Bati *et al. Environmental Systems Research* (2023) 12:9
https://doi.org/10.1186/s40068-023-00290-y

Environmental Systems Research

RESEARCH

Open Access

Effects of landscape attributes and climate variables on catchment hydrology

Hirpo Gudeta Bati[1*], Tena Alamirew Agumassie[2], Tenalem Ayenew Tegaye[3] and Mulugeta Dadi Belete[4]

Abstract

Catchments characteristics, such as geomorphology, geology, soil, land use, and climatic variables, play an important role in total stream flow responses, a critical resource for people and the environment. Most of the previous literatures were applied a conventional statistical regression model to assess the relationship between landscape-climate descriptors, and streamflow and PET. However, a conventional statistical regression model didn't consider dependence of explanatory variables that were collected or extracted across both space and time. This paper investigated the impacts of landscape attributes and climate variables on catchment scale temporal variation of total streamflow and spatio-temporal variation of potential evapotranspiration (PET) in the Mille catchment using multiple linear regression techniques, and the importance of this study was to test spatial autocorrelation in the spatial regression model which is required to properly assess and quantify the relationship between hydrological regime response components and Landscape-climate descriptors in a catchment with topographically complex, and high spatio-temporal climatic variation like in our case study area, the Mille catchment. Statistical regression analysis revealed significant relationships between streamflow and climate variables, especially with rainfall. Mean maximum temperature is the most dominant factor controlling temporal variation of potential evapotranspiration at a monthly scale, whereas NDVI is the most significant factor that controls the spatial variability of PET. The multiple regression model shows that 91.1% of temporal variation in streamflow was accounted for rainfall, whereas, 96.6% and 78.4% of temporal and spatial variation in potential evapotranspiration was accounted for in maximum temperature and NDVI, respectively. Methods also can be applied to catchments with similar landscape attributes and climate variables.

Keywords Landscape, Climate, NDVI, PET, Stream flow

Introduction

Catchment landscape components, such as geomorphology, geology, soil, and climatic characteristics, play a significant role in total stream flow, base flow and surface runoff spatiotemporal variation (Zhu and Day 2009). Land use and land cover, which is an integral component of the landscape, critically alter the hydrological regime response descriptors like surface runoff and evaporation (Mao and Cherkauer 2009). The land use and land cover change affect the rainfall and temperature pattern, the fundamental driving forces of the hydrological cycle (Kittel 2000). Subsequently, it alters the catchments' water balance that exists between evaporation, groundwater recharge, and stream flow (DeFries and Eshleman 2004).

Streamflow with various magnitudes is extremely disturbed by topographic attributes like elevation, drainage network, and climate variables (Mohamoud 2004; Wei and Zhang 2010), and watershed soil moisture variability (Qiu et al. 2001). Hydrological response descriptors like baseflow are also significantly affected by climate

*Correspondence:
Hirpo Gudeta Bati
hirpo.gudeta@aau.edu.et
[1] Africa Center of Excellence for Water Management, Addis Ababa University, Addis Ababa, Ethiopia
[2] Water and Land Resource Center (WLRC), Addis Ababa University, Addis Ababa, Ethiopia
[3] Schools of Earth Sciences, Addis Ababa University, Addis Ababa, Ethiopia
[4] School of Water Resource Engineering, Institute of Technology, Hawassa University, Hawassa, Ethiopia

descriptors like rainfall, and landscape components like geology, relief, land use, and soil attributes (Price 2011). According to Woods et al. (1997), spatial variability of soil moisture content strongly affects catchments' hydrology, which may be predominantly controlled by catchment topography. Evaluating the impacts of landscape-climate descriptors on catchments' hydrological regime response is important for the management and development of catchments.

One of the major challenges with climate variables and their changes is their impact on water resources and extreme hydrological events (Andersson et al. 2011). Semiarid regions of developing countries, including Ethiopia, which faces major water resource stress, are likely to be the most severely impacted (Misra 2014; Nations 2007). Developing countries like Ethiopia, where water is a stressed resource, and their river basins of arid and semi-arid provinces, are facing extreme land use land cover changes as a result of rapidly increasing population number and socio-economic developments (Chatterjee 2018; Kassas 1976), and consequently, they drastically affect stream flow variability (Pervez and Henebry 2015). Additionally, in recent years many of the river basins in developing countries are facing extremely adverse stream flow conditions, such as floods and droughts (Garg et al. 2017).

Understanding the factors that drive the temporal fluctuation of catchments' streamflow and potential evapotranspiration is essential for predicting how those descriptors will respond to landscape-climate predictors (Chiverton et al. 2015; Ersi et al. 2022; Hatfield and Prueger 2014; Xiao et al. 2019), and helps water resource planning and management related issues like the delivery of drinking and agricultural irrigation water, and industrial water use (Verstraeten et al. 2008). Most of the aforementioned literature used a conventional statistical regression model at the regional scale to assess and quantify the landscape-climate explanatory variables as a factor that affects hydrological regime response descriptors like stream flow and potential evapotranspiration. However, the conventional statistical regression model was not considered for the spatially referenced data that was collected or extracted across both space and time, because there is an assumption that the explanatory variables are independent, but in reality that is not (Lei Ji and A. J.Peters 2004). Instead, a spatial regression method that can correct for spatial autocorrelation in the regression model (Lei Ji and Peters 2004; Tiefelsdorf 2000) is required to properly assess and quantify the relationship between hydrological regime response components and Landscape-climate descriptors in a catchment with topographically complex, and high spatiotemporal climatic variation like in our case study area, Mille catchment.

This research aims to analyze mean monthly stream flow and potential evapotranspiration (PET) with limited landscape-climate descriptors such as mean monthly rainfall, potential dryness index (PDI), mean monthly temperature, maximum temperature, land use land cover-normalized difference vegetation index (LULC-NDVI), soil water content (SWC), baseflow index (BFI), easting and elevation (for spatial PET) using temporal and spatial regression method that adjusted for spatial autocorrelation.

This research attempts to answer the following questions regarding stream flow-potential evapotranspiration and limited landscape-climate predictors at the catchment scale: (1) which landscape-climate descriptors are likely to affect the stream flow temporal variation? (2) Which landscape-climate descriptors are the factors that drive spatiotemporal variation of potential evapotranspiration within the Mille catchment? and (3) which landscape or climate predictor variable has the dominant impact on streamflow and potential evapotranspiration? The results can help improve our understanding of how limited landscape attributes and climate variables affect the stream flow and potential evapotranspiration variability, thus assisting water resource management and planning throughout the catchment.

Materials and methods

Study area

Mille River catchment is located in the Awash river basin of Northeast Ethiopia lies between 11°26′ and 11°46′N latitude and 39°38′ to 40°46′E longitude**,** and covers an area of 442 km^2 (Figure 1). Complex mountainous terrain and various climatic conditions, including humid to arid weather, characterize the catchment. The intra-annual spatial rainfall distribution in the study area is bimodal, which is mainly affected by the Intertropical Convergence Zone (ITCZ) (Fekadu 2015). Bati et al. (2022) showed that the combination of altitude and easting as an explanatory variable highly affects the spatial pattern of the Mille catchment's rainfall and temperature.

From Fig. 2, we can observe that the catchment's water loss as potential evapotranspiration is on average ten folds than the mean areal rainfall. Except for July and August, the rest months from September to June the catchment were under water deficiency. The lowest PET is in December which is 116.58 mm, while the highest PET was in June, 173.60 mm.

Datasets and methods

Datasets

Daily meteorological datasets were obtained (Fig. 1) from the National Meteorological Agency (NMA), Ethiopia, from 1 January 1983 to 31 December 2002. These

