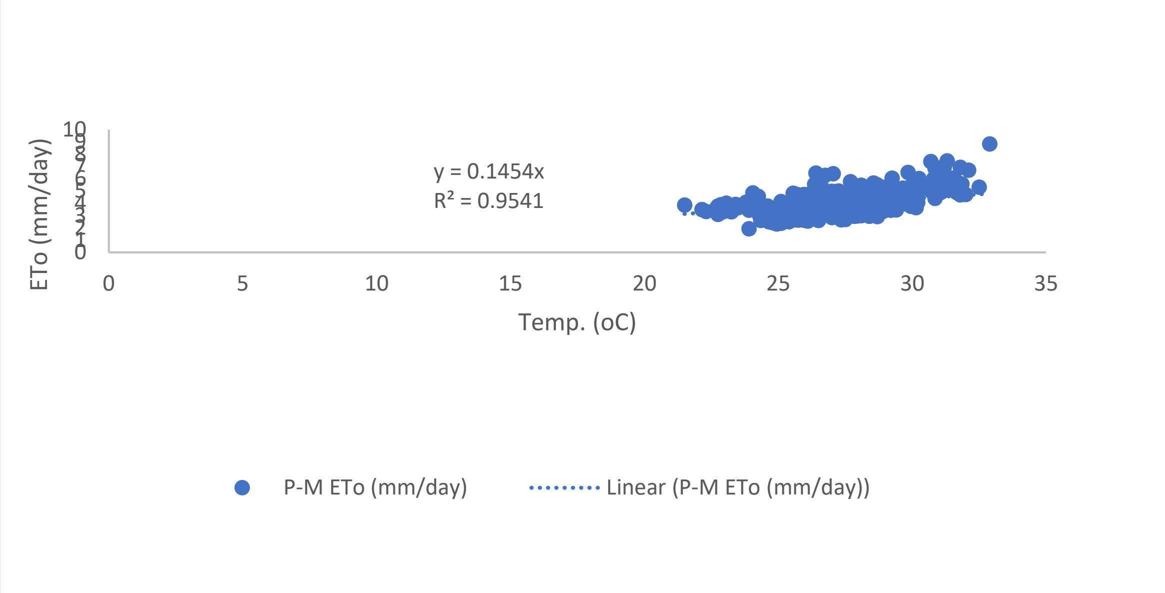


Figure 3: P-M Models' Results with Temperature

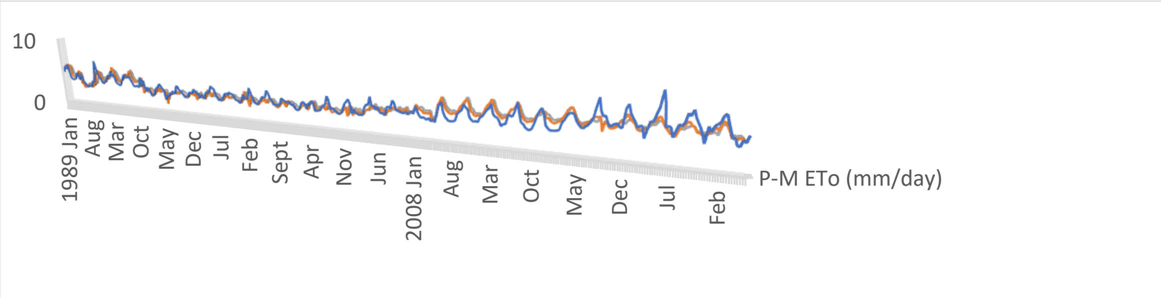


Figure 4: Evapotranspiration Trend

Relationship between Climate Elements and ET

Results of the statistical analysis using multiple regression of the component variables is displayed in Table 2. At 95% confidence level and 251 degree of freedom, the R-square value (0.98279) with standard error as low as 0.10 shows that about 98% of the variability in ET can be explained by the explanatory variables considered (temperature, wind speed, relative humidity, net radiation, vapour pressure deficit, atmospheric cloudiness and rainfall). The table shows that temperature, rainfall and atmospheric cloudiness (R_s/R_{so}) are not statistically significant (so also with the standardized values in Table 3) since their P-values are above 0.05 significant level. The insignificance of these variables may be due to several factors including higher strength of the observed significant variables overshadowing effects of others on ET, which also have temperature, rainfall and or cloudiness as their components (variable interactions). For instance, temperature is a very vital component of vapour pressure while rainfall and cloudiness play significant roles on relative humidity and net radiation.

