BILINGUAL PUBLISHING GROUP

Journal of Atmospheric Science Research

https://journals.bilpubgroup.com/index.php/jasr

ARTICLE

Analysis of Evaporation Variables in Open Water and Penman-Monteith Evaporation Variables Using Ridge Regression [Case Study Forest Fire Model]

Endar H. Nugrahani [1], Sri Nurdiati [1*], Ardhasena Sopaheluwakan [2], Pandu Septiawan [3], Willy Pratama [1]

[1]*Department of Mathematics, IPB University, Bogor 16680, Indonesia*

[2]*Climate and Air Quality Research Agency official Meteorological, Agency for Meteorology Climatology and Geophysics, Jakarta 10720, Indonesia*

[3]*Engineering Mechanics and Energy, University of Tsukuba, Tsukuba 305-8573, Japan*

ABSTRACT

Forest and land fires constitute a calamity influenced by actual evapotranspiration (act) derived from ERA5-land data, evaporation from open water (pev) derived from ERA5-Land data, and Penman-Monteith evaporation (*et0*). This study aims to scrutinize the characteristics, spatial and temporal aspects of these three variables, and assess the efficacy of *pev* and *et0* in modeling forest fires through Ridge Regressions. Forest fire-related data, acquired via the Empirical Orthogonal Function method based on Singular Value Decomposition, will be employed. The analysis characteristic, temporal, spatial, and ridge regression assessments, with hotspot data serving as an indicator for forest and land fires. The findings reveal that the *pev* variable requires the fewest computation indicators. The characteristics of the *act* variable closely resemble those of the *et0* variable. All three variables exhibit similar spatial and temporal patterns in forest fire-related data. The performance of the *pev* in modeling forest fires is better than the reference evapotranspiration variable in most local region analysis. This result highlights the potential applications of *pev* in modeling the general characteristics of fire events, making it robust across different regions. However, for modeling extreme events, *et0* proves to be superior to other analyzed variables, particularly for events not included in the training data. Thus, et0 has higher potential performance in analysing

*CORRESPONDING AUTHOR:

Sri Nurdiati, Department of Mathematics, IPB University, Bogor 16680, Indonesia; Email: nurdiati@apps.ipb.ac.id

ARTICLE INFO

Received: 20 July 2024 | Revised: 14 September 2024 | Accepted: 10 October 2024 | Published Online: 15 October 2024

DOI: https://doi.org/10.30564/jasr.v7i4.6916

CITATION

Nugrahani, E.H., Nurdiati, S., Sopaheluwakan, A., et al., 2024. Analysis of Evaporation Variables in Open Water and Penman-Monteith Evaporation Variables Using Ridge Regression [Case Study Forest Fire Model]. Journal of Atmospheric Science Research. 7(4): 51–65. DOI: https://doi.org/10.30564/jasr.v7i4.6916

COPYRIGHT

Copyright © 2024 by the author(s). Published by Bilingual Publishing Group. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License (https://creativecommons.org/licenses/by-nc/4.0/).

and modeling future impact of climate changes related to fires risk rise due to more frequent drought condition.
Keywords: Evaporation; Evapotranspiration; Forest and Land Fires; Characteristics; Performance; Modelling

1. Introduction

Forest fires represent a natural phenomenon frequently observed in Indonesia[1]. Various factors contribute to the occurrence of forest fires in Indonesia, including the presence of hotspots, El Niño, and drought[2]. Drought leads to a reduction in vegetation water content, resulting in plant mortality and an elevated risk of fires[3]. One repercussion of drought is an increased susceptibility to forest fires due to the drying out of organic material, making it more prone to ignition[4]. Presently, not only is the intensity of drought on the rise, but the impact and geographic distribution of drought are also expanding[5]. Consequently, measures need to be implemented to mitigate the effects of drought and reduce its impact. One approach involves monitoring drought using the drought index. Drought index is influenced by various parameters such as potential evaporation and actual evapotranspiration[6]. Low air humidity levels are another contributing factor to forest fires, causing the surrounding air to become dry due to diminished water content[7].

Predicting forest fires can be achieved by analyzing available data on potential evaporation and actual evapotranspiration. Evaporation, the process of converting liquid molecules into gas through heat energy utilization (e.g., solar heat), and evapotranspiration, the evaporation process occurring in plants through transpiration[8], play crucial roles in forest fire analysis.

Numerous studies have explored evaporation and evapotranspiration, linking them to forest fires. In fact, most of the fire's mitigation index uses combination of evapotranspiration and precipitation[9]. Clear physical meaning of evapotranspiration facilitates the understanding of its relationship to fire severity[10]. Besides the usability to understand fire dynamics, evapotranspiration is also important since its affect ecosystem recovery after wildfires[10]. Especially in dense, mid-elevation forests such as in Nevada which suffers high dropped of evapotranspiration after wildfires[11]. In semi-arid to arid regions, increased evapotranspiration affects water availability and soil salinization, which also lead to higher probability of forest fire. Meanwhile, significant changes after wildfires can influence groundwater recharge if a critical canopy-loss threshold is reached, with site-specific outcomes. The recovery of evapotranspiration post-fire seems to be influenced by climate and fire severity[12]. This structured qualitative review lays the groundwork for interdisciplinary researchers and water managers to utilize evvapotranspiration in addressing complex environmental issues in the fires analysis[13].

For Indonesia forest fire, research in[8] investigated the relationship between evapotranspiration, rainfall, and altitude in tropical Indonesia. Fundamentals research of evaporation and evapotranspiration, aiming to maximize the use of evapotranspiration data in managing irrigation and estimating daily water use[14]. Comparative research on evaporation data from open water[15] and Penman-Monteith modeling[16] is infrequently conducted, offering an interesting avenue for further analysis. Despite the importance, the availability of research that analyse evapotranspiration and fires event in Indonesia is very limited resulted in relatively low understanding level of both phenomena. However, Comparison between various variables reveals that evapotranspiration reduce the difference between multiple characteristic fires pattern in Indonesia[17]. This research shows the possibility of developing fire analysis that applicable for all Indonesia regions by giving one principal component pattern that covers all fires region of Indonesia, covering more than 95% of historical fire event. Thus, objective of this research is conducted to assess potential uses of evapotranspiration in estimating forest fire in Indonesia.

One reason that could contribute to the limited amount of research in evapotranspiration is the complexity of measurement. Different method of measurement and calculation could result in different characteristic of variables, which could lead to inconsistent result. Thus, this research will also assess the difference between evaporation, evapotranspiration, and potential evapotranspiration especially when linked to fires event. Empirical Orthogonal Function (EOF) and Singular Value Decomposition (SVD) methods are employed to extract data values related to forest and land fires.

Evaporation is a process of molecule transformation

from liquid to gas, utilizing heat energy (commonly from the sun)[8]. The loss liquid usually came from open and wet surface[18]. Evapotranspiration is an evaporation process that happened on plant from transpiration processes. 20% of water that absorbed through the root for photosynthesis generally are loss during the transpiration process[8, 19]. Potential evaporation (potential evapotranspiration) are uses commonly in hydrology[20], introduced to differentiate between actual and potential water loss during the process[21]. Analysis that uses to assess in this research is distribution and regression analysis to understand difference between each variable as well as their relation to the fires event in Indonesia. By doing so, this research will reveal difference between characteristic of used variable and potential uses each of them in the fires. Provide better understanding of how evapotranspiration impact fires event in Indonesia, thus become valuable knowledge when moving forward to analysis more complex index such as standardized precipitation evapotranspiration index and fire weather index in Indonesia's fires event research.

