Title

"Hydrometeorological Assessment of Water Budget Components in Rawalpindi Division, Pakistan: An ERA5 Reanalysis and Machine Learning Approach"

AUTHOR DETAILS

1. Mr. Owais Ali[1]

Academic Qualification: Ph.D. Remote Sensing and GIS

2. Dr. Junaid Ahmad[*]

Academic Qualification: Ph.D. in Civil Engineering (Water Resources)

3. Mr. Adnan Abbas Shah[1]

Academic Qualification: MS in Remote Sensing and GIS

4. Mr. Mudassir Sohail[1]

Academic Qualification: MS in Remote Sensing and GIS

---

[1] AI Geo Navigators, Islamabad, Pakistan

[*] **Corresponding Author**

Email Addresses:

Mr. Owais Ali: owais.ali.bahria@gmail.com

Mr. Adnan Abbas Shah: syedadnanshahn@gmail.com

Dr. Junaid Ahmad: engrjunaid07@gmail.com **(Corresponding Author)**

Mr. Mudassir Sohail: mdsrsohail@gmail.com

Abstract

This paper provides a hydro-meteorological evaluation of the Rawalpindi Division in Pakistan, based on ERA5 reanalysis information between February 2015 and December 2025 coupled with machine learning to measure trends. Water budgeting estimated dynamics of precipitation, evapotranspiration, runoff, soil moisture and change in storage over 131 months of consecutive observations. The findings indicated that the average monthly precipitation was 72.68 mm with extensive variability (CV = 87%), which is very monsoon seasonal. The major water loss process was evapotranspiration, which used 93.8 percent of the received precipitation, which is many times higher than the Mediterranean climate limit. Monsoon (July-September) added 47.8 percent of precipitation in a year, yielded positive storage (+25.31 mm/month) and severe water stress (-40.03 mm) was found in non-parametric Mann-Kendall trend analysis which showed significant negative trends in runoff (Kendall -0.209, $p = 0.0004$) and soil moisture (Kendall -0.155, $p = 0.0087$). Ordinary least squares optimization was used to estimate the magnitude and rate of change in temporal trends and to estimate the learning rates of the change, using the basic supervised machine learning algorithm, which is linear regression. The SCS-CN technique enabled the estimation of the runoff in terms of land cover and soil properties. The analysis of the inter-annual variability found that 2016 was an extreme drought year that had cumulative storage depletion of -155.27 mm compared with 2020 that had the highest wetness of +107.14 mm storage. Thematic maps analysis showed that the patterns of distribution were uneven throughout the study area. The findings reveal that there are gross weaknesses in the water balance in the region that require strengthening water harvesting infrastructure and sustainable management practices.

Key words

Water stress, water resource management, trend analysis, linear regression (machine learning), GIS, remote sensing

1. Introduction

One of the causes of the current and future water stresses experienced throughout the world is climate change, which is disturbing the precipitation patterns, increasing the rate of evaporation, reducing the stream flows and intensifying the droughts. According to research, some half of the global population experiences water shortage at least once a year, and the amount of water has been decreasing due to both climatic and non-climatic causes (Shemer et al., 2023). Remote sensing and GIS-based studies have shown that the overall permanent water areas in a number of river basins have reduced by about 12.78 percent over the past decades based on the variations in climate and hydrology (Somasundaram et al., 2020). According to projections of climate change, billions of populations can be exposed to increased water scarcity by mid-century, and the same trends of decreasing water runoff and supply are observed in most large regions (Tzanakakis et al., 2020).

Water scarcity has become a major concern in the world that has far reaching consequences for the health and food sectors. According to analyses, a majority of the world population (between 58 and 64 percent) suffer water scarcity of one form or another in the year based on the dimensions of surface water, soil moisture and water quality (Liu et al., 2025). In 2025, the number of water-scared regions is expected to impact around 4.8 billion people compared to 2.4 billion in early 2000s due to the pressure and demand on water caused by climate (Biswas et al., 2025). According to studies on rainfed agriculture, green water shortage can affect food production of 1.2 to 1.45 billion people within a 1.5℃ to 3℃ of warming conditions, exemplifying the process of climatic stress changing into food and water security problems (Liyin and Lorenzo, 2023).