Table 2: Multiple Regression Statistics of all the Explanatory Variables

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.991357904							
R Square	0.982790494							
Adjusted R Square	0.982296778							
Standard Error	0.103711171							
Observations	252							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	7	149.8764641	21.41092344	1990.601466	2.8499E-211			
Residual	244	2.624465725	0.010756007					
Total	251	152.5009298						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	0.683460577	0.27878464	2.451571855	0.014925447	0.134328995	1.232592	0.134329	1.232592
Ta oC	0.016787261	0.01135668	1.478183829	0.140648268	-0.005582378	0.039157	-0.00558	0.039157
Rainfall (mm)	-2.63308E-05	9.46583E-05	-0.278166636	0.781120017	-0.000212783	0.00016	-0.00021	0.00016
U m/s @2m	0.309991451	0.014500083	21.37859794	7.59202E-58	0.281430144	0.338553	0.28143	0.338553
RH %	-0.018512744	0.004911406	-3.76933704	0.000205349	-0.028186908	-0.00884	-0.02819	-0.00884
Rn MJm-2day-1	0.256518568	0.01789361	14.33576408	3.1809E-34	0.221272917	0.291764	0.221273	0.291764
VPD=(es-ea) Kpa	0.816467699	0.135485995	6.026214741	6.15297E-09	0.54959633	1.083339	0.549596	1.083339
Rs/Rso	0.185915076	0.234111832	0.794129345	0.427891794	-0.275222958	0.647053	-0.27522	0.647053

Table 3: Multiple Regression Statistics of Normalized Explanatory Variables

Regression Statistics								
Multiple R	0.991358							
R Square	0.98279							
Adjusted R Square	0.982297							
Standard Error	0.103711							
Observations	252							
ANOVA								
	df	*SS*	*MS*	*F*	*gnificance F*			
Regression	7	149.8764641	21.41092	1990.601	2.8E-211			
Residual	244	2.624465725	0.010756					
Total	251	152.5009298						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	3.576209	0.00653319	547.391	0	3.563341	3.589078	3.563341	3.589078
Ta oC	0.016787	0.01135668	1.478184	0.140648	-0.00558	0.039157	-0.00558	0.039157
Rainfall (mm)	-2.6E-05	9.46583E-05	-0.27817	0.78112	-0.00021	0.00016	-0.00021	0.00016
U m/s	0.309991	0.014500083	21.3786	7.59E-58	0.28143	0.338553	0.28143	0.338553
RH %	-0.01851	0.004911406	-3.76934	0.000205	-0.02819	-0.00884	-0.02819	-0.00884
Rn MJm-2day-1	0.256519	0.01789361	14.33576	3.18E-34	0.221273	0.291764	0.221273	0.291764
VPD Kpa	0.816468	0.135485995	6.026215	6.15E-09	0.549596	1.083339	0.549596	1.083339
Rs/Rso	0.185915	0.234111832	0.794129	0.427892	-0.27522	0.647053	-0.27522	0.647053

Table 4: Multiple Regression Statistics of Statistically Significant Explanatory Variables