2. Materials and Methods

2.1. Data

The data used in this research was downloaded from https://cds.climate.copernicus.eu/ with the dataset name ERA5-Land monthly averaged data from 1950 to present and the variables used were total evaporation and evaporation from open water from January 2001 until December 2020[21]. Penman-Monteith potential evaporation data for 2001–2020 has been calculated previously[17]. Apart from that, hotspot data for the Indonesian region for 2001–2020 comes from the Moderate Resolution Imaging Spectroradiometer (MODIS). The hotspot data used was obtained from MODIS with a spatial resolution of 0.1°×0.1°. This data is data on the number of monthly hotspots from 2001–2020[22].

ERA5-Land data for the variable's total evaporation and evaporation from open water is Indonesian territory data at coordinates 6° N–10° S and 95° E–141° E which has a spatial resolution of 0.1°×0.1° with units of m of water equivalent (m.w.e). According to the guidelines from ERA5-Land, the total evaporation variable is also referred to as actual evapotranspiration. Actual evapotranspiration, defined as standard evapotranspiration from the land surface with an unlimited water supply[15], accumulates four types of evaporation: from open water, canopy, open ground, and vegetation transpiration. This variable is the sum of four other variables, namely evaporation from open water which is also known as potential evaporation, evaporation from vegetation transpiration, evaporation from bare soil, and evaporation from top of canopy. In this study, the total evaporation variable is referred to as actual evapotranspiration (act) and the evaporation from open water variable is referred to as evaporation from open water (pev). The data obtained is in NetCDF format, so an extraction process is needed to obtain the data matrix. In this research potential evaporation data (et0) data for the Indonesian region is calculated using ERA5 data, explained in previous research[17] has a spatial resolution of 0.25°×0.25° with units of m.w.e.

Potential evaporation, often considered as evaporation from open water, can be calculated using the Penman-Monteith model, requiring multiple indicators such as rainfall, radiation, temperature, vapor pressure, ground heat flux, air density, altitude, and wind speed[16].

The *et0* data has a spatial resolution of 0.25° so that resolution equalization with other data needs to be done before analysis. In this study, the resolution of all data was made the same so that the resolution of the *et0* data had to be changed to 0.1°. This process is called interpolation. The method used for is nearest-neighbor interpolation. This interpolation is available in MATLAB software using the "interp1" function. This method is used for several reasons, namely a fast-computing process, sufficient calculations requiring two data points and minimal memory usage. Comparison of spatial graphs between before and after the downscaling process using the nearest-neighbor interpolation method is shown in **Figure 1**.

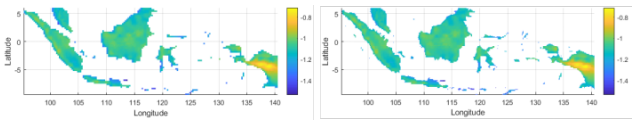

Figure 1. Spatial graph of et0 data before Interpolation (**left**) and after downscaling (**right**).

From the analytical perspective, potential evapotranspiration could be calculated using Penman-Monteith model which used in Food and Agricultural Organization (FAO) guideline for uses in hydrological research. The model follow equation bellow:

$$PET = \frac{\Delta (R_n - G) + \rho_a C_p \left(\frac{e_s - e_a}{r_a} \right)}{\Delta + \gamma \left(1 + \frac{r_s}{r_a} \right)}, \quad (1)$$

with: PET is potential evaporation, Δ is the slope of the saturated vapor pressure, R_n is net solar radiation, G is the ground heat flux, ρ_a is the average air density under isobaric conditions, C_p is the specific air heat, $e_s - e_a$ is the lack of vapor pressure, γ is the pisometric constant, r_s is the surface resistance, r_a is aerodynamic endurance [23, 24] .

2.2. Singular Value Decomposition

EOF analysis is carried out to transform n original variables that are correlated with each other into k orthogonal (uncorrelated) components [25]. Let The SVD of X with rank (X) = *r* is so that we get:

$$X = A_{m \times m} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}_{mxn} B^T_{n \times n}, \quad (2)$$

Matrix B is the EOF matrix or vector coefficients and AΣ is the EOF score matrix or principal components. The principal component score is written as $z_i = a_i \sigma_i$, $i = 1, 2, \ldots, r$. Furthermore, the variance of the ith principal component, i = 1,2, …, r is:

$$\mu_i = \frac{\sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \quad (3)$$

with i = 1,2,…,r being the singular values of X. In its use, k mode EOF 1 or the first principal component with $k \ll r$ explains the largest proportion of variance. EOF 2 mode is a linear combination of all observed variables with orthogonal properties to EOF1 mode and has the second largest variance and so on. Therefore, the kth EOF mode has the kth maximum variance and is uncorrelated with the previous EOF modes. EOF analysis is used to find $m \times n$ component score matrices from m observation objects and n times [26].

2.3. Data Characteristic Analysis

Time series data is a series of observations based on time sequence. A time series data can consist of one of four main components, namely trend, cycle, seasonality and irregularity. According to previous research [27], one way to compare temporal plots is to use distance measurements. One commonly used distance measurement method is Minkowski Distance or L_p-norm. The Minkowski Distance value of the two temporal data vectors is calculated using the formula in equation below:

$$d(A, B) = \sqrt[p]{\sum_{i=1}^{n} (A_i - B_i)^p}, \quad (4)$$

with p: Order of Minkowski Distance. Furthermore, the probability density function is a basic characteristic that describes the nature or behavior of a random variable [28]. This function can be estimated using the non-parametric kernel method from a random variable. According to [28], the kernel density function estimation formula can be written as follows:

$$f(x_j) = \frac{1}{hn} \sum_{i=1}^{n} K \left(\frac{x_j - X_i}{h} \right), \quad (5)$$

with n is the sample size, h is the smoothing parameter (Bandwidth), x_j is the *jth* value to be estimated (j = 1,2,…,n), X_i is the ith sample point (i = 1,2,3,…,n), K is a kernel function.

Real world data is known to have high variability of data points. Therefore, distribution analysis is very important in the modeling. This research uses quantile-quantile plot and Jensen-Shannon Divergence to analyze the difference between analyzed variables. QQ plot with two data shows the difference in two or more distributions [29]. If the two data have a similar distribution, the resulting QQ plot will be close to linear. Meanwhile, Kullback-Leibler Divergence (KL-Div) is a measure of the large amount of information lost when using an approximation (estimation) [30]. Suppose there are two distributions P and Q, the KL-Div value of Q from P can be written as follows:

$$KL(P, Q) = \int P(x) \cdot log \frac{P(x)}{Q(x)} dx. \quad (6)$$

KL-Div is an asymmetric measurement and does not satisfy the triangle inequality [31]. Therefore, KL(P,Q) ≠ KL(Q,P) holds.

Jensen-Shannon Divergence (JS-Div) is a measure of the size of the difference between two distributions. The JS-Div calculation requires KL-Div to obtain a symmetric normalized value so that the value JS(P,Q) = JS(Q,P). The JS-Div value can be calculated in the following way [32]:

$$JS(P, Q) = \frac{1}{2} \left(KL \left(P, \frac{1}{2} (P + Q) \right) + KL \left(Q, \frac{1}{2} (P + Q) \right) \right). \quad (7)$$

2.4. Ridge Regression

Modeling in this research utilizing Ridge regression which is the improvement method from commonly used Ordinary Least Square (OLS). The improvement was developed caused by few assumption that restrict the uses of OLS, which are: uncorrelated error, have normal distribution, and not having multicollinearity among each independent variables. For the regular expression of regression of $X\beta + \epsilon$, Regression parameter of OLS [33] could be calculated using equation below:

$$\beta^{OLS} = (X^T X)^{-1} \cdot X^T Y. \quad (8)$$

Ridge regression calculation is done by adding c bias constant at diagonal matrix of $X^T X$ [34]. This method was perform to eliminate the ill-condition caused by multicollinearity of each dependent variables. During ill-condition, $X^T X$ matrix have form close to singular matrix so that result in unstable estimated parameters [35].