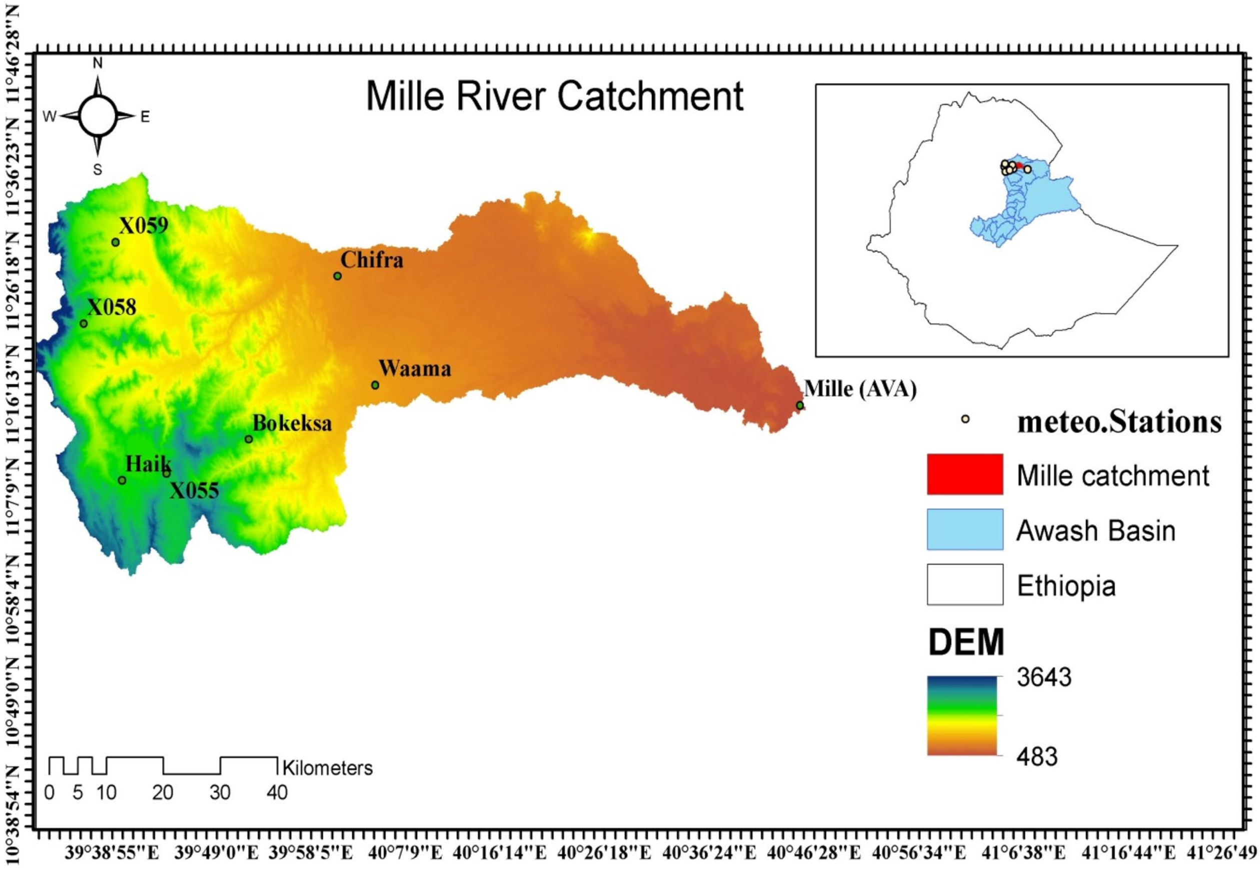

Fig. 1 Study area, Mille catchment

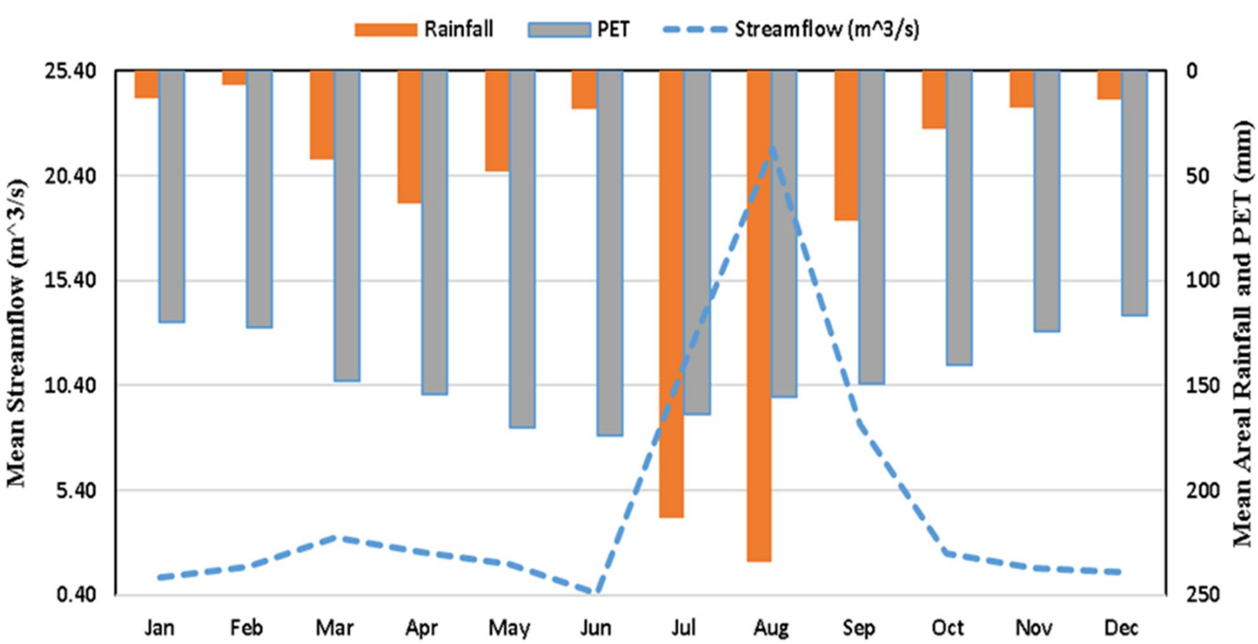

Fig. 2 Mean monthly streamflow (m.^{3}/s), areal rainfall, and PET (mm)

meteorological datasets include rainfall (R), maximum and minimum air temperature (T_{min} and T_{max}), relative humidity (RH), solar radiation (SR), and wind speed (W_s). According to Hargreaves and Samani (1985), a recommendation based on two main reasons such as the problem related to the quality of climatological data, and the possible errors in connection with the more sophisticated methods, Potential evapotranspiration at each meteorological station was estimated using the Hargreaves method stated as:

$$ET_0 = 0.00023 * RA * TD^{0.50}\left(T^0C + 17.8\right) \quad (1)$$

where ET_0 is potential evapotranspiration (reference crop evapotranspiration), RA is extraterrestrial radiation, TD is mean maximum minus mean minimum temperature, and T^0C is mean temperature

The mean monthly areal rainfall (mm), temperature(°c), and potential evapotranspiration (mm) were estimated using kriging with external drift (KED) (Goovaerts 2000; Bati et al. 2022; Webster 2015), a geostatistical interpolation technique with elevation as an auxiliary variable, respectively. The potential dryness index (PDI) was calculated as the ratio between mean monthly areal potential evapotranspiration and mean monthly areal rainfall. Average daily stream flow at the catchment outlet (Lat: 11°25′ N and Long: 40°46′ E) from 1983 to 2002 was provided by the Department of Hydrology, Ministry of Water and Energy, Ethiopia. Average soil water content (SWC) (mm) extracted from SWAT+ model calibrated dataset (not shown here).

Monthly point and spatial areal NDVI was extracted from MOD13A1-006 MODIS/Terra Vegetation Indices having 500m spatial and 16 days temporal resolution from NASA (https://doi.org/10.5067/MODIS/MOD13A1.006), accessed 12 November 2022 using Application for Extracting and Exploring Analysis Ready Samples (AρρEEARS) software (https://appeears.earthdatacloud.nasa.gov). The data was extracted from 2000 to 2002, and then the monthly mean NDVI was averaged.

Baseflow (mm) was estimated from total stream flow records using the Base flow Digital Filter Program (https://swat.tamu.edu/software/), an automated base flow separation technique(Arnold et al. 1995; Arnold, 1999), and then baseflow index (BFI) was calculated from the historical records as the ratio of base flow to total streamflow used as a surrogate for geological attributes.

Monthly landscape-climate descriptors datasets such as mean monthly areal soil water content (SWC), NDVI, mean monthly areal rainfall, mean monthly areal temperature, potential dryness index (PDI), and hydrological regime response like mean monthly total streamflow and potential evapotranspiration were prepared for regression analysis (Table 1). In addition, Elevation and Easting were extracted from 30m SRTM DEM provided by NASA/USGS (https://earthexplorer.usgs.gov), accessed 8 June 2021.

Statistical analysis and modeling technique

Statistical analysis and modeling were conducted using R programming software (https://www.R-project.org/). Regression analysis was computed to examine relationships between hydrological regime responses like stream flow and potential evapotranspiration (PET), and landscape-climate descriptors like elevation, easting, PDI, mean monthly rainfall, mean monthly temperature, LULC_NDVI, BFI and S_SWC (Average). The regression analysis procedure involved the following two steps:

(a) **A multiple linear regression analysis** was performed using the response variable such as mean monthly streamflow and potential evapotranspiration, and a set of explanatory variables.