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.9912557							
R Square	0.9825878							
Adjusted R Square	0.9823058							
Standard Error	0.1036846							
Observations	252							
ANOVA								
	df	*SS*	*MS*	*F*	*gnificance F*			
Regression	4	149.8455552	37.46139	3484.617	6.8E-216			
Residual	247	2.655374598	0.010751					
Total	251	152.5009298						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	3.5762092	0.006531518	547.5311	0	3.563345	3.589074	3.563345	3.589074
U m/s	0.3089588	0.014357741	21.51862	1.45E-58	0.28068	0.337238	0.28068	0.337238
RH %	-0.013188	0.002145176	-6.14781	3.13E-09	-0.01741	-0.00896	-0.01741	-0.00896
Rn MJm-2day-1	0.2718801	0.004585835	59.28695	4.5E-148	0.262848	0.280912	0.262848	0.280912
VPD Kpa	0.9828973	0.058344229	16.84652	6.43E-43	0.867982	1.097813	0.867982	1.097813

From Table 3 above, it is obvious that standardizing the variables would not change the values of R-square value (0.98279) and standard error (0.103711). This is in support of the previous researchers' conclusion that multicollinearity does not affect model predictions or goodness-of-fit.

The statistically significant explanatory variables (wind speed u, Relative Humidity RH, net radiation Rn and vapour Pressure Deficit VPD) are isolated and another multiple regression analysis performed on them. The result returned a general intercept of 3.576 with its associated error as shown in the table. The intercept should have been what is estimated to be released as ET when all the variables are zero. From the equation below, the model is corrected by adding an error factor Ɛ.

$$ET = 3.576 + 0.983VPD + 0.309u_2 + 0.272R_n - 0.013RH + \varepsilon \quad (19)$$

ET is evapotranspiration (mm/day);
VPD is vapour pressure deficit (kPa);
u is wind speed (m/s) at 2 meters height above ground level;
Rn is net radiation (MJ/m^2day)
RH is relative humidity.

Correction Error $\mathcal{E}$

The correction error Ɛ is the amount the regression line missed the value of ET. It is calculated as the mean value of the residual, by way of validation, using the 2010 to 2019 data collected. The residual is the difference between the ET (Penman-Monteith) and the predicted ET using equation 19 without the correction error Ɛ. The result shows the mean value Ɛ to be -3.0371.

The model equation becomes;

$$ET = 3.576 + 0.983VPD + 0.309u_2 + 0.272R_n - 0.013RH - 3.0371 \quad (20)$$

$$ET = 0.539 + 0.983VPD + 0.309u_2 + 0.272R_n - 0.013RH \quad (21)$$

All variables carrying their usual meanings and units.

Equation 21 shows that an average ET of 0.5 mm/day is expected when VPD, u, Rn and RH are all near zero (in association with their coefficients). This may be taken to be the 2% contribution to the ET from the lurking variables not accounted for by the considered explanatory variables.

Validation

The resulting equation above is used to predict ET with set of data (2010 – 2019) and the result compared with the FAO Penman-Monteith values taken as standard (observed). Below are the 3-D line chart and scatter plots of the results. The chart and scatter plots depict strong, positive, linear relationship between the predicted and "observed" values. The residual scatter plot shows no particular pattern which further confirms the correctness of the developed model.