In the ridge regression process, each variable is standardized before used in the estimation process. The estimated value of the ridge parameter could be obtained using penalized least square method [36], following equation below:

$$PLS = \sum_{i=1}^{n} \left(V_i - \hat{\beta}_0 - \hat{\beta}_1 U_{i1} - \cdots - \hat{\beta}_k U_{ik}\right)^2 + c \sum_{j=0}^{k} \left(\hat{\beta}_j\right)^2. \quad (9)$$

The c value in the equation is called ridge estimator with value of c≥0 [37]. Value of c could be calculated using equation below [38]:

$$c = \frac{k \cdot \sigma^2}{\beta^{OLS^T} \beta^{OLS}} \quad (10)$$

with k is number of independent variables, $\sigma^2 = \frac{\left(Y-\hat{Y}\right)^T \left(Y-\hat{Y}\right)}{n-k-1}$, while β^{OLS} is regression parameter of Ordinary Least Square. Ridge regression parameter obtained by minimizing Eq. x, thus obtained equation:

$$\beta^R = \left(U^T U + cI\right)^{-1} \cdot U^T V \quad (11)$$

with I is identity matrix that have order of k.

when c = 0, the ridge parameter is same with OLS parameter. When c > 0, ridge parameter become bias while maintain stability compared to regular regression [39]. To compare the modeling result, this research uses determination coefficient. The determination value could be calculated using equation below [37]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2 + \sum_{i=1}^{n} \left(\hat{y}_i - \bar{y}\right)^2} \quad (12)$$

with $\hat{y}$ is the estimated regression value and $\bar{y}$ is the average value of y.

3. Results

3.1. Analysis of Characteristics of Original Data and Data Related to Forest and Land Fires

Data characteristics analysis aims to determine the characteristics of the data being used. The characteristics of the data obtained include descriptive statistics (mean, median, standard deviation) and data distribution. Based on **Table 1**, the mean, median and standard deviation values of the *act* variable are closer to the *et0* variable than the *pev* variable for both types of data.

Distribution analysis is an analysis carried out to see the differences in the distribution of two or more data. The distribution of the three original data and data related to forest and land fires is expressed using a kernel density function graph. The graph was created using the "ksdensity" function in MATLAB software.

Based on the distribution plot in **Figure 2**, the *et0* and *act* variables in the original data have a more similar distribution compared to the *pev* variable. However, based on data related to forest and land fires, the distribution of the three variables has a similar pattern. This must be further confirmed through several approaches, including QQ plot and Jensen-Shannon Divergence. QQ plot is carried out to identify the similarity of two distributions. If the QQ plot is linear, the two distributions are considered the same.

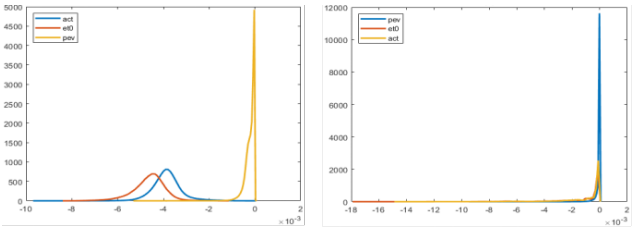

Figure 2. Kernel Probability Density Function Graph from original data (**left**) and data related to forest and land fires (**right**).

Based on **Figures 3** and **4**, the variables *et0* and *act*

Table 1. Descriptive statistics for three variables.

Data	Variable	Mean	Median	Standard Deviation
Original Data	*act*	−0.0038	−0.0039	$0.6408 \cdot 10^{-3}$
	pev	−0.0002	−0.0002	$0.2329 \cdot 10^{-3}$
	et0	−0.0046	−0.0046	$0.7134 \cdot 10^{-3}$
Data Related to Forest and Land Fires	*act*	−0.0011	$-0.2424 \cdot 10^{-3}$	0.0018
	pev	−0.0001	$-0.0149 \cdot 10^{-3}$	0.0002
	et0	−0.0013	$-0.2923 \cdot 10^{-3}$	0.0021

have the most similar distribution (more linear) in the original data and data related to forest and land fires. In contrast to visual observations, the results of the QQ plot show that the distribution of variables related to forest and land fires for the *pev* variable is different from the other two variables. The magnitude of the difference between the three variable distributions is obtained through the Jensen-Shannon Divergence (JSdiv) value. In general, this method measures the size of the difference between two probability distribution curves. The JSdiv calculation results for the three variables for each type of data are presented in **Tables 2** and **3**.

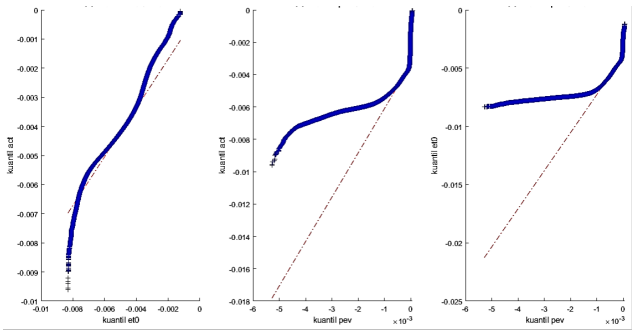

Figure 3. QQ plot between variables on the original data (**left**) et0 and act; (**middle**) pev and act; (**right**) pev and et0.

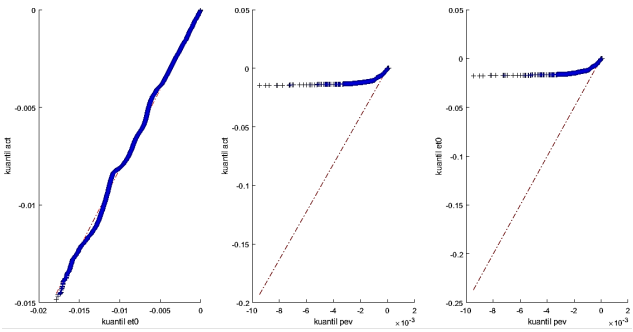

Figure 4. QQ plot between variables in data related to forest and land fires (**left**) *et0* and act; (**middle**) *pev* and act; (**right**) *pev* and et0.

Based on **Tables 2** and **3**, the large differences in distribution shown by the JS-Div values for the original data and data related to forest and land fires both show that the *act* distribution is more similar to the *et0* distribution compared to the *pev* distribution. This is in accordance with the QQ plot results in **Figures 3** and **4**. From the three approaches carried out on both types of data, it was found that the distribution of the *pev* variable was different from the distribution of the other two variables.

Table 2. JS-Div values original data distribution.

Variable	*pev*	*et0*	*act*
pev	0	0.6726	0.6621
et0		0	0.1323
act			0

Table 3. JS-Div value distribution of data related to forest and land fires.