The multiple linear regression model mathematical equations are expressed as:

Table 1 Statistical results for all collected data in the study

Statistics	Streamflow	PET	Rainfall	PDI	NDVI	BFI	SWC	Tmean
Mean	2.91	144.73	64.98	5.54	0.37	0.64	56.05	24.69
Standard Error	1.09	5.78	20.92	1.10	0.02	0.09	3.57	0.76
Median	1.31	148.45	38.22	4.81	0.36	0.76	54.29	25.07
Standard Deviation	3.77	20.01	72.47	3.81	0.08	0.30	12.36	2.62
Skewness	2.19	-0.11	1.65	0.23	0.90	-0.94	0.84	-0.13
Minimum	0.28	116.58	12.97	0.79	0.28	0.10	38.47	20.54
Maximum	13.13	173.60	221.52	11.51	0.52	0.95	80.12	28.34

NB: PET = Potential Evapotranspiration (mm), NDVI = Normalized Difference Vegetation Index, BFI = Baseflow Index, SWC = Soil Water Content (mm), streamflow (mm), Rainfall (mm), Tmean = Long year-based mean annual temperature (°c)

Table 2 Correlation coefficients of streamflow versus landscape-climate explanatory variables

	Streamflow	Rainfall	PDI	SWC	NDVI	BFI	T_mean
Streamflow	1.000	0.932	− 0.687	0.751	0.611	− 0.815	0.361
Rainfall		1.000	− 0.779	0.630	0.436	− 0.903	0.552
PDI			1.000	− 0.623	− 0.485	0.913	− 0.497
SWC				1.000	0.839	− 0.566	0.315
NDVI					1.000	− 0.366	− 0.022
BFI						1.000	− 0.534
T_mean							1.000

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_i x_i \tag{2}$$

where y is the response variable. β_0 is the intercept, x_1 is the 1st explanatory variable, β_1 is the slope coefficient, β_i is the slope coefficient for the ith explanatory variable, ε is the remaining unexpected noise in the data (error)

The values for β_0, β_1, β_2..., β_i were estimated using the ordinary least square method, which minimizes the sum of squared residuals (RSS):

$$RSS = \sum \left(Y_i - \hat{Y}_i \right)^2 \tag{3}$$

where: yi: The actual response value for the ith observation. ŷi: The predicted response

Multiple regression models were developed and fitted with ordinary least squares (OLS) Procedures, a technique used to estimate the regression model parameters (Kroll and Song 2013).

(b) **Checking multicollinearity**

One important assumption of multiple linear regressions is that no explanatory variable in the regression model is highly correlated with another explanatory variable in the same model but in reality that is not. Multicollinearity is the condition where at least one explanatory variable is closely related to one or more other explanatory variables (Helsel et al. 2020), which influences the coefficient estimates and the p-values. Such adverse consequences are magnified when sample sizes are smaller (like in our case), correlations between variables are higher, and model error variances are higher (Kroll and Song 2013). Before the selection of the most useful predictor variables, we detected multicollinearity in the model using the variance inflation factor (VIF) (Kroll et al. 2004). Predictors that have VIF values that are greater than 10 are the cause of multicollinearity, and removed one or more of the highly correlated variables (Kroll and Song 2013).

$$\mathbf{VIF_i} = \frac{1}{\left(1 - \mathbf{R_i^2}\right)} \tag{4}$$

where R_i^2 is the R^2 from a regression of the ith explanatory variable on the other explanatory variables, and the equation is used for adjustment of x_i in multiple regression equation (Eq. 1).

R tool was used to do multiple linear regression analyses because it is free, powerful, and widely available (https://www.R-project.org/).

Results

Streamflow regression analysis

Table 2 shows Pearson correlation coefficients between mean monthly stream flow and limited landscape-climate explanatory variables. Stream flow is correlated positively to rainfall, SWC, NDVI, and temperature, and negatively correlated to PDI, and BFI.

The estimation of the full multiple linear regression model (Eq. 2) is shown in Table 3. The analysis of variance (ANOVA) indicates that overall the model was performed significantly (p-value < 0.002305, Adjusted R^2 = 0.9131). Also, the predictor variable, rainfall is statistically significant (p-value = 0.025) at the 0.05 significance level, while the PDI, T_{mean}, SWC, NDVI, and BFI are not. However, we can see that the variance inflation factor (VIF) value for BFI is greater than 10 (Fig. 3). This indicates that it is highly correlated with other explanatory variables in this

Table 3 Coefficient estimates of the full regression model for stream flow

	Coefficients Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	− 0.5862	5.73941	− 0.102	0.923
R	0.04229	0.01336	3.164	0.025*
PDI	0.36878	0.2731	1.35	0.235
T_{mean}	− 0.20945	0.19027	− 1.101	0.321
SWC	0.0708	0.06559	1.079	0.33
NDVI	6.10533	10.76269	0.567	0.595
BFI	− 4.30561	5.06655	− 0.85	0.434

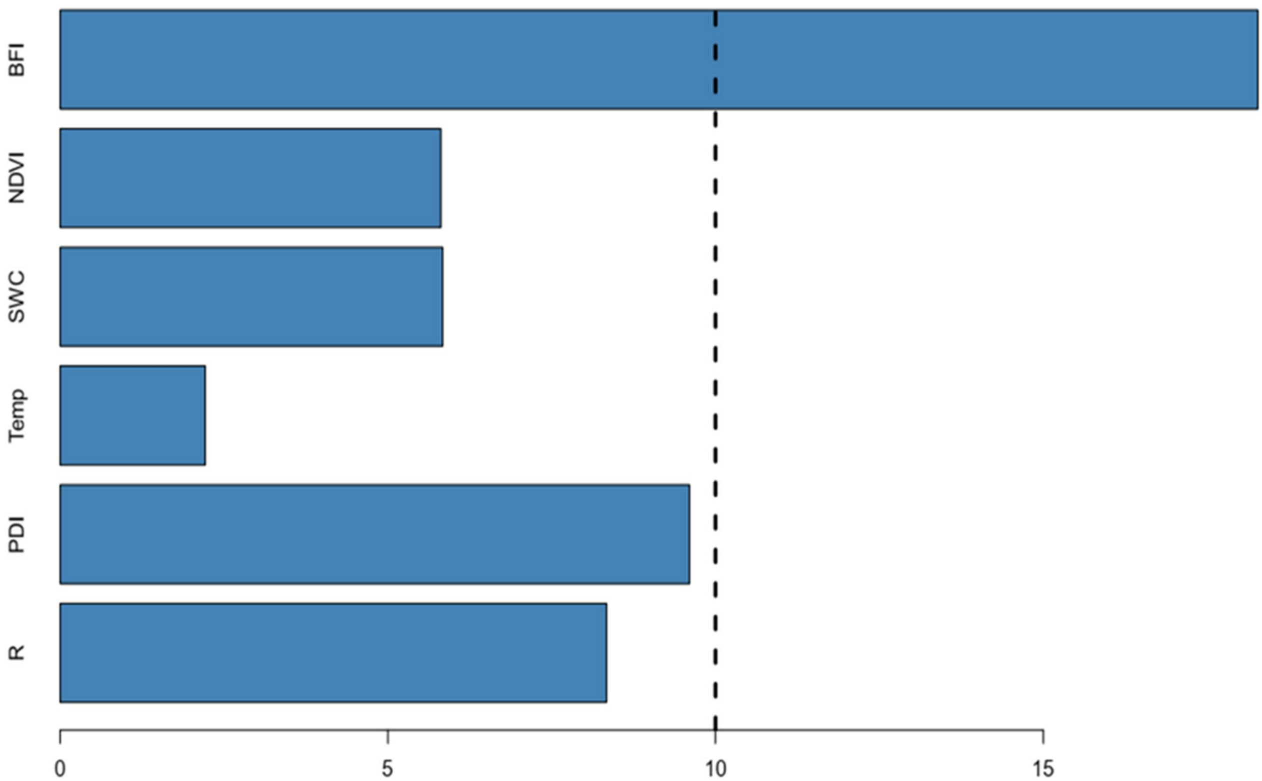

Fig. 3 VIF values of explanatory variables for the full regression model

regression, which means that multicollinearity is likely to be a problem in this regression model.

In such cases, the coefficient estimates and p-values in the regression output are likely unpredictable. Thus, based on Kroll and Song (2013), we remove one variable at a time, which is the cause for multicollinearity, and repeat the process until all problem-causing explanatory variables are removed.

From Fig. 3, BFI has a VIF value greater than 10 that causes multicollinearity to have been removed from the regression. So, finally, we were left with a list of variables that have no or very weak correlation between them (Table 4).

From the regression model output, the adjusted R-Squared value for the model is 0.9171, the overall F-statistic is 25.35 and the corresponding p-value is 0.0005753<0.001 significance level, which shows that the overall regression model is significant with the VIF value for all predictor variables were less than 10. Nevertheless, predictor variables such as PDI, T_{mean}, SWC, and NDVI are still showing less significance on the response variable, stream flow, whereas rainfall is highly significant at $\alpha = 0.001$ level of significance. The regression model equation using the final regression model output is written as:

Table 4 Coefficient estimates of the subset regression model for streamflow

	Coefficients Estimate	Std. error	t value	Pr(>|t|)	VIF
(Intercept)	− 0.523	5.605	− 0.093	0.929	
R	0.0513	0.008	6.441	0.000663***	3.10
PDI	0.175	0.146	1.194	0.278	2.89
Tmean	− 0.248	0.181	− 1.373	0.219	2.09
SWC	0.076	0.064	1.183	0.282	5.80
NDVI	2.81	9.805	0.287	0.784	5.05

NB:—Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Adjusted R-squared: 0.9171; F-statistic: 25.35, p-value: 0.0005753

$$Q = -0.52 + 0.05R + -0.25\mathrm{T}_{\mathrm{mean}} + 0.18PDI + 0.08\mathrm{SWC} + 2.81\mathrm{NDVI} \tag{5}$$

where Q is streamflow (mm)

In multiple regression Model 5, a per unit increase of mean rainfall depth, potential dryness index, soil water content and normalized difference vegetation index produce an increase in percentage stream flow values of 0.05, 0.18, 0.08, and 2.81, respectively, whilst per unit