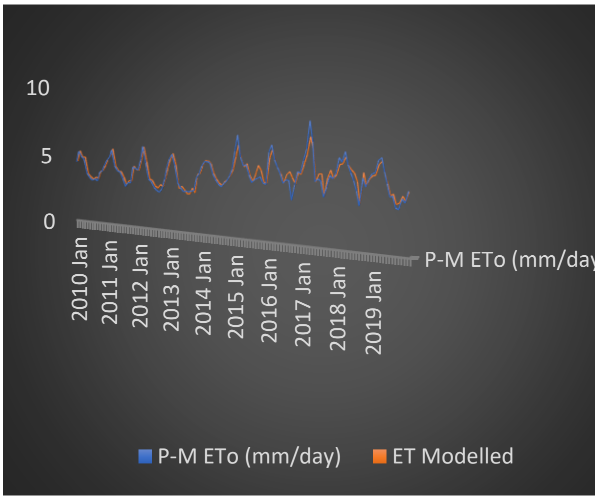


Figure 5: 3-D Line Chart of $ET_{p\text{-}m}$ and $ET_{predicted}$

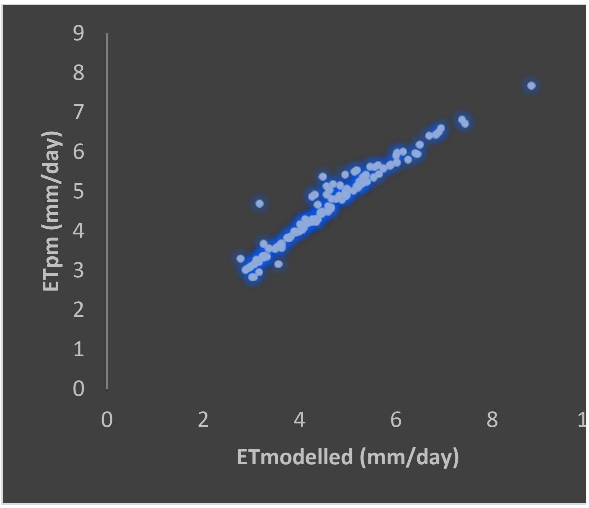


Figure 6: Scatter Plot of the ET Modelled

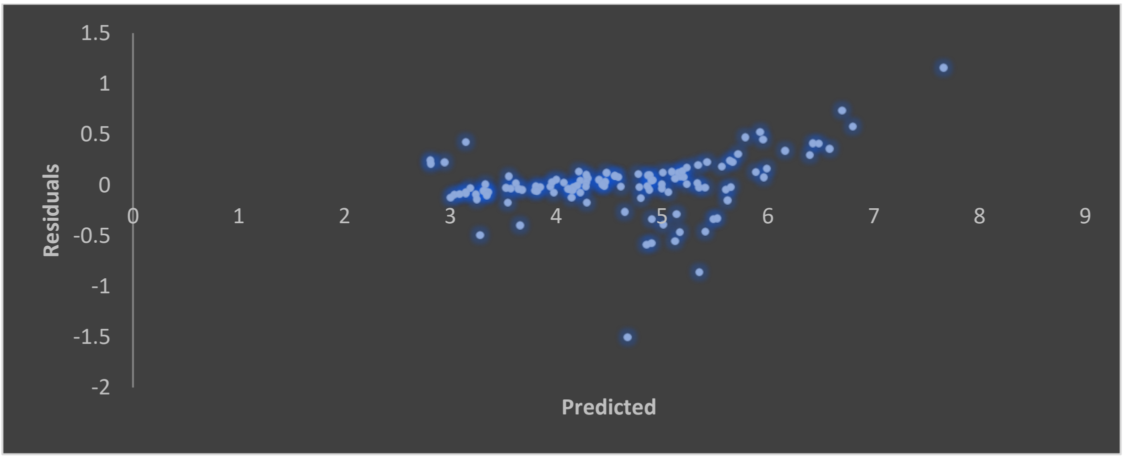


Figure 7: Residual Scatter Plot

RSME and MAE

The result of the Root Squared Mean Error (RSME) and Mean Absolute Error (MAE) are presented in table 5 below. The RSME of approximately 0.001 and MAE of 0.005 indicate that the developed model is performing well in predicting ET, with the error between the predicted and observed values being very low so can be considered reliable for making future predictions.

Table 5: RSME and MAE Results

Year	ETpm(mm/day)	ETpredicted(mm/day)	RSME	MAE
2010	3.995278	3.999996	0.001492	0.000472
2011	4.127839	4.136022		
2012	3.95203	4.010011		
2013	3.876911	3.813746		
2014	4.530383	4.524014		
2015	5.070424	5.066522		
2016	5.071683	5.139474		
2017	5.552576	5.465518		
2018	5.328667	5.341141		
2019	4.731162	4.701561		

Variance Inflation Factor (VIF) Test Result

Table 6 shows the summary of the VIF test conducted on the statistically significant variables. A VIF value of 1 indicates that there is no correlation between this independent variable (wind speed) and others therefore declared useful without need for corrective measure. VIF between 1 and 5 (Net Radiation) shows there is a moderate correlation but not severe enough to warrant corrective measures while values greater than 5 (Relative Humidity and Vapour Pressure Deficit) suggest critical levels of multicollinearity where the coefficients are poorly estimated (becoming inflated or reversed in sign) and the p-values are questionable. Simply centering the variables helps reduce the correlation between predictor variables because it changes the scale of the variables and makes them more comparable to each other, and help improve the stability and accuracy of regression coefficients which in turn make it easier to interpret the results of the analysis.