Variable	*pev*	*et0*	*act*
pev	0	0.6726	0.6621
et0		0	0.1323
act			0

3.2. Spatial and Temporal Analysis of Original Data and Forest and Forestry Related Data

This analysis is carried out by comparing the spatial patterns of the two types of data. In this analysis, the data is converted into a cumulative distribution function (CDF) for each data element. This process is carried out using the "ecdf" function in MATLAB software. The three CDF data vectors have a similar data range, namely from 0 to 1. Then this range is divided into three intervals of the same length, as described in **Table 4**. Based on the criteria in **Table 4**, the location points of the *pev* and *et0* variables which have similar criteria to the *act* variable are displayed on the spatial graph and the number of location points is counted. The spa-

tial graph and the number of point locations can be seen in **Figure 5**. The black dot in **Figure 5** shows the location of the *act* variable whose criteria are the same as the *et0* variable or variables. The number of black dots in **Figure 5** shows that the number of location points that have the same criteria between the *act* variable and the *et0* variable is greater than the *pev* variable. Relative to the total number of location points, the percentage of similarity of the spatial pattern of the *et0* and *pev* variables with the *act* variable is presented in **Table 5**.

Table 4. CDF data criteria.

Interval cdf (x)	Interpretation
$0 \leq x \leq 0.33$	Low
$0.33 < x < 0.6$	Medium
$0.66 \leq x \leq 1$	High

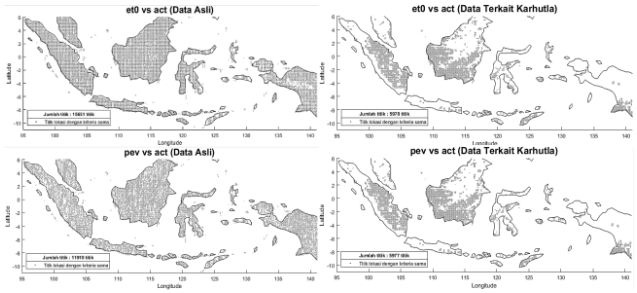

Figure 5. Location points with the same criteria for the *pev* variable and *et0* variable as the *act* variable in the original data and data related to forest and land fires.

Based on **Table 5**, the *et0* variable produces more location points, namely 86.58% of all spatial points in the original data. This percentage is 20.69% higher than the location points produced by the *pev* variable. However, in data related to forest and land fires, the percentage of location points produced by the *et0* variable and the *pev* variable are not much different, namely 100% and 99.98% respectively. These results indicate that in data related to forest and land fires, the spatial criteria for the three variables are similar.

The spatial analysis carried out above shows that the process of taking factors related to forest and land fires has an impact on the similarity of the spatial criteria for the elements of the three variables. The use of the CDF approach to the data and the division of criteria carried out above can only indicate uniformity of values in the interval for each variable. Therefore, other analyzes need to be carried out to see the behavior of the data from a different perspective.

Note that the evaporation rate is also influenced by temperature, solar radiation and rainfall, while the observed areas in Indonesia have varying temperatures, radiation and rainfall. Therefore, a temporal analysis using a data mode approach to see the evaporation level of most of all regions in Indonesia for each month was carried out.

There are differences between the temporal graph of the original data and data related to forest and land fires (**Figure 6**). In the original data, the patterns of the three variables do not show similarities. This is different from the temporal graph of data related to forest and land fires. The three variables appear to have similar temporal patterns. The differences in the temporal patterns of the three variables above are carried out using several approaches, namely difference graphs and using metric distance calculations, namely Minkowski Distance.

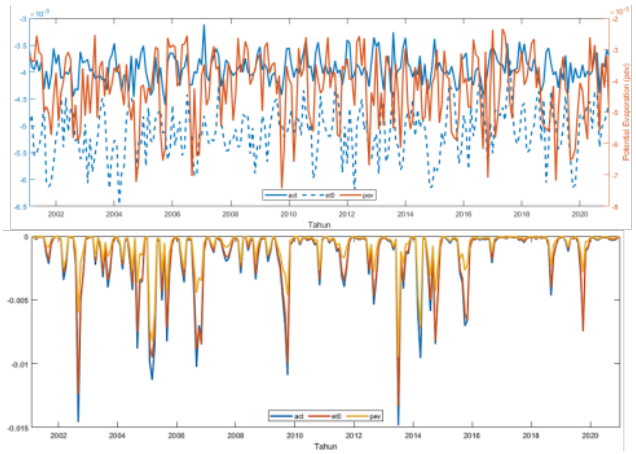

Figure 6. Monthly temporal (timeseries) graph of the combined mode of the three variables for original data (**top**) and data related to forest and land fires (**bottom**).

Based on **Table 6**, the value of the difference between the *et0* variable and the *act* variable is closer to zero than the difference between the *pev* variable and the *act* variable. This shows that for both types of data, the *et0* variable and the *act* variable have higher temporal graph similarity. However, in data related to forest and land fires, there is a change in the distance between variables in the original data and data related to forest and land fires. The distance between variables in data related to forest and land fires is smaller than in the original data.

The temporal analysis carried out above shows that the process of taking factors related to forest and land fires has an impact on increasing the similarity of the temporal

Table 5. Number of points and percentage of locations for the *et0* variable and the *pev* variable which have the same criteria as the *act* variable.

Variable	Original Data			Data Related to Fires Event		
	Black Dot	Total of Dot	Percentage	Black Dot	Total of Dot	Percentage
et0	15651	18076	86.58%	5978	5978	100%
pev	11910		65.89%	5977		99.98%

Table 6. Minkowski distance (p = 5) between variables from both types of data.

	Original Data			Data Related to Fires Event		
	pev	***et0***	***act***	***pev***	***et0***	***act***
pev	0	0.0155	0.0118	0	0.0078	0.0096
et0		0	0.0046		0	0.0028
act			0			0

patterns of the three variables. The improvements that occur include smaller distances between temporal graphs and similar temporal patterns. Although, the temporal graph of the *et0* variable is still closer to the *act* variable than the *pev* variable.

3.3. Regression Analysis

Ridge regression analysis was carried out to compare the performance of the *pev* and *et0* variables in modeling forest fires. The forest fire indicator used is hotspots. Similar to the analysis carried out previously, ridge regression analysis was carried out on both types of data. This analysis also uses two types of observation locations, namely the location of Indonesia as a whole and the location of the cutting results. The locations of the cuts are shown in **Figure 7**. The four locations in **Figure 7** were selected based on the high density of hotspot points (red dots) from the filtered locations in **Figure 6**. The location of Indonesia is hereinafter referred to as location 5.

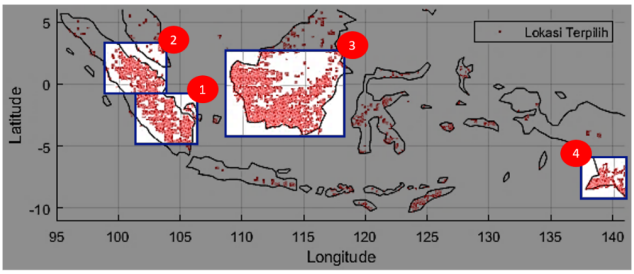

Figure 7. 4 focused locations.

Regression estimation is carried out with the help of the "ridge" function in MATLAB. Comparison of the performance of each variable in modeling forest fires was carried out by comparing the coefficient of determination of each regression. The coefficient of determination values obtained are presented in **Table 7**. The table shows the performance of the variables act, pev, and et0 from each location to model each type of data, denoted by act_{ij}, $et0_{ij}$, $pev_{ij}$with index i = 1,2 respectively being the data index original and data related to forest and land fires and index j = 1,2,3,4,5 is the index for the five locations.