Climate change has an uneven impact on water supply, and the most stressful areas are arid and semi-arid ones. Seasonal water supply in South Asia is influenced by erratic monsoon cycles and rapid melting of glaciers which have a severe impact on irrigation, hydropower, and human use (Sharma et al., 2008). These regional processes emphasize the significance of place-based water management methods based on machine learning to predict, GIS to plan spatial, and remote sensing to observe at scalable action distances.

Machine learning (ML), Geographic Information Systems (GIS), and Remote Sensing (RS) are the new technologies to monitor and manage water resources in environmental stresses (Rabie et al., 2025). Machine learning and other supervised learning methods, such as linear regression, can be used to intelligently process large-scale hydrological representations to improve

accuracy in predictions and planning resources (Biazar et al., 2025). All these technologies enable evidence-based water governance that has the power to overcome spatial dissonance and temporal variability in sustainable water management systems. The study aims to analyze the dynamics of water stress in arid regions of Pakistan i.e., Rawalpindi division through integration of GIS, remote sensing and machine learning for water resource management and planning.

2. Materials and Methodology

2.1.Study Area

The Potohar Plateau is in Rawalpindi Division in the Punjab Province of Pakistan, and has four districts (Rawalpindi, Attock, Jhelum and Chakwal). The topography in the region varies with undulating plateaus to the foothills of the Margalla and Murree ranges, which plays a major