Table 5 Correlation coefficients of the PET and predictor variables

	PET	Rainfall	PDI	NDVI	BFI	SWC	Temp
PET	1.000	0.477	− 0.465	− 0.018	− 0.470	0.330	0.983
Rainfall		1.000	− 0.779	0.436	− 0.903	0.630	0.552
PDI			1.000	− 0.485	0.913	− 0.623	− 0.497
NDVI				1.000	− 0.366	0.839	− 0.022
BFI					1.000	− 0.566	− 0.534
SWC						1.000	0.315
Temp							1.000

Table 6 Statistic results of the full regression model for mean monthly PET

	Coefficients Estimate	Std. error	t value	Pr(>|t|)	VIF
(Intercept)	− 49.53463	21.05247	-2.353	0.065323	
R	− 0.03779	0.04902	− 0.771	0.475585	8.34
PDI	− 0.47231	1.00174	-0.471	0.657145	9.60
SWC	0.26713	0.2406	1.11	0.317403	5.84
NDVI	− 23.04782	39.47817	− 0.584	0.584691	5.81
Temp	7.68126	0.69791	11.006	0.000108	2.21
BFI	6.13762	18.58439	0.33	0.754595	18.28

Table 7 Statistical estimates of the subset regression model for the temporal regression model

	Coefficients Estimate	Std. error	t value	Pr(>|t|)	VIF
(Intercept)	− 49.626	19.425	− 2.56	0.0432	
Temp	7.736	0.626	12.37	1.71E-05	2.09
R	− 0.051	0.028	− 1.83	0.1163	3.10
SWC	0.260	0.22122	1.18	0.2837	5.80
PDI	− 0.196	0.50744	− 0.386	0.7129	2.89
NDVI	− 18.35	33.98202	− 0.54	0.6087	5.05

decreases of stream flow value attributed to mean temperature is 0.25%.

Potential evapotranspiration

Regression model for temporal variation

The relationship between mean monthly potential evapotranspiration and predictor variables in the Mille catchment has been analyzed by developing the correlation coefficients. Table 5 showed that the correlation coefficient of mean monthly PET is positively correlated to rainfall, SWC, and mean T_{mean}, negatively correlated to PDI and BFI, and almost no relationship between NDVI.

Table 6 indicated that the VIF value for PET of the temporal regression model was like that of the stream flow regression model (see Fig. 3). Overall, the full temporal regression model for PET estimated with the OLS procedure was significant at F-statistics = 43.27, Adjusted R-squared = 0.9584 and p-value = 0.0003762 < 0.001 significance level, respectively.

From the regression model output last column, the VIF value for the BFI predictor variable is greater than 10, and this shows that the predictor variable is likely suffering from a multicollinearity problem, which affects the regression coefficient and the p-value of the predictor.

By removing the predictor in the model that has a high VIF value from the dataset (Kroll and Song 2013) and by doing so, it can solve multicollinearity (Table 7).

$$\begin{aligned} \text{PET} = & -49.63 + 7.74\text{Temp} + \\ & -0.051\text{R} + 0.26\text{SWC} \\ & -18.35\text{NDVI} - 0.196\text{PDI} \end{aligned} \tag{6}$$

Overall performance of the subset regression model was highly significant (F-statistic = 60.95, p= 4.62E-05< 0.001, Adj_R^2 = 0.9646), specifically the mean monthly temperature is statistically highly significant at the 0.001 significance level, and the rainfall is statistically less significant at 0.1 significance level while SWC, NDVI, and PDI are not statistically significant predictors. The final regression model is written as:

Spatial multiple regression model

To improve a better understanding of the correlation between the response variable (in our case PET) and spatial explanatory variables, we created a correlation matrix to view the linear correlation coefficients between each pair of variables (Table 8). From the correlation matrix,

Table 8 Correlation matrix of PET versus landscape-climate predictor variables

	PET	Rainfall	elev_90m	Mille Easting	Tmax	NDVI
PET	1.000	− 0.807	− 0.881	0.815	0.854	− 0.885
Rainfall		1.000	0.910	-0.974	− 0.843	0.913
elev_90m			1.000	− 0.875	− 0.981	0.951
Mille Easting				1.000	0.817	− 0.917
Tmax					1.000	− 0.937
NDVI						1.000

Table 9 Statistical results for full regression of the relationship between spatial PET and landscape-climate predictors for the mean annual. The climatic variables included in the analysis were mean elevation, mean annual rainfall, easting, NDVI, and maximum temperature

	Coefficients Estimate	Std. error	t value	Pr(>|t|)	VIF
(Intercept)	2.04E+03	9.46E+03	0.215	0.85	
R	1.24E+00	2.17E+00	0.571	0.625	38.58
Elev	− 9.23E-01	1.33E+00	− 0.694	0.559	81.09
Easting	4.34E-03	1.12E-02	0.386	0.737	27.99
NDVI	− 1.18E+03	2.84E+03	− 0.416	0.718	19.11
Tmax	− 5.84E+01	1.19E+02	− 0.492	0.671	52.36

we can conclude that PET is strongly positively correlated with easting and temperature, and negatively correlated with rainfall, elevation, and land cover (NDVI). Each predictor variable exhibits a high correlation with the other, which shows there is a multicollinearity problem.

Table 9 shows that all predictors have a VIF value greater than 10, and all predictors are statistically less significant at the 0.05 significance level. So, we removed a predictor that has a high VIF value sequentially until the predictor/s left with a VIF value of less than 10 (Kroll and Song 2013).

For instance, if we look at the coefficient estimated for the temperature, the model is telling us that for each additional one unit increase in maximum temperature, the average increase in PET is -5.84E+01 mm, assuming rainfall, mean Elevation, easting, and NDVI are held constant. This doesn't seem to make sense, considering we would expect the mean temperature with high magnitude and thus have a higher PET. The reason behind this effect was a multicollinearity that affects the estimates of coefficients (Helsel et al. 2020). Following the same procedure applied for stream flow and temporal PET, we removed predictors in the model having a high VIF value step by step from the datasets (Kroll and Song

Table 10 Statistical estimates of the subset regression model for the linear regression model

	Coefficients. Estimate	Std. error	t value	Pr(>|t|)
(Intercept)	2508.2	163.4	15.35	4.84E-06
NDVI	− 1986.3	425.6	− 4.667	0.00344

2013), and lastly, we solved an issue of multicollinearity (Table 10). Accordingly, land use land cover NDVI is statistically highly significant (F-statistic = 21.78, Adj.R^2 = 0.75, p-value = 0.003443 < 0.01 significance level). The final spatial regression model is written as:

$$PET_{spatial} = 2508.2 - 1986.3\text{NDVI} \tag{7}$$

where $PET_{spatial}$ is the spatial potential evapotranspiration

Discussion

Landscape-climate predictors' effect

Explaining the spatiotemporal variability of catchment hydrological regime response in connection with catchment landscape attributes and climate variables is a major concern of catchment hydrology. In this study, further analysis and discussion on the results of these multiple linear regression models are summarized below.

Streamflow temporal variation

In this paper, a multiple regression equation, for long year-based mean monthly stream flow for the Mille River Catchment was constructed using multiple linear regression techniques. This regression model expressed the quantitative relationship between mean monthly stream flow, and landscape-climate predictors that were used to quantify the impacts of limited landscape-climate predictors on stream flow temporal variation at catchment scale (see Table 4). The multiple linear regression analysis