Table 7. Coefficient of determination value (R^2).

Variabel	R^2	Variable	R^2
act_{11}	0.8738	act_{21}	0.8929
$et0_{11}$	0.8697	$et0_{21}$	0.8929
pev_{11}	0.8750	$et0_{21}$	0.8930
act_{12}	0.7548	act_{22}	0.7798
$et0_{12}$	0.7343	$et0_{22}$	0.7804
pev_{12}	0.7523	pev_{22}	0.7796
act_{13}	0.8358	act_{23}	0.8176
$et0_{13}$	0.8426	$et0_{23}$	0.8158
pev_{13}	0.8418	pev_{23}	0.8186
act_{14}	0.8793	act_{24}	0.9465
$et0_{14}$	0.3928	$et0_{24}$	0.9464
pev_{14}	0.8794	pev_{24}	0.9465
act_{155}	0.6993	act_{25}	0.7628
$et0_{15}$	0.7061	$et0_{25}$	0.7643
pev_{15}	0.6979	pev_{25}	0.7643

Based on **Table 7**, the coefficient of determination obtained by the majority is above 0.7 (or 70%). This shows that the three variables in each type of data in most locations can model forest fires well. **Table 7** shows that the *pev* variable performs better than the *et0* variable in six out of ten cases.

The performance of the *et0* variable is above the *pev* variable at location 3 for the original data, location 2 for data related to forest and land fires and total location for the original data. On the other hand, the performance of these two variables is the same for total locations for data related to forest and land fires. Performance of the *act* variable is more similar to the *pev* and *et0* variables in four cases and five cases. This means that in general the *et0* variable has performance that is slightly more similar to the *act* variable. However, the *pev* variable also shows similar performance to the *act* variable in several cases.

Next, the effect of using the SVD-based EOF method is analyzed. Performance comparison analysis was carried out on original data and data related to forest and land fires. A comparison of the performance of the three variables on different types of data is presented in **Figure 8**. The three variables experienced improved performance on data related to forest and land fires compared to the original data in all cases except Location 3 (**Figure 8**). The performance of the three variables in the cut Location data is better than the Location 5 data.

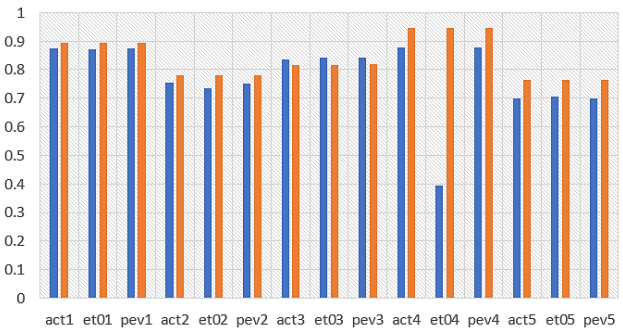

Figure 8. Comparison graph of the coefficient of determination values for the three variables in the two types of data, original (blue) and related to fires event (orange).

3.4. Hotspot Estimation Analysis

Ridge regression analysis shows that overall the *pev* variable has better performance than the *et0* variable in modeling hotspot occurence. Based on the similarity of performance with the *act* variable, the *et0* variable slightly outperforms the *pev* variable. In addition, the performance of variables in the original data is worse than the performance of variables in data related to forest and land fires. Furthermore, the cut areas have higher performance than the Indonesian region. The last section of analysis is to investigate robustness of each variable toward all Indonesian regions in single relation. Thus, the analysis is conducted only toward filtered data that related to forest fire event obtained in the previous SVD analysis. Investigation is performed toward 2 analysis, which is joint distribution and ridge regression analysis by separating several years as testing data (20%) and rest as training data (80%) to build the ridge regression model. When the data is predicted, the data is removed from the training. For example, the 2001–2004 data is predict using ridge regression model that built using 2005–2020 data. This approached is used to avoid self-predicted value from the training data. The error value is shown to get an insight of how the result spread toward the modelling.

Figure 9 shows the possibility of severe dry condition (represented by each variable) occurred together with hotspot occurrence. The blue area represented low probability while yellow area represents high probability. As expected, *act* and *et0* gives similar result due to similar character of distribution obtained in the previous analysis. While both variables could give high correlation in low-left area (low number of hotspot during wet condition) and high-right area (high number of hotspot during severe dry condition), the difference of both variables is insignificant and barely noticeable.

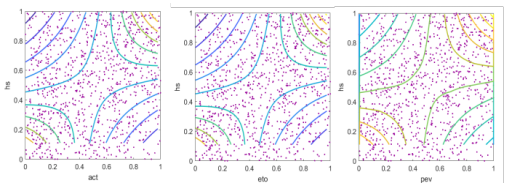

Figure 9. Comparison graph of the joint distribution between each variable against hotspots (*hs*).

Using symmetric copula family, both variables is best fitted using frank copula with Log Likelihood value of 7.9357e+03 and 9.1241e+03 for *act* and *et0* respectively. However, best fitted of et0 is obtained when using frankhac copula which is asymmetric modification of frank copula. The frankhac modifier gives et0 slightly higher correlation on high-right area than low-left area, indicating potential higher performance when estimating high fire event. Meanwhile, *pev* gives significant different by increasing the probability of area around the 0.4 – 0.6 cdf which result of higher correlation of hotspot occurrence during normal drought condition.

This result is very essential in term of modelling to not only anticipate the extreme scale of fires event (such as happened in 2015, 2019), but also medium scale fires event

that occurrent more often. Result in **Figure 9** shows the key difference between evaporation (*pev*) and evapotranspiration (*act* and *et0*) is on the diagonal of the distribution which has higher correlation across the positive diagonal (**Figure 9**). This result support analysis in **Table 7** where even though *pev* didn't have higher R^2 value, *pev* is more consistent performance across 5 focused regions. However, this result does not mean that *pev* is the most accurate variables in terms of estimating hotspot occurrence.

Figure 10 shows the performance of each variable when estimating total monthly hotspot of 4 consecutive years. In this analysis, all variables could perform similarly in reasonable performance except when predict 2005–2008 data and 2017–2020 data. Only *et0* could give good result when predict 2017–2020 fires event where other variables very overestimated the condition resulting low amount of accuracy. The result is supported by performance matrix in **Table 8** which highlight the important of *et0* in the modeling of fires event, especially in extreme condition. While *pev* and *act* gives relatively similar value of R^2 for each testing year with less than 0.05 R^2 value of difference. The largest gap is obtained when estimating 2017–2020, while the smallest gap is presence in estimating 2013–2016 data. Meanwhile, et0 outperform both variables by consistently giving higher R^2 value and smaller MSE. Moreover, et0 gives best result when estimating 2017–2020 while pev and act gives the worst result.

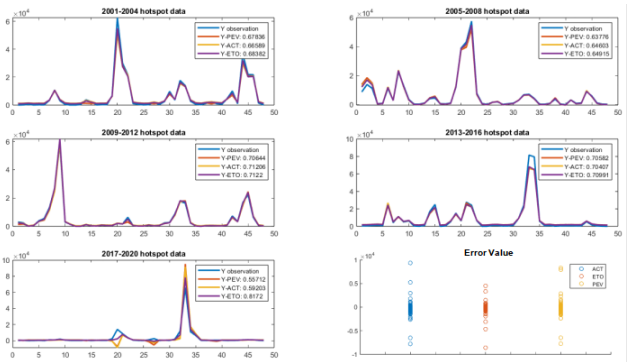

Figure 10. Ridge Regression performance comparison between each data, modelling each respective year as well as error value comparison.