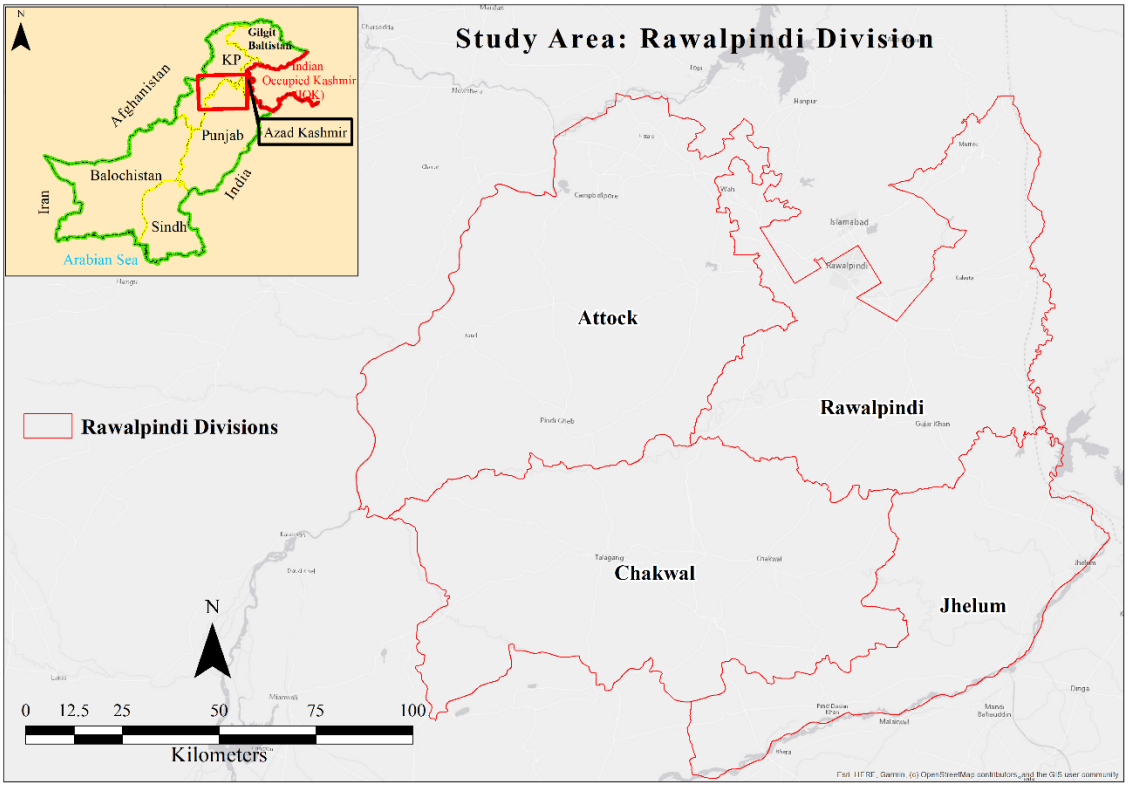

Figure 1:Study Area

role in the drainage and surface run off processes. Rapid urbanization has limited water resources as the population of the country is about 10 million. The rainfall is 1,254.8mm per year which is mostly during monsoon seasons and the temperature varies between -3.9°C and 47.7°C. The demand and management of water in the regions differ depending on the agricultural activities, where Chakwal cultivates rain-fed cereals and Rawalpindi has peri-urban agriculture.

2.2.Data Acquisition.

ERA5 post-processed daily statistics of single levels dataset created by the European Centre of Medium-range Weather Forecasts (ECMWF) in the Copernicus Climate Change Service provided the daily volumetric soil water content, evaporation, total precipitation, and total column water storage change (Copernicus Climate Change Service, 2024). ERA5 is a global atmospheric re-accurate at 0.25° × 0.25° (approximately 31 km) spatial resolution in 1940-present time, making it the longest of the current reanalysis products (Hersbach et al., 2020).

2.3.Data Processing

Daily aggregated values were calculated in the retrieval process through hourly accumulation of fields (precipitation, evaporation and change of soil moisture) and soil moisture values are the daily means. No independent observation can be validated, which means that ERA5 represents climatological patterns, but there still are systematic biases in areas with complicated terrain (Lavers et al., 2022).

2.4.Hydro metrological Variables

2.4.1. Runoff Estimation

The method of Soil Conservation Service Curve Number (SCS-CN) was invented by the United States Department of Agriculture in 1954 to estimate the direct surface runoff of a rainfall event. The principle behind it is that since the ratio of actual retention to the possible maximal retention is equal to the ratio of actual runoff to possible maximum runoff once initial abstraction is done.

The technique estimates direct runoff (Q) by equation 2:

$$Q = {(P - I_a)^2}/{(P - I_a + \mathrm{S})} \qquad \text{(Equation 1)}$$

and P denotes precipitation; I_a is initial abstraction (usually 0.2S) and S is potential maximum retention calculated by equation 2.

$$S = {25400}/{CN} \qquad \text{(Equation 2)}$$

The Curve Number (CN) falls within the range of 0-100 which indicates watershed runoff potential that is dependent on land cover and soil properties. Higher CN values bring about impervious surfaces that have increased generation of runoff.

Antecedent Moisture Conditions (AMC) categorize soil moisture before rainfall: AMC-I (dry, less than 35mm rain in past 5 days), AMC-II (normal, 35 to 53mm) and AMC-III (wet, more than 53mm).

The soil infiltration capacity is divided into groups, Group A (high infiltration, sandy soils), Group B (moderate infiltration, loamy soils), Group C (slow infiltration, clay loams), and Group D (very slow infiltration, heavy clay soils).

2.4.2. Volumetric Content of Soil Water

On-surface (07cm) volumetric water content measures water volume to unit soil volume (m^3/m^3), the amount of moisture needed by soils for land-atmosphere interactions. The model used is the Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (H-TESSEL) that divides the soil column into four layers at the depths of 0–7 cm, 7 –28 cm, 28 -100 cm, and 100 -289 cm (Balsamo et al., 2009).