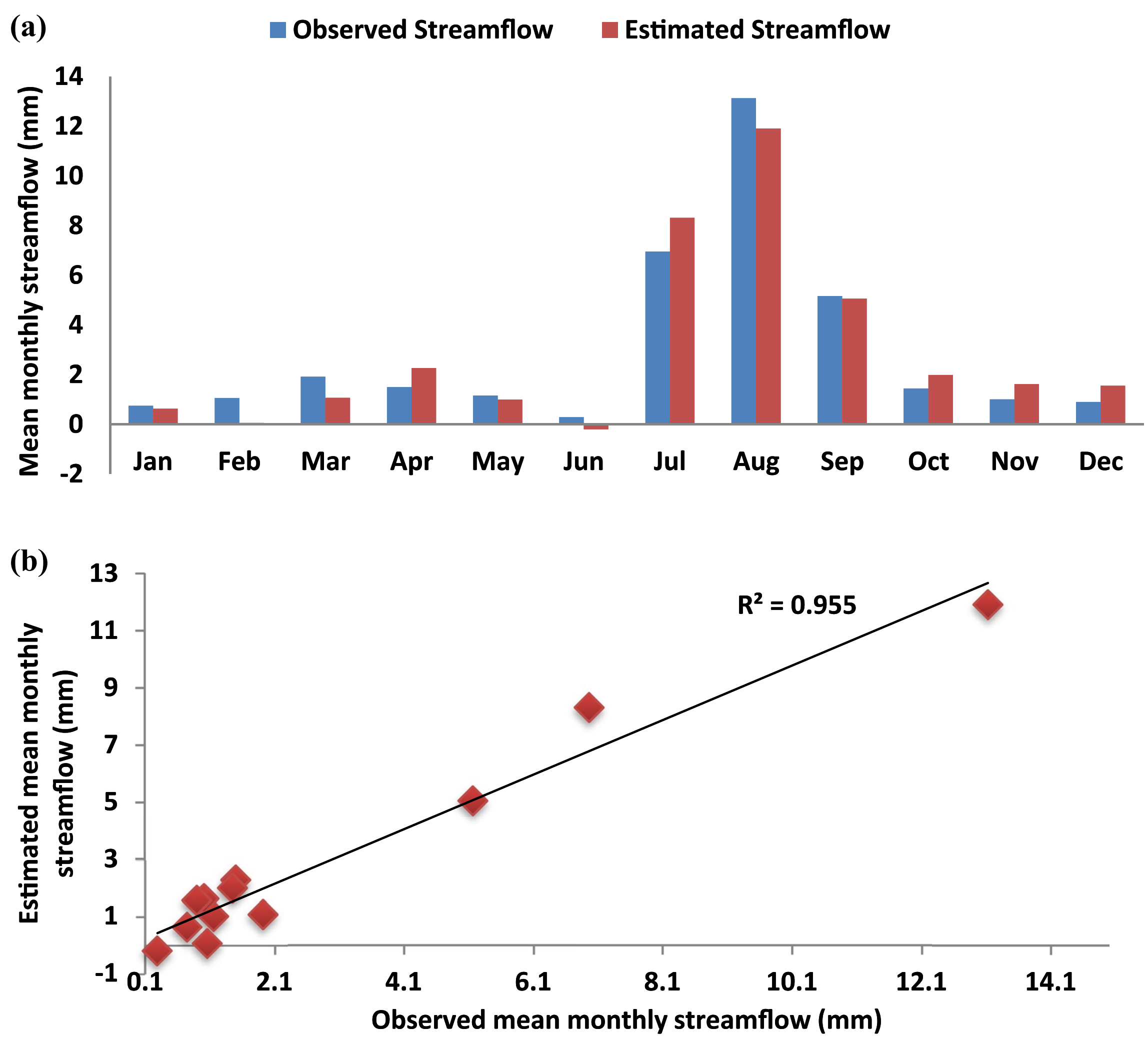

Fig. 4 Mean monthly streamflow chart **a**, comparison of observed and estimated mean monthly streamflow **b**

of mean monthly stream flow with limited landscape attributes and climate variables for mean monthly stream flow prediction (Fig. 4) was analyzed, and our regression analyses demonstrated that the regression model was overall statistically significant ($F_{statistics} = 25.35$, Adj.$R^2 = 0.9171$, P-value $= 0.0005753 < 0.001$ significance level). However, among all the explanatory variables, only rainfall is strongly significant for explaining temporal variation in stream flow.

The regression model showed that mean monthly rainfall is the most important contributor to stream flow temporal variation, while mean temperature, PDI, SWC, NDVI, and BFI are relatively less important contributors. As shown in Figure 5, the mean monthly streamflow increased with increasing observed mean monthly rainfall, with a coefficient of determination of 0.911. The analysis showed that the temporal variation of the response variable, mean stream flow was significantly influenced by rainfall temporal variation. This could be because of the orographic effect of the catchment on mountainous terrain on spatial and temporal variation of rainfall (Goovaerts 2000; Bati et al. 2022).

PET temporal variation

Temporal variation of potential evapotranspiration in the topographic complex catchment depends on a few dominant landscape-climate descriptors (Hatfield and Prueger 2014; Zhang et al. 2010). In this paper, we combined a correlation and regression analysis to identify the dominant influence factors on potential evapotranspiration temporal variation. Specifically, we developed a regression model to analyze the influence of mean monthly areal rainfall, temperature, PDI, SWC, NDVI, and BFI on mean monthly PET. The regression model depicts that

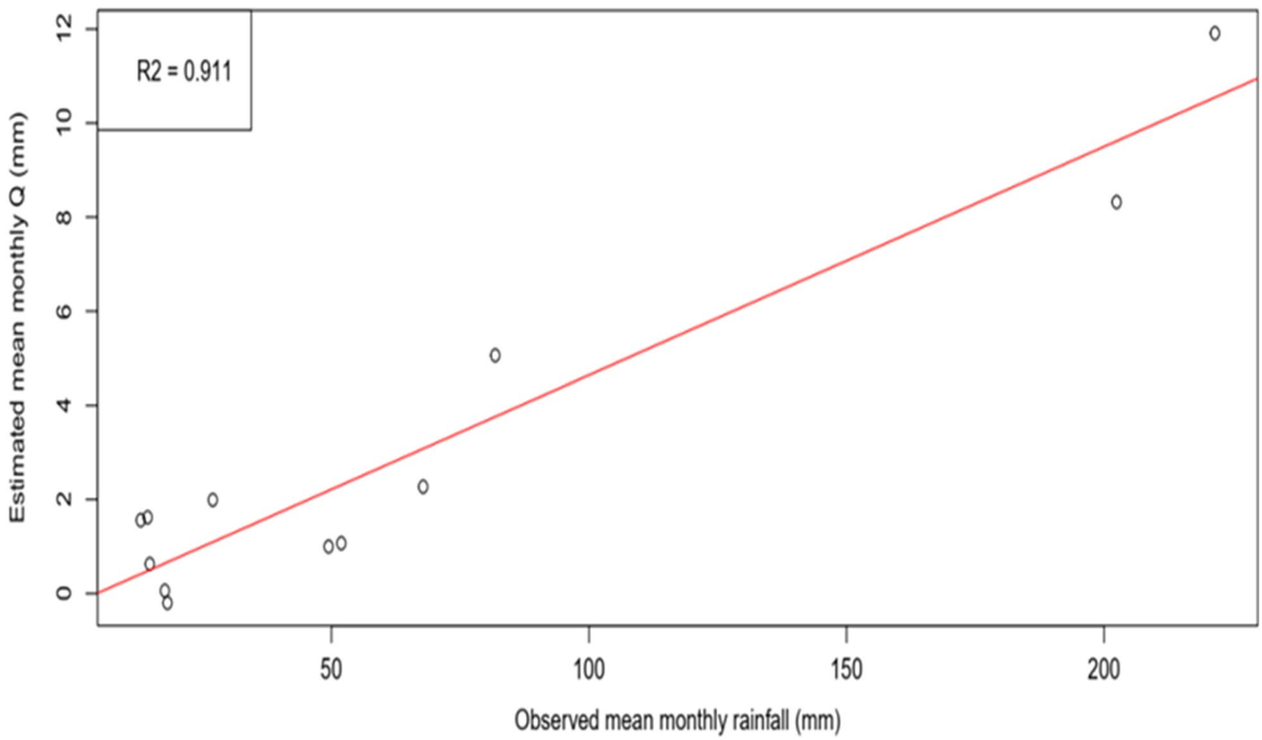

Fig. 5 Relationship between mean monthly estimated streamflow (Q) with mean monthly rainfall

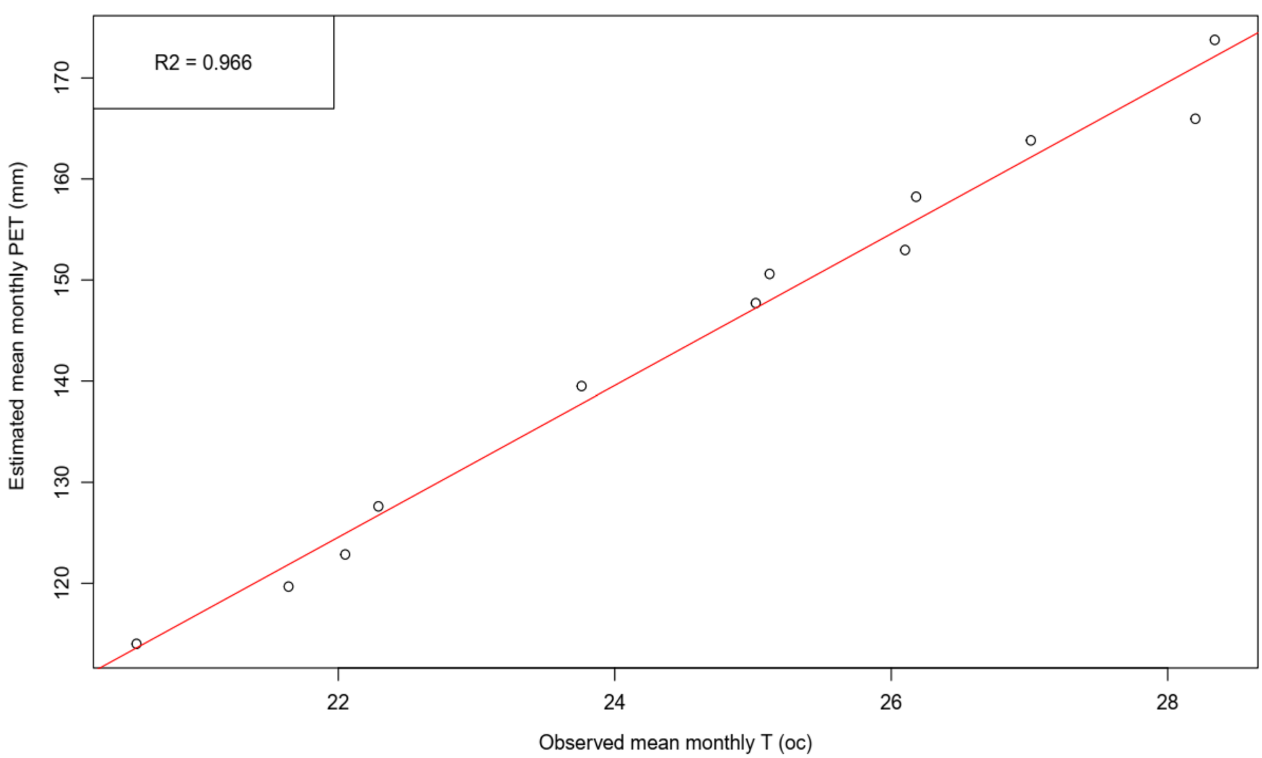

Fig. 6 Plot of the true linear relation between the estimated PET and the observed mean monthly temperature, and 12 estimations of the PET for mean monthly maximum temperature values from 1 through 12