Table 6: VIF Result Summary

Variable	VIF Value
Wind Speed	1.00
Relative Humidity	12.03*
Net Radiation	1.60
Vapour Pressure Deficit (VPD)	12.41*

Table 7: VIF Multiple Regression Statistics for Relative Humidity

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.957518
R Square	0.91684
Adjusted R Square	0.915834
Standard Error	3.069204
Observations	252

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	3	25756.36	8585.452	911.4056	1.4E-133
Residual	248	2336.163	9.420012		
Total	251	28092.52			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	86.23718	1.379689	62.50477	8.5E-154	83.51978	88.95458	83.51978	88.95458
U m/s @2m	0.047422	0.424998	0.111582	0.911245	-0.78964	0.884487	-0.78964	0.884487
Rn MJm-2day-1	1.267609	0.109307	11.59678	3.94E-25	1.052321	1.482898	1.052321	1.482898
VPD=(es-ea) Kpa	-26.0402	0.498521	-52.2348	7.9E-136	-27.0221	-25.0583	-27.0221	-25.0583

Table 8: VIF Multiple Regression Statistics for Net Radiation

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.610929							
R Square	0.373235							
Adjusted R Square	0.365653							
Standard Error	1.435721							
Observations	252							
ANOVA								
	df	*SS*	*MS*	*F*	*gnificance F*			
Regression	3	304.4169	101.4723	49.22743	5.38E-25			
Residual	248	511.2014	2.061296					
Total	251	815.6184						
	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	-18.6711	2.360658	-7.90926	8.55E-14	-23.3206	-14.0216	-23.3206	-14.0216
RH %	0.27738	0.023919	11.59678	3.94E-25	0.23027	0.324489	0.23027	0.324489
U m/s @2m	0.06866	0.198764	0.345436	0.730059	-0.32282	0.460141	-0.32282	0.460141
VPD=(es-ea) Kpa	7.755731	0.640425	12.11028	7.89E-27	6.494364	9.017097	6.494364	9.017097

Table 9: VIF Multiple Regression Statistics for VPD

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.958848							
R Square	0.919389							
Adjusted R Sc	0.918414							
Standard Erro	0.112847							
Observations	252							
ANOVA								
	df	*SS*	*MS*	*F*	*gnificance F*			
Regression	3	36.01957	12.00652	942.8354	3E-135			
Residual	248	3.158152	0.012734					
Total	251	39.17772						
	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*pper 95.0%*
Intercept	3.081486	0.06945	44.37012	5.9E-120	2.9447	3.218272	2.9447	3.218272
Rn MJm-2day	0.047914	0.003956	12.11028	7.89E-27	0.040122	0.055707	0.040122	0.055707
RH %	-0.0352	0.000674	-52.2348	7.9E-136	-0.03653	-0.03388	-0.03653	-0.03388
U m/s @2m	0.000701	0.015626	0.044889	0.964232	-0.03008	0.031479	-0.03008	0.031479

CONCLUSION

In this research, we have conducted a statistical analysis of Lokoja weather variations, evapotranspiration and its explanatory variables over three decades (1989 – 2019). It was observed that the wet months (April – September) witnessed lower ET compared with the ET during the dry months (October – March) due to variations in humidity and net radiation. The results also revealed that temperature alone is not enough to explain the variability in ET, temperature contributes only about 29% explanation according to statistical result from Penman-Monteith model. The explanatory variables that are statistically significant include wind speed, net radiation, relative humidity and vapour pressure deficit.