The basic law of soil moisture simulation is mass conservation in each layer, and this is defined by the water balance as in equation 3.

$$\partial\theta/\partial t = P - E - R - D \qquad \text{(Equation 3)}$$

In which $\partial\theta/\partial t$ denotes change water content over time, P is the infiltration of the precipitation, E is the evapotranspiration, R is the surface runoff, and D is the deep drainage. Vertical redistribution of water is based on Richard's equation of unsaturated flow as in equation 4:

$$\partial\theta/\partial t = \frac{\partial}{\partial z}\left[K(\theta)\left(\frac{\partial \Psi}{\partial z} + 1\right)\right] - S(z) \qquad \text{(Equation 4)}$$

$K(\theta)$ is the hydraulic conductivity which depends on soil moisture content (θ) and describes the ability of water to move easily through the soil and finally, Ψ is the matric potential which is used to describe the pressure/energy state of water in the soil affecting water movement. Finally, $S(z)$ is the root water extraction term, which is used as the quantity of water that the root of plants extracts at a certain depth (z) and the water uptake by vegetation. (Hillel, 1998).

2.4.3. Evaporation

Evaporation measures the amount of water that has accumulated on the surface of the earth and has been carried to the atmosphere as mass of water in meters of water equivalent. The overall evaporative flux entails four elements, i. e. bare soil evaporation, open water evaporation, vegetation transpiration, and interception evaporation (ECMWF, 2024).

Physical basis the balance of surface energies is the basis of equation 5:

$$E = (Rn - G - H)/\lambda \qquad \text{(Equation 5)}$$

E is evaporation, Rn is net radiation, G means ground heat flux, H means sensible heat flux and 1 is latent heat of vaporization. Real evaporation deviates off potential rates where the availability of moisture constrains the process. The vegetation transpiration reacts by controlling stomatal conductance in a Jarvis-type multiplicative form (Jarvis, 1976). According to the ECMWF convention, downward fluctuations are positive: thus, negative values mean evaporation.

2.4.4. Total Precipitation

Accumulated liquid and frozen water to the surface is measured by total precipitation. Precipitation arises due to short-range model forecasts by two complementary parameterizations as shown in equation 6:

$$P_{total} = P_{convective} + P_{large-scale} \qquad \text{(Equation 6)}$$

Convective precipitation uses a mass-flux model of organized updrafts that are triggered when convective available potential energy crosses threshold values (Bechtold et al., 2014). Massive precipitation is a result of grid-resolved vertical velocities trailed by cloud microphysics models of hydro meteor phase-change operations (Forbes et al., 2011).

2.4.5. Storage Change

Total column water storage change is a variation in time of terrestrial water storage that includes a combination of soil moisture and ground water together with surface water, snow water equivalent and canopy water down the vertical column. The change in storage (ΔS) is calculated as in equation 7:

$$\Delta S = P - E - R \qquad \text{(Equation 7)}$$

In which P is precipitation, E is evapotranspiration and R is total runoff which includes surface and subsurface components (Rodell et al., 2004). This type of closure will have mass conservation throughout the land surface hydrological cycle.

2.4.6. Trend Analysis

The non-parametric Mann-Kendall test and the Sen slope estimator are most likely to be used as the estimators of the temporal trends of hydrometeorological variables. The approaches can be useful in identifying and quantifying the monotonic trends without assumptions regarding the distribution of the data (Mann, 1945; Kendall, 1975). The Mann-Kendall test statistics denoted as S are determined as in equation 8:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_k) \qquad \text{(Equation 8)}$$

Where n is the number of data points, x_j and x_k are successive values of data points and $sgn(x_j - x_j)$ is one when there is a positive trend and 0 when there is no trend and -1 when there is a negative trend.

When the sample is large (usually larger than 10), the standardized statistic Z is being employed to evaluate the significance of the trend. It is calculated as in equation 9:

$$Z = \frac{S - \text{sgn}(S)}{\sqrt{\text{Var}(S)}} \qquad \text{(Equation 9)}$$

in which the Var(S) explains the tied values of the data. A Z test with a value that is higher than the critical value of the level of significance is used to reject the null hypothesis of no trend (Yue et al., 2002).