Overall, *et0* perform best in the hotspot estimation analysis supported with low spread of error value indicating lower uncertainty of events when estimating using *et0* variables. Performance between *pev* and *act* is very similar in accuracy result in almost identical estimation. However, the spread of error value is slightly lower in *pev*, indicating that it has lower uncertainty.

4. Discussion

Evapotranspiration and evaporation are two of the common indicators, determining drought condition with precipitation data. Evapotranspiration (ET) plays a crucial role in connecting ecosystem functioning, carbon and climate feedback, agricultural management, and water resources. It also emphasizes important scientific and practical questions, along with necessary actions particularly from a space-based viewpoint to further our understanding [40]. However, it is difficult to get observation data ET over a large scale and for a long time. Thus, many models were developed over the past few decades to estimate the actual value and their usability in scientific research. Multiple ET datasets exist, each with associated uncertainties stemming from diverse assumptions related to their algorithms, parameters, and input data. Furthermore, some researcher also trying to synthesis those many different datasets to obtained better estimation [41, 42]. Meanwhile, Pan evaporation, observed globally as a routine practice, serves as an excellent indicator of how terrestrial evaporation responds to global warming [43]. Because of different measurement method and physical indicators that influenced it, pan evaporation, as well as reference and actual evapotranspiration have different characteristic and trend [44]. Thus, usability for each of them is quite different in the research and application.

Analysis in this research provide multiple characteristics of the 3 variables and their performance when representing drought condition, especially in related to forest fire event (indicated with hotspot occurrence). As expected, the value of reference evapotranspiration (calculated using Penman-Monted procedure) significantly similar to the actual evapotranspiration (calculated using ERA5-LAND Guideline). Unfortunately, the similarity is decreases when analyzed only in condition related to hotspot occurrence showing the complexity of the fires event in Indonesia.

While evapotranspiration itself is used in most of the popular fire danger index, evapotranspiration itself has complex relation especially in respond of climate change [45]. Most of the climate models show increasing reference evapotranspiration because of increased air temperature [46]. Vari-

Table 8. Coefficient of DETERMINATION (R^2) and Mean Square Error (MSE) of the Regression.

Testing Year	R^2 *pev*	MSE *pev*	R^2 *act*	MSE *act*	R^2 *et0*	MSE *et0*
2001–2004	0.6784	$1.1063e+07$	0.6659	$1.1492e+07$	0.6838	$1.0875e+07$
2005–2008	0.6378	$1.2026e+07$	0.6460	$1.1752e+07$	0.6491	$1.1648e+07$
2009–2012	0.7054	$7.1683e+06$	0.7121	$7.0311e+06$	0.7122	$7.0278e+06$
2013–2016	0.7058	$2.1186e+07$	0.7041	$2.1312e+07$	0.7099	$2.0891e+07$
2017–2020	0.5591	$1.0429e+07$	0.5920	$9.6064e+06$	0.8172	$4.3043e+06$

ations in evapotranspiration impact water availability and ecosystem health. During droughts, increased evaporative demand boosts evapotranspiration, but the reduced moisture supply limits it, making it difficult to predict even the direction of evapotranspiration anomalies[47]. Popular fire weather index such as Canadian Fire Weather Index[48], McArthur Fores Fire Danger Index[49], and Keetch-Byram drought index[50] are product of several key drought indicators, which is relative humidity, temperature, wind speed, and precipitation. Except precipitation, those indicators are calculated in the reference evapotranspiration (calculated using Penman-Monted equation). This is key reason that allows reference evapotranspiration has consistently higher performance in the hotspot prediction analysis[51].

In the ERA5-LAND Guideline, the open water evaporation also called pan evaporation which often considered as potential evaporation substitute indicators[52, 53]. However, evaporation from open water could not be used to represent reference or actual evapotranspiration (shown by the analysis). This result is in line with previous research investigating various evapotranspiration method over mainland China data[54], even though both pan evaporation and potential evapotranspiration in the analysed area represent the combined effects of radiation, wind, temperature, and humidity on evaporation, and display strong correlations between them. The performance of it is similar with total evapotranspiration from ERA5 Land (considered as actual evapotranspiration).

Supported with consistent high result in region analysis, open water evaporation (pan evaporation) could be good as alternative instead of using reference evapotranspiration to predict already known phenomena, especially for spatial analysis in local region with similar character of drought-fires event. The bad performance in the analysis is caused by high number of hotspot occurrence happened in 2006[55] and 2019[56] caused by extreme positive IOD which rarely affect Indonesian forest fire. In this case, potential evapotranspiration has relatively higher performance in terms of estimation shown by **Table 8** even though the performance still differs depend on the region (**Figure 8**). However, this result is mainly affected by Ridge regression behavior in the estimation. Ridge regression did not drop independent variable that has low coefficient value to minimize the number of features in the regression[57]. While it could become heavier to execute and harder to interpret the result, it allows the models has contribution surrounding area/lower correlation point in the occurrence of fire event. Thus, further analysis with multiple different regression method with various behavior needs to be done to get other insight in the performance with different assumption.

Main highlight of the open water evaporation is the ability to have higher correlation with hotspot occurrence in Indonesia, during the medium scale of fires event in Indonesia which shown by the joint distribution analysis. Pan evaporation gives more consistent performance when used to analyses fire event in regions that has multiple characteristics of forest fire. This behavior is very important since there are multiple characteristic of fires event and need to be handled separately to obtained optimal result of analysis[58–60]. Moreover, it could give lower uncertainty during the error value analysis compared to actual evapotranspiration (from ERA5-Land). While the performance is comparable to actual evapotranspiration, it has lower ability of estimating extreme fires event that never happened. In this regard, reference evapotranspiration is more useful especially since it has higher correlation during high probability of fires event and drought condition. However, analysis in this research is very surface level experiment that needs to be explore further to get better understanding of the indicator characteristics. Additionally, considering there are different trend in each type of evapotranspiration[61–63], in depth analysis regarding uncertainty of different type evapotranspiration when predict/project future fires event needs to be explored to anticipate more severe drought condition and higher fires

risk in the future[64, 65]. Moreover, further improvement of the research could be done by analyzing how each variable behaves in different drought/fire danger index such as standardized precipitation evapotranspiration index, fire weather index, and fire danger index, as well as integrating the impact of bigger climate phenomena such as El Nino Southern Oscillation and Indian Ocean Dipole.

5. Conclusions

Theoretically, the evaporation from open water variable has fewer calculation indicators than the actual (calculated following ERA5 guideline) and reference evapotranspiration (calculated using Penman-Monted model). Descriptive statistics for the open water evaporation are different from both actual and the reference evapotranspiration. The distribution of the actual is more similar to the reference variable, especially in original data. However, all three variables show similar distribution when filtered only for data related to fires event. Thus, the three variables show similar temporal patterns in data related to forest and land fires. The distance between temporal actual data is closer to the reference evapotranspiration.

Based on the cumulative distribution function (CDF) value of each indicator, the spatial patterns of the three indicators have a high level of similarity to data related to forest and land fires. The use of the SVD-based EOF method increases the similarity of the temporal and spatial patterns of the three variables and effectively improves the performance of the three variables in modeling forest fires. Cutting the total Location to four Locations had a good impact in improving the performance of the three variables.