Generally, the analyzed explanatory variables in this research provided about 98% explanation for the variability in ET. It is observed that ET is affected positively by average wind speed (u), net radiation (Rn), and vapour pressure deficit while average relative humidity (RH) has negative effect on ET from the multiple regression analysis. The result shows a mean daily increase of ET, starting from 0.54mm/day, by about 0.31mm/day for every unit increase in wind speed (u) at two meters height, 0.27mm/day for a unit increase in net radiation (Rn), 0.98mm/day for a unit increase in vapour pressure deficit (VPD), and a decrease by 0.01mm/day for every unit increase in relative humidity (RH).

Assuming linear and independent relationship between the explanatory variables, and perhaps ET, we have proposed a linear model with location-based coefficients for estimating ET, and the problem of rigorous steps involved in Penman-Monteith equation is simplified in this study. The fact that multicollinearity problem does not affect model predictions and goodness-of-Fit is further confirmed in this research. Though this provides easier way of estimating ET in the confluence region thereby improving water resource management and planning in the area, factors such as limited data availability, non-linear relationships, regional and temporal variabilities and some other external factors like land use changes and irrigation practices may limit the performance of this model in other regions. We recommend further testing of this model in other regions, exploring additional explanatory variables, and integrating satellite data for more comprehensive ET analysis and improvement of the linear model.

The newly developed linear ET model can be integrated into existing water management systems by incorporating it into existing irrigation scheduling tools or software. Farmers could input the data into the model, which would then calculate the optimal irrigation schedule based on evapotranspiration rates. This information could be used to more efficiently allocate water resources and reduce water waste. Alternatively, local farmers could directly use the model themselves to optimize their irrigation practices. By regularly monitoring ET rates and adjusting irrigation schedules accordingly, farmers could ensure that their crops are receiving the right amount of water at the right time, leading to improved crop yields and water usage efficiency. In both cases, the integration of the linear ET model could help to improve overall water management practices, leading to better resource allocation, increased crop productivity, and potentially cost savings for farmers.

REFERENCES

Animashaun, I. M., Adeoye, P. A., Otache, M. Y., & Abatan, O. A. (2020). Rainfall variability and trend analysis over Lokoja, Nigeria. *Covenant Journal of Engineering Technology, 2020, vol. 4(2), 2682-5317.*

Ashaolu, E. D., & Iroye, K. A. (2018). Rainfall and potential evapotranspiration patterns and their effects on climate water balance in the Western Lithoral hydrological zone of Nigeria. *Ruhuna Journal of Science, Vol. 9(2): 92-116. ISSN: 2536-8400*. http://doi.org/10.4038/rjs.v9i2.45

Banik, P., Kumar, N. T., & Ranjan, S. (2012). Comparative crop water assessment using meteorological data and modeling techniques. Global Sustainability Transitions: Impacts and Innovations. *ISBN: 978-93-83083-77-0*

IPCC, (2023): *Climate Change 2023: Synthesis Report.* Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, 35-115. doi: 10-59327/IPCC/AR6-9789291691647

Joshua, M. M., Ambrose, A. Z., Emeka, D. O., & Mohammed, B. Y. (2023). Climate change variability in the upper Benue river basin, Nigeria. *Journal of Agriculture and Forest Meteorology Research, JAFMR, 6(2): 522-535.* www.scitcentral.com

Kelly, L., Sophie, B., & Rebecca, C. (2022). 6 Big findings from the IPCC 2022 report on climate impacts, adaptation and vulnerability. World Resources Institute, February 27, 2022.

Lokoja at a Glance. Accessed January 20, 2024. From https://www.fulokoja.edu.ng>aboutus

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to linear regression analysis.John Wiley & Sons.

Mundo-Molina, M. (2015). Climate change effects on evapotranspiration in Mexico. *American Journal of Climate Change. 2015,4,163-172.* http://dx.doi.org/10.4236/ajcc.2015.42012

Satish D. (2018). Penman Monteith equation. www.slideshare.net

Yates, D., & Strzepek, K. (1994): Potential Evapotranspiration Methods and their Impact on the Assessment of River Basin Runoff under Climate Change; WP-94-46. International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria. www.pure.iiasa.ac.at