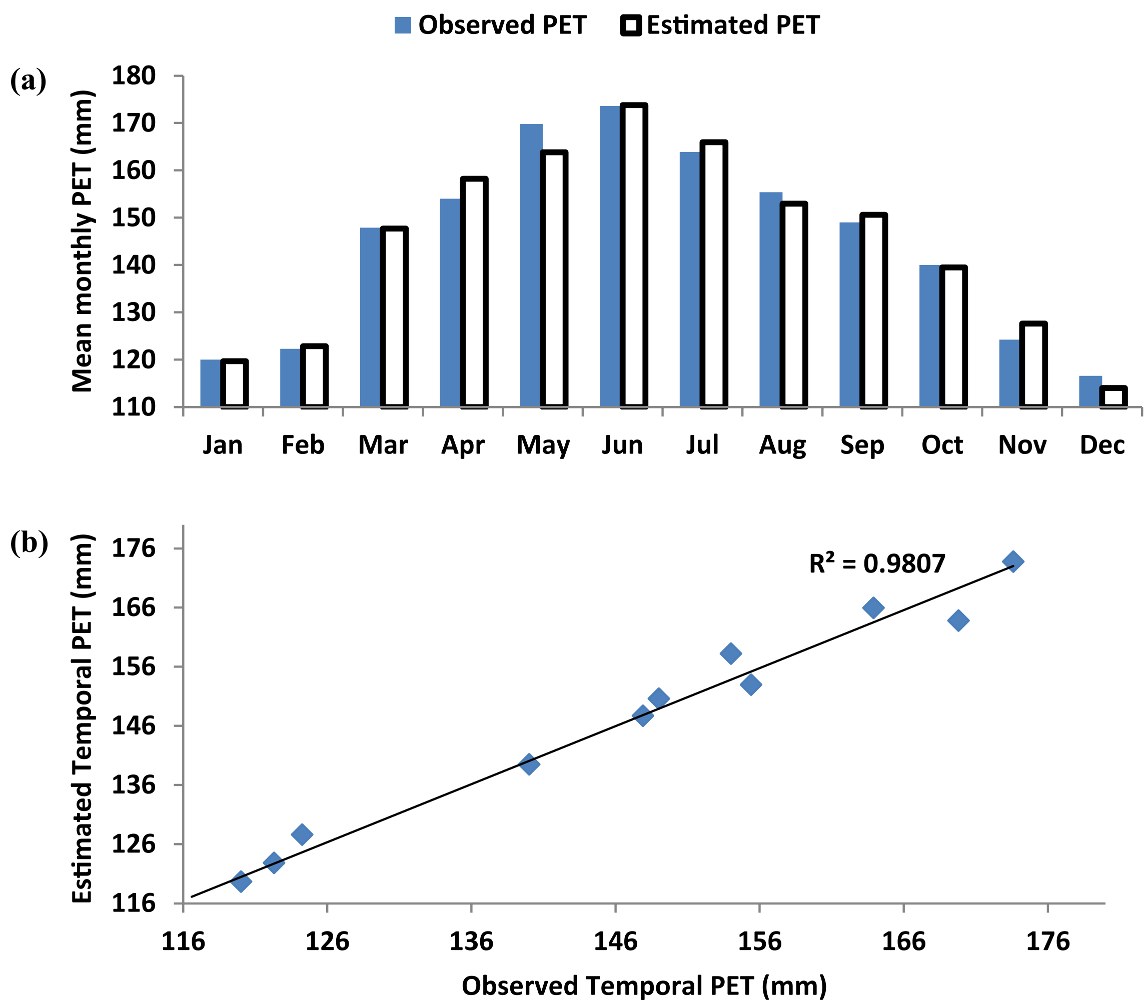

Fig. 7 Estimated and observed monthly time series of the PET **a**, and Comparisons between PET predicted by Landscape-climate predictors and that derived from the measured temperature time series **b**

the mean monthly temperature is the most significant contributor to potential evapotranspiration deviation ($p< 0.001$) (see Table 7).

From Fig. 6, 96.6% of temporal variation in PET accounted for the mean monthly maximum temperature. However, no other predictors like rainfall, SWC, PDI, and NDVI were statistically significant for the regression model. The reason lack of significance may be because of the dominance of aridity to semi-aridity regime across the study area, Mille catchment. Figure 7a, b represent the observed and estimated mean monthly PET chart and scatterplot that shows the estimated and observed PET relationship, respectively. Thus, variation among months for potential evapotranspiration in the Mille catchment depends on the mean monthly air temperature across the regime, and this study agrees with conclusions drawn by other studies (Ye et al. 2014; H. Zhang and Wang 2021).

PET spatial variation

Spatial variation of potential evapotranspiration within the Mille catchment is strongly significant based on the results shown in Fig. 8a, b. We observed that spatial variation of potential evapotranspiration across the regime significantly increases from the uppermost head to the lower part of the landscape in the Mille catchment. We found that the spatial pattern of mean annual PET varied from 900 mm in the western study area to 2200 mm in the eastern part of the catchment. To investigate the potential drivers of the spatial PET variation, two meteorological variables (rainfall and maximum temperature), two topographic attributes (elevation and easting), and one land use land cover (NDVI) attribute were analyzed using OLS regression analysis.

As a result, the effect of the mean annual NDVI on the spatial distribution of PET is highly significant (see Table 10).