The performance of the open water evaporation in modeling forest fires is better than the reference evapotranspiration variable in regions with multiple characteristics of fires event. This result indicates the potential uses of open water evaporation when modeling general characteristic of the fires event that can used robustly across different region. However, in modeling extreme event, reference evapotranspiration is superior compared to other analyzed variables, especially in event that never happened in the training data. This result indicates higher ability of reference evapotranspiration to uses in analyzing and modeling future impact of climate changes that potentially increases fires risk in Indonesia.

Author Contributions

E.H.N.: Writing—review & editing, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. S.N.: Visualization, Writing—review & editing, Validation, Supervision, Resources, Project administration, Methodology, Conceptualization. A.S.: Supervision, Resources, Data curation, Conceptualization. P.S. Writing—review & editing, Software, Formal analysis. W.P.: Writing—original draft, Software.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The datasets used during current study are available from corresponding author on reasonable request.

Acknowledgments

The authors would like to thank The Department of Mathematics, IPB University and The Meteorological, Climatological and Geophysical Agency for their strong support and invaluable assistance throughout this research.

Conflict of Interest

There is no conflict Interest in this research.

References

[1] Gellert, P.K. 1998. A brief history and analysis of Indonesia's forest fire crisis. Southeast Asia Program Publications at Cornell University. 65:63-85.

[2] Cahyono, S.A., Warsito, S.P., Andayani, W., Darwanto, D.H. 2015. Factors Influencing Forest Fires in Indonesia and Their Policy Implications (Translated: Indonesia). Jurnal Sylva Lestari. 3(1):103-112.

[3] Suharjo, B.H., Velicia, W.A. 2018. The Role of Rainfall Towards Forest and Land Fires Hotspot Reduction in Four Districs in Indonesia on 2015-2016. Jurnal Silvikultur Tropika. 09(1).

[4] William, A.P., et al. 2014. Correlations between components of the water balance and burned area reveal new insights for predicting forest fire area in the southwest United States. International Journal of Wildland Fire. 24(1):14-26. https://doi.org/10.1071/WF14023

[5] Hadiyanto, S. 2007. Pattern of drought vulnerability levels in Central Java (Translated: Indonesia) [Graduate Thesis]. Jakarta (ID): University of Indonesia

[6] Tsakiris, G., Vangelis, H. 2005. Establishing a drought index incorporating evapotranspiration. European Water. 9(10):3-11.

[7] Vembrianto, N., Yoza, D., Sribudiani, E. 2015. Ecological characteristics Location of forest and land fires in Rantau Bais Village, Tanah Putih District, Rokan Hilir Regency (Translated: Indonesia). Jom Faperta. 2(1):1-9.

[8] Marganingrum, D., Santoso, H. 2019. Evapotranspiration of Indonesian tropical area. Jurnal Presipitasi. 16(3): 106.

[9] Yu, G., Feng, Y., Wang, J., Wright, D. B. 2023. Performance of fire danger indices and their utility in predicting future wildfire danger over the conterminous United States. Earth's Future. 11(e2023EF003823). https://doi.org/10.1029/2023EF003823

[10] Quintano, C., Fernández-Manso, A., Fernández-Guisuraga, J.M., Roberts, D.A. 2024. Improving Fire Severity Analysis in Mediterranean Environments: A Comparative Study of eeMETRIC and SSEBop Landsat-Based Evapotranspiration Models. Remote Sens. 16 (361). https://doi.org/10.3390/rs16020361

[11] Ma, Q., Bales, R.C., Rungee, J., Conklin, M. H., Collins, B.M., Goulden, M.L. 2020. Wildfire controls on evapotranspiration in California's Sierra Nevada. Journal of Hydrology. 590. https://doi.org/10.1016/j.jhydrol.2020.125364

[12] Nolan, R.H., Lane, P.N.J., Benyon, R.G., Bradstock, R.A., Mitchell, P.J. 2019. Changes in evapotranspiration following wildfire in resprouting eucalypt forests. Ecohydrology. 7:1363-1377.

[13] Dimitriadou, S., Nikolakopoulos, K.G. 2021. Evapotranspiration Trends and Interactions in Light of the Anthropogenic Footprint and the Climate Crisis: A Review. Hydrology. 8 (163). https://doi.org/10.3390/hydrology8040163

[14] Lhomme JP. 1997. Towards a rational definition of potential evaporation. Hydrol Earth Syst Sci. 1(2): 257-264.

[15] Han, S., Tian, F., Hu, H. 2014. Positive or negative correlation between actual and potential evaporation? Evaluating using a nonlinear complementary relationship model. Water Resources Research. 50(2): 1322-1336.

[16] Ghiat, I., Mackey HR, Al-Ansari T. 2021. A review of evapotranspiration measurement models, techniques and methods for open and closed agricultural field application. Water.13(18):2523.

[17] Nurdiati, S., Sopahelukawan, A., Septiawan, P., Ardhana, M.R. 2022. Joint spatio-temporal analysis of various wildfire and drought indicators in Indonesia. Atmosphere. 13(1591). https://doi.org/10.3390/atmos13101591.

[18] Abtew W, Melesse A. 2013. Evaporation and Evapotranspiration: Measurements and Estimations. Berlin (DE): Springer.

[19] Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. Geographical Review 38. Nr. 1: 55-94.

[20] Granger RJ. 1989. An examination of the concept of potential evaporation. Journal of Hydrology. 111(1-4): 9-19. https://doi.org/10.1016/0022-1694(89)90248-5

[21] Muñoz, S. J. 2019. ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac

[22] ORNL DAAC. 2018. MODIS and VIIRS Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1379

[23] Penman, H.L. 1948. Natural Evaporation from Open Water, Bare Soil And Grass. Proc. R. Soc. London, Ser. A. 193: 120–146

[24] Monteith, J. L. 1965. Evaporation and Environment. 19th Symposium of the Society for Experimental Biology: 205-234. Cambridge Univ. Press, Cambridge

[25] Hannachi, A. 2004. A primer for EOF analysis of climate data: Department of Meteorology, University of Reading, Reading RG6 6BB, UK

[26] Navarra, A., Simoncini, V. 2010. A Guide to Empirical Orthogonal Function for Climate Data Analysis: Springer

[27] Cassisi, C., Montalto, P., Aliotta, M., Cannata, A., Pulvirenti, A. 2012. Advances in Knowledge Discovery and Data Mining. Switzerland (CH): Springer

[28] Chen, S.X. 2000. Probability density function estimation using gamma kernels. Annals of the Institute of Statistical Mathematics. 52(3): 471–480. https://doi.org/10.1023/a:1004165218295

[29] Marden, J.I. 2004. Positions and QQ plots. Statistical Science. 19(4): 606-614.

[30] Joyce, J.M. 2011. International Encyclopedia of Statistical Science. New York (US) : Springer, p720–722. https://doi.org/10.1007/978-3-642-04898-2_327

[31] Nguyen, H.V., Vreeken, J. 2015. Non-parametric Jensen-Shannon Divergence. Lecture Notes in Computer Science: 173–189. https://doi.org/10.1007/978-3-319-23525-7_11

[32] Lin, J. 1991. Divergence measures based on the Shannon entropy. IEEE Trans. Inf. Theory, 37:145–151.