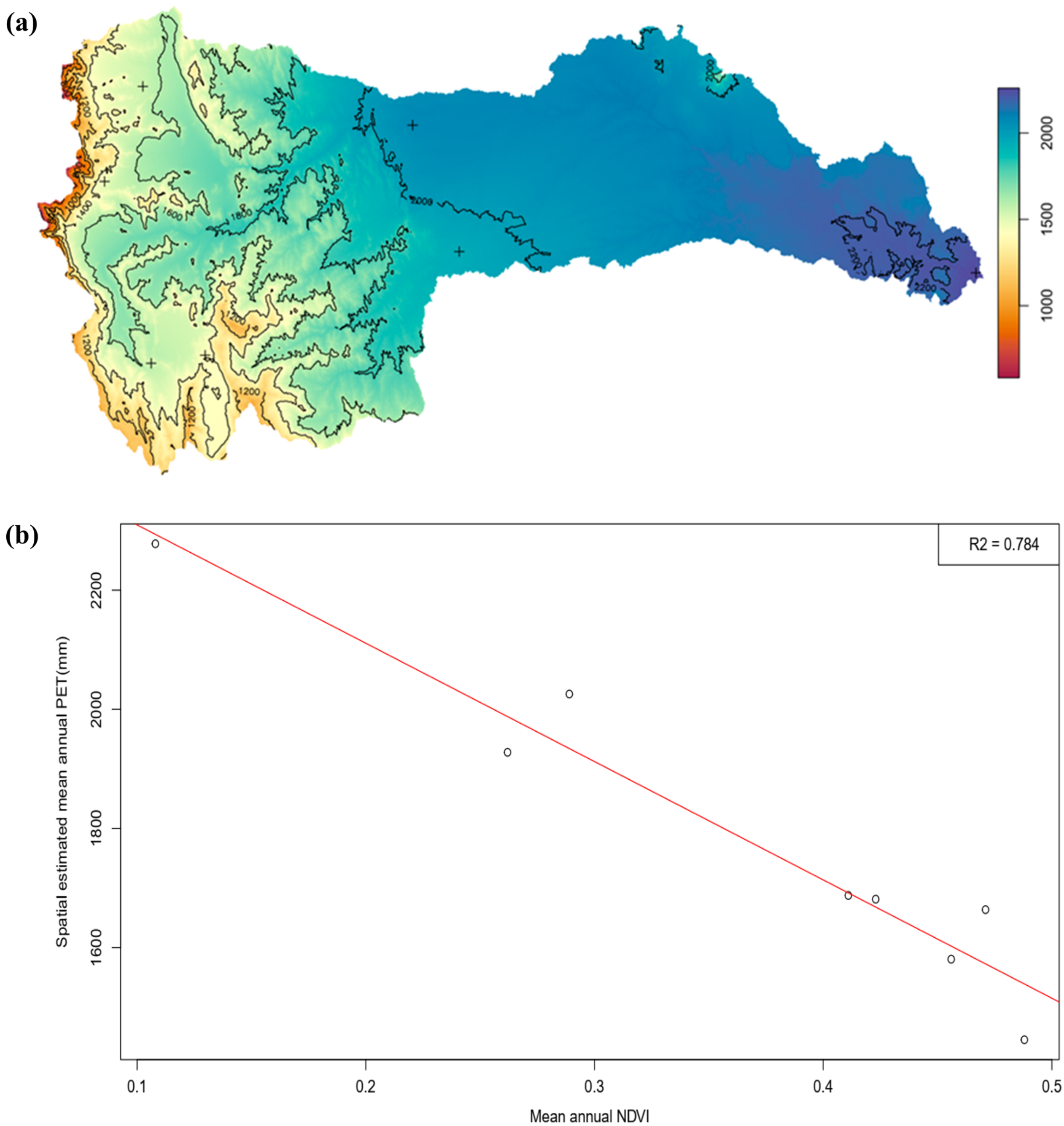

Fig. 8 Spatial patterns of observed mean annual PET in Mille catchment during 1983–2002 **a** and scatterplot between the observed PET and the land use land cover mean annual NDVI at the catchment scale **b**. As depicted in Fig. 8a, the mean annual PET is elevated from the head of the catchment to the lower part of the region

The location of the catchment lay in the core of the Great Rift Valley, which may be the main contributor to high diversity in the Agro ecological zone in the region, which resulted in high PET (Resources, 2009). Thus, we conclude that mean annual NDVI is the main factor influencing the spatial distribution of PET in the Mille catchment.

We developed a variogram model to show the occurrence of spatial autocorrelation of the regression residuals with a sample variogram (Fig. 9). The mean annual PET residuals experimental variogram was fitted with

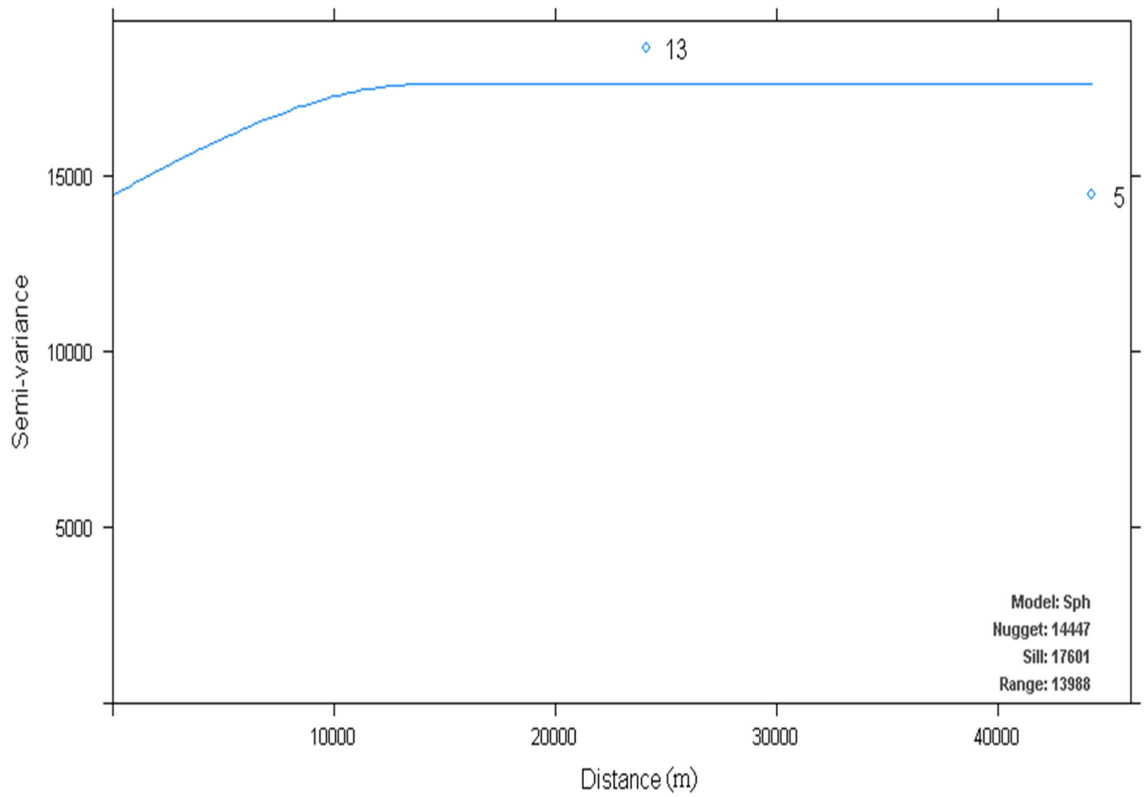

Fig. 9 Empirical variogram fitted with theoretical variogram using spherical variogram model for mean annual PET regression residuals

theoretical variograms using a spherical variogram model with a range of 13.99 km, a sill of 17601 mm^2, and a nugget effect of 14447 mm^2.

From Fig. 9, the residuals in regression analysis between any pair of weather stations within a distance range are spatially correlated to each other which violate the assumptions of linear regression, which states that there is no spatial relationship between consecutive residuals in time series data.

Conclusions

This research applied multiple linear regression techniques to explore the relationship between the catchment's landscape-climate predictors and hydrological regime response spatiotemporal variation, specifically stream flow and potential evapotranspiration at monthly and annual time scales. The results show that the dominant climate predictor on mean monthly stream flow temporal variation was the mean monthly areal rainfall that accounts for 95.5% of the stream flow variation. Other studies using multiple regression models have also demonstrated that the precipitation is the main climatic factor for streamflow temporal variation (e.g. Ali et al. 2021), which supports our findings. The complex mountainous area of the upper catchment resulted in high spatial areal rainfall (Bati et al. 2022), and this may possibly be a source for a significant relationship with lowland gauged stream flow (Viviroli et al. 2007), in our case the Mille River.

A comprehensive assessment of the effects of landscape attributes and climate variables on the spatio-temporal variation of potential evapotranspiration was also presented via regression analysis that provides the basis for the identification of dominant factors which affect the spatio-temporal variation of the response variable. In temporal regression analysis, we found that among limited landscape-climate descriptors, only long year-based mean monthly temperature has a significant influence on potential evapotranspiration temporal variation showing that the climate component is a key for hydrological cycle temporal alteration, and these findings are like that reported by other authors (e.g. Neubert and Rannow 2014). An increase in mean monthly temperature increases the potential evapotranspiration which means that it enhances the moisture-holding capacity of the atmosphere thus temperature is the key change in the hydrological system, specifically alteration in the temporal variation of potential evapotranspiration at a monthly scale whereas, spatial variability in catchment's land cover depicted that the spatial correlation between potential evapotranspiration and land use land cover (NDVI) component is highly significant.

In multiple linear regressions, there is the possibility that some of the independent variables are correlated with one another, so it is important to check these before developing the regression model. In this paper, multicollinearity in the model was detected by using the variance inflation factor (VIF) so that one or more of the highly correlated predictors that have VIF values which are greater than 10 were removed.

Overall, understanding factors causing spatio-temporal variation in hydrological regime response like total streamflow and potential evapotranspiration will lead to improving capabilities for water management in various water use systems, and the results provide water resource planners at various sectors a sign of which catchment landscape-climate descriptors may aid in describing the hydrological regime response spatiotemporal variation. This information should aid in the modeling and management of total stream flow and potential evapotranspiration. Establishment of more flow gauging stations at different reaches of the river and with more distributed meteorological stations, a much better understanding of the spatial and temporal variability of the hydrology of the catchment could be achieved. However, with the existing limited data, the research demonstrated the dominant role of rainfall, temperature, and NDVI on hydrological system response specifically on the total stream flow and potential evapotranspiration of the catchment.

Abbreviations

PET	Potential evapotranspiration (mm)
NDVI	Normalized difference vegetation Index
BFI	Baseflow index
SWC	Soil water content (mm), streamflow (mm), rainfall (mm)
Tmean	Long year-based mean annual temperature (°c)
VIF	Variance inflation factor
Q	Total stream flow (mm)
OLS	Ordinary list square
SWAT+	Soil water assessment tool plus

Acknowledgements
The observed climate and stream flow datasets used for further applications were obtained from the National Metrology Institute of Ethiopia (NMIE), and the Ministry of Water and Energy (MoWE), respectively. We greatly appreciate the Awash Basin Administrative office, Melka Worer, Ethiopia for offering the vehicle for fieldwork.

Author contributions
HG conceptualization of the research, develop the methodology, and field data collected and analyzed, and wrote the original draft and final manuscript. TA, TA, and MD substantively edited the manuscript. All authors read and approved the final manuscript.

Funding
The research was financed through the Ph.D. scholarship (Grant No. ACEWM/GSR/4334/11) awarded to the first author by the Africa Center of Excellence for Water Management, Addis Ababa University, Ethiopia.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors have declared that no competing interests exist.

Received: 31 January 2023 Accepted: 5 March 2023
Published online: 29 March 2023

References

Ali M, Zaman Q, Ali S, Qasim M, Khalil U, Ahmad S, Ismail M, Muhammad S, Ali S (2021) Science of the total environment regression analysis of hydro-meteorological variables for climate change prediction : a case study of Chitral basin, Hindukush region. Sci Total Environ 793:148595. https://doi.org/10.1016/j.scitotenv.2021.148595

Andersson L, Samuelsson P, Kjellström E (2011) Assessment of climate change impact on water resources in the Pungwe river basin. Tellus Series Dynamic Meteorol Oceanogr 63(1):138–157. https://doi.org/10.1111/j.1600-0870.2010.00480.x

Arnold JG, Allen PM (1999) Automated methods for estimating baseflow and ground water recharge from streamflow records 1. JAWRA J Am Water Resour Associat 35(2):411–424

Arnold JG, Allen PM, Muttiah RGB (1995) Automated base flow separation and recession analysis techniques. Ground Water. https://doi.org/10.1111/j.1745-6584.1995.tb00046.x

Bati HG, TAT, TAA (2022) Performance assessment of interpolation techniques for optimal areal rainfall—temperature estimation: the case of two contrasting river catchments, Akaki and Mille, in Ethiopia. J Water Climate Change. https://doi.org/10.2166/wcc.2022.089

Chatterjee U (2018) Water scarcity in semi-arid regions of Bankura district, west Bengal, India—water scarcity in semi-arid regions of Bankura district, west Bengal, India—problems and prospects. Khoj an Internati Peer Revie Jour of Geog. https://doi.org/10.5958/2455-6963.2018.00007.3

Chiverton A, Hannaford J, Holman I, Corstanje R, Prudhomme C, Bloomfield J, Hess TM (2015) Which catchment characteristics control the temporal dependence structure of daily river flows ? Hydrol Process 1369(2014):1353–1369. https://doi.org/10.1002/hyp.10252

DeFries R, Eshleman KN (2004) Land-use change and hydrologic processes: a major focus for the future. Hydrol Process. https://doi.org/10.1002/hyp.5584

Ersi C, Bayaer T, Bao Y, Bao Y, Yong M (2022) Temporal and spatial changes in evapotranspiration and its potential driving factors in mongolia over the past 20 years. Remote Sens. https://doi.org/10.3390/rs14081856

Fekadu K (2015) Ethiopian seasonal rainfall variability and prediction using canonical correlation analysis. Earth 4(3):112–119. https://doi.org/10.11648/j.earth.20150403.14

Garg V, Aggarwal SP, Gupta PK, Nikam BR, Thakur PK, Srivastav SK, Senthil Kumar A (2017) Assessment of land use land cover change impact on hydrological regime of a basin. Environ Earth Sci. https://doi.org/10.1007/s12665-017-6976-z

Goovaerts P (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J Hydrol 228(1–2):113–129. https://doi.org/10.1016/S0022-1694(00)00144-X

Gudeta H (2022) Digital Elevation Model 's spatial resolution, and its influence on the accuracy of spatial Rainfall-Temperature prediction at catchment

scale: The case of Mille catchment, Ethiopia. https://doi.org/10.26491/mhwm/149231
Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. Appl Eng Agric 1:96–99
Hatfield JL, Prueger JH (2014) Spatial and temporal variation in evapotranspiration. In: Gerosa G (ed) Evapotranspiration—from measurements to agricultural and environmental applications. InTech, London
Helsel DR, Hirsch RM, Ryberg KR, Archfield SA, Gilroy EJ (2020) Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chap. A3, p 458. https://doi.org/10.3133/tm4a3. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chap. A3, version 1.1.]
Ji L, Peters AJ (2004) International journal of remote a spatial regression procedure for evaluating the relationship between AVHRR-NDVI and climate in the northern great plains. Int J Remote Sens. https://doi.org/10.1080/0143116031000102548
Kassas M (1976) Arid and semi-arid lands: problems and prospects. Agro-Ecosystems 3:185–204
Kittel TGF, Chase TN, Pielke Sr. RA, Nemani RR, Running SW (2000) Simulated impacts of historical land cover changes on global climate in northern winter. Climate Dynamics 16:93–105. https://doi.org/10.1007/s003820050007
Kroll CN, Song P (2013) Impact of multicollinearity on small sample hydrologic regression models. Water Resour Res 49(6):3756–3769
Kroll C, Luz J, Allen B, Vogel RM (2004) Developing a watershed characteristics database to improve low streamflow prediction. J Hydrol Eng. https://doi.org/10.1061/(ASCE)1084-0699(2004)9:2(116)
Mao D, Cherkauer KA (2009) Impacts of land-use change on hydrologic responses in the Great Lakes region. J Hydrol 374(1–2):71–82. https://doi.org/10.1016/j.jhydrol.2009.06.016
Misra AK (2014) Climate change and challenges of water and food security. Int J Sustain Built Environ 3(1):153–165. https://doi.org/10.1016/j.ijsbe.2014.04.006
Mohamoud Y (2004) Comparison of hydrologic responses at different watershed scales Research Triangle Park, NC, USA: Office of Research and Development, United States Environmental Protection Agency. 1–81. http://www.epa.gov/athens/publications/reports/Mohamoud_600_R04_103_Comparison_hydrologic.pdf
[MoWIE 2009] Federal Democratic Republic of Ethiopia Ministry of Water, I. and E. (2009) Federal democratic republic of Ethiopia ministry of water resources Mille and Dirma integrated sub-watershed management study annex e: land cover and land use (final)
Neubert M, Rannow S (2014) Managing protected areas in central and eastern europe under climate change, vol 58. Springer, Netherlands, Dordrecht. https://doi.org/10.1007/978-94-007-7960-0
Pervez MS, Henebry GM (2015) Assessing the impacts of climate and land use and land cover change on the freshwater availability in the Brahmaputra River basin. J Hydrol Reg Stud 3:285–311. https://doi.org/10.1016/j.ejrh.2014.09.003
Price K (2011) Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: a review. Prog Phys Geogr 35(4):465–492. https://doi.org/10.1177/0309133311402714
Qiu Y, Fu B, Wang J, Chen L (2001) Soil moisture variation in relation to topography and land use in a hillslope catchment of the Loess Plateau. China J Hydrol 240(3–4):243–263. https://doi.org/10.1016/S0022-1694(00)00362-0
Tiefelsdorf M (2000) Modelling spatial processes: the identification and analysis of spatial relationships in regression residuals by means of moran's I. Springer-Verlag, Berlin
United Nations Framework Convention on Climate Change (UNFCCC) (2007) Climate change: impacts, vulnerabilities and adaptation in developing countries
Verstraeten WW, Veroustraete F, Feyen J (2008) Assessment of evapotranspiration and soil moisture content across different scales of observation. Sensors. https://doi.org/10.3390/s8010070
Viviroli D, Du HH, Messerli B, Meybeck M, Weingartner R (2007) Mountains of the world, water towers for humanity: typology, mapping, and global significance. Water Resour Res 43:1–13. https://doi.org/10.1029/2006WR005653
Webster R, Oliver MA (2015) Basic steps in geostatistics: the variogram and kriging (No. 11599). Springer International Publishing, Cham, Switzerland
Wei X, Zhang M (2010) Quantifying streamflow change caused by forest disturbance at a large spatial scale: a single watershed study. Water Resour Res 46(1):15. https://doi.org/10.1029/2010WR009250
Woods RA, Sivapalan M, Robinson JS (1997) Modeling the spatial variability of subsurface runoff using a topographic index. Water Resour Res 33(5):1061–1073
Xiao D, Shi Y, Brantley SL, Dibiase R, Davis K, Li L (2019) Stream fl ow generation from catchments of contrasting lithologies: the role of soil properties. Topogr Catchment Size Water Resour Res. https://doi.org/10.1029/2018WR023736
Ye X, Li X, Liu J, Xu C, Zhang Q (2014) Variation of reference evapotranspiration and its contributing climatic factors in the Poyang Lake catchment China. Hydrol Process 6162(2013):6151–6162. https://doi.org/10.1002/hyp.10117
Zhang H, Wang L (2021) Analysis of the variation in potential evapotranspiration and surface wet conditions in the Hancang River Basin China. Sci Reports. https://doi.org/10.1038/s41598-021-88162-2
Zhang X, Kang S, Zhang L, Liu J (2010) Spatial variation of climatology monthly crop reference evapotranspiration and sensitivity coefficients in Shiyang river basin of northwest China. Agric Water Manag 97(10):1506–1516. https://doi.org/10.1016/j.agwat.2010.05.004
Zhu Y, Day RL (2009) Regression modeling of streamflow, baseflow, and runoff using geographic information systems. J Environ Manage 90(2):946–953. https://doi.org/10.1016/j.jenvman.2008.02.011

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.