[33] Acito, F. 2023. Ordinary Least Squares Regression. In: Predictive Analytics with KNIME. Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-45630-5_6

[34] Hilt, D. E., Seegrist, D.W. 1977. Ridge, a computer program for calculating ridge regression estimates. https://doi.org/10.5962/bhl.title.68934

[35] Hoerl, A. E., Kennard, R. W. 1970. Ridge Regression: Applications to Nonorthogonal Problems. Technometrics. 12(1): 69–82. https://doi.org/10.1080/00401706.1970.10488635

[36] Tutz, G., Ulbricht, J. 2006. Penalized regression with correlation based penalty. Statistic and Computing. 19(3): 239-253.

[37] Paulson, D.S. 2007. Handbook of Regression and Modeling. New York (US): Chapman & Hall/CRC.

[38] El-Dereny, M., Rashwan, N.I. 2011. Solving multicollinearity problem using ridge regression models. Int. J. Comtemp. Math. Sciences. 6(12): 585-600.

[39] Kutner, M., Christhoper, Nachtsheim, J.C., Neter, J., Li, W. 2005. In Applied Linear Statistical Models, Fifth Edition. NewYork (US): McGraw-Hill.

[40] Fisher, J. B., et al. (2017), The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, Water Resour. Res. 53: 2618–2626, https://doi.org/10.1002/2016WR020175

[41] Zhang, K., Chen, H., Ma, N. et al. 2024. A global dataset of terrestrial evapotranspiration and soil moisture dynamics from 1982 to 2020. Sci Data. 11(445). https://doi.org/10.1038/s41597-024-03271-7

[42] Elnashar, A., Wang, L., Wu, B., Zhu, W., Zeng, H. 2021. Synthesis of global actual evapotranspiration from 1982 to 2019. Earth Syst. Sci. Data. 13: 447–480. https://doi.org/10.5194/essd-13-447-2021

[43] Du, J.Z., Xu, X.L., Liu, H.X., Wang, L.Y., Cui, B.S. 2023. Deriving a high-quality daily dataset of large-pan evaporation over China using a hybrid model. Water Res. 238, 120005.

[44] Liu, Y. J., Chen, J., Pan, T. 2019. Analysis of Changes in Reference Evapotranspiration, Pan Evaporation, and Actual Evapotranspiration and Their Influencing Factors in the North China Plain During 1998–2005. Earth and Space Science. 6: 1366–1377. https://doi.org/10.1029/2019EA000626

[45] Fisher, J. B., et al. 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resour. Res. 53 :2618–2626. https://doi.org/10.1002/2016WR020175

[46] McEvoy, D. J., Pierce, D. W., Kalansky, J. F., Cayan, D. R., Abatzoglou, J. T. 2020. Projected changes in reference evapotranspiration in California and Nevada: Implications for drought and wildland fire danger. Earth's Future. 8(e2020EF001736). https://doi.org/10.1029/2020EF001736

[47] Zhao, M., A, G., Liu, Y. et al. 2022. Evapotranspiration frequently increases during droughts. Nat. Clim. Chang. 12:1024–1030. https://doi.org/10.1038/s41558-022-01505-3

[48] Van Wagner, C. E. 1987. Development and Structure of the Canadian Forest Fire Weather Index System, Technical Report 35, Canadian Forestry Service, Ottawa, ON.

[49] McArthur, A. G. 1967. Fire Behaviour in Eucalypt Forests. Department of National Development Forestry and Timber Bureau, Canberra, Leaflet 107

[50] Ketch, J.J., Byram, G.M. 1968. A drought index for forest fire control. Res. Pap. SE-38. Asheville, NC. U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, 32 p.

[51] Yu, X., Qian, L., Wang, W., Huo, X.; Hu, X., Wang, Y. 2023. Assessing and Comparing Reference Evapotranspiration across Different Climatic Regions of China Using Reanalysis Products. Water. 15(2027). https://doi.org/10.3390/w15112027

[52] Liu, B., 2004. A spatial analysis of pan evaporation trends in China, 1955–2000. J. Geophys. Res. 109. D15102. https://doi.org/10.1029/2004JD004511.

[53] Xie, R., Wang, A., 2020. Comparison of ten potential evapotranspiration models and their attribution analyses for ten chinese drainage basins. Adv. Atmos. Sci. 37: 959–974. https://doi.org/10.1007/s00376-020-2105-0.

[54] Xu, C., Wang, W., Hu, Y., Liu, Y. 2024. Evaluation of ERA5, ERA5-Land, GLDAS-2.1, and GLEAM potential evapotranspiration data over mainland China. Journal of Hydrology: Regional Studies. 51(101651). https://doi.org/10.1016/j.ejrh.2023.101651

[55] Pan, X., Chin, M., Ichoku, C. M., Field, R. D. 2018. Connecting Indonesian fires and drought with the type of El Niño and phase of the Indian Ocean dipole during 1979–2016. Journal of Geophysical Research: Atmospheres. 123: 7974–7988. https://doi.org/10.1029/2018JD028402

[56] Dafri, M., Nurdiati, S., Sopaheluwakan, A. 2021. Quantifying ENSO and IOD impact to hotspot in Indonesia based on Heterogeneous Correlation Map (HCM). Journal of Physics: Conference Series. 1869. 012150. https://doi.org/10.1088/1742-6596/1869/1/012150

[57] Zhang, Y., Politis, D.N. 2021. Ridge Regression Revisited: Debiasing, Thresholding and Bootstrap. Mathematics Statistics Theory. https://doi.org/10.48550/arXiv.2009.08071

[58] Nurdiati, S., Sopaheluwakan, A., Julianto, M.T., Septiawan, P., Rohimahastuti, F. 2021. Modelling and anal-

ysis impact of El Nino and IOD to land and forest fire using polynomial and generalized logistic function: Cases study in South Sumatra and Kalimantan, Indonesia. Model. Earth Syst. Environ. 8:3341–3356.

[59] Ardiyani, E., Nurdiati, S., Sopaheluwakan, A., Septiawan, P., Najib, M.K. 2023. Probabilistic Hotspot Prediction Model Based on Bayesian Inference Using Precipitation, Relative Dry Spells, ENSO and IOD. Atmosphere. 14 (286). https://doi.org/10.3390/atmos14020286

[60] Nurdiati, S., Bukhari, F., Sopaheluwakan, A., Septiawan, P., Hutapea, V. 2023. ENSO and IOD Impact Analysis Of Extreme Climate Condition In Papua, Indonesia. Geographia Technica. 19:1-18. https://doi.org/10.21163/GT_2024.191.01

[61] Zhang, Y., Peña-Arancibia, J., McVicar, T. et al. 2016. Multi-decadal trends in global terrestrial evapotranspiration and its components. Sci Rep. 6(1924) https://doi.org/10.1038/srep19124

[62] Li, Z., Wang, S. Li, J. 2020. Spatial variations and long-term trends of potential evaporation in Canada. Sci Rep. 10(22089). https://doi.org/10.1038/s41598-020-78994-9

[63] Shen, J., Yang, H., Li, S., Liu, Z., Cao, Y., Yang, D. 2022. Revisiting the pan evaporation trend in China during 1988–2017. Journal of Geophysical Research: Atmospheres. 127(e2022JD036489). https://doi.org/10.1029/2022JD036489

[64] Clarke, H., Nolan, R.H., De Dios, V.R. et al. 2022. Forest fire threatens global carbon sinks and population centres under rising atmospheric water demand. Nat Commun. 13(7161). https://doi.org/10.1038/s41467-022-34966-3

[65] Jain, P., Castellanos-Acuna, D., Coogan, S.C.P. et al. 2022. Observed increases in extreme fire weather driven by atmospheric humidity and temperature. Nat. Clim. Chang. 12:63–70. https://doi.org/10.1038/s41558-021-01224-1