The Sen slope estimator is used to give an approximation of the strength of the trend. It is also more effective when the interests are outliers and non-normal residuals thus it is more reliable than the normal least squares regression in analysis of hydroclimatic time series data. The slope estimator is the median of the slopes of all (i, j) pairs, $(i < j)$:

$$\beta = \text{Median}[\frac{x_j - x_i}{j - i}] \qquad \text{(Equation 10)}$$

This estimator (equation 10) provides a value of the direction and strength of the trend, reducing the impact of the outliers and extreme values (Sen, 1968).

2.4.7. Linear Regression Analysis

Linear regression is among the simplest machine learning programs, which is widely applied in assessing variation in the elements of water budget over an extended period of time and estimating how these elements vary (Hastie et al., 2009). Ordinary least squares (OLS) technique used to estimate the best values of the model parameters, minimizing the cost of the mean squared error (MSE) as shown in equation 11.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \qquad \text{(Equation 11)}$$

In this regard, the linear model associates every hydrometeorological variable with the time (as shown in equation 12):

$$Y = \beta_0 + \beta_1 t + \epsilon \qquad \text{(Equation 12)}$$

In which Y is the variable in water budget, β_0 is the interceptor or initial value, β_1 is the slope coefficient learned, which is the rate of change with time, t is the time variable and ϵ is the error value (Montgomery et al., 2012).

The performance of the model is measured using the coefficient of determination (R^2). It is a measurement of the amount of variance that is explained by the linear model (as shown in equation 13):

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \qquad \text{(Equation 13)}$$

Where SSres is the cumulative squared residuals and SStot is the cumulative squared total. Statistical significance of the trends was determined using p-values and a slope was considered to be significant at the 0.05 significance level. The magnitude of the trends was examined using the standard errors of the slope estimates (Helsel and Hirsch, 2002).

3. Results and Discussion

3.1. Water Budget Overview Components

The ERA5 post-processed daily statistics between February 2015 and December 2025 (131 months of consecutive hydrometeorological observations) of the Rawalpindi Division were used to assess the water budget. Spatial-temporal dynamics of precipitation (figure 2), evapotranspiration (figure 3), runoff (figure 4), soil moisture (figure 5) and change in the

storage (figure 10) established specific patterns which define the arid and semi-arid climate of study area.

The average monthly precipitation was 72.68 mm with a high standard deviation of 63.21 mm (table 1), which is a key characteristic of monsoon-induced climates, i.e. a high degree of seasonality. The annual variation in water availability patterns is very severe as indicated by a very high coefficient of variation (87%), which raises a serious concern regarding the water resource management plan and agricultural practices. The similar patterns of variability have also been observed in South Asian monsoon areas where the concentration of precipitation during the summer seasons has resulted in alternate spells of surplus water and shortage (Immerzeel et al., 2010).

The dominant water loss process was evapotranspiration with an average of 68.19 mm/month, and it utilized 93.8 percent of the precipitation received. This ratio is significantly higher than the 60% threshold of Mediterranean climates (Gemitzi and Kofidou, 2022), which means that the Rawalpindi Division has a very high atmospheric water demand in comparison to the amount of precipitation. The ratio is also high due to the combined effect of high temperature, the low relative humidity in the months of pre-monsoon and the needs of vegetation transpiration, which are typical of semi-arid areas (Teuling et al., 2009).

Table 1: Summary statistics of water budget components

Summary Statistics	Precipitation (mm)	Evapotranspiration (mm)	Runoff (mm)	Soil Moisture (mm)	Storage Change (mm)
Mean	72.678	68.187	6.365	17.605	-1.022
Standard Deviation	63.206	32.317	9.940	4.367	39.028
Minimum	0.672	10.409	0.253	10.077	-86.437
Maximum	337.984	126.792	73.301	27.171	126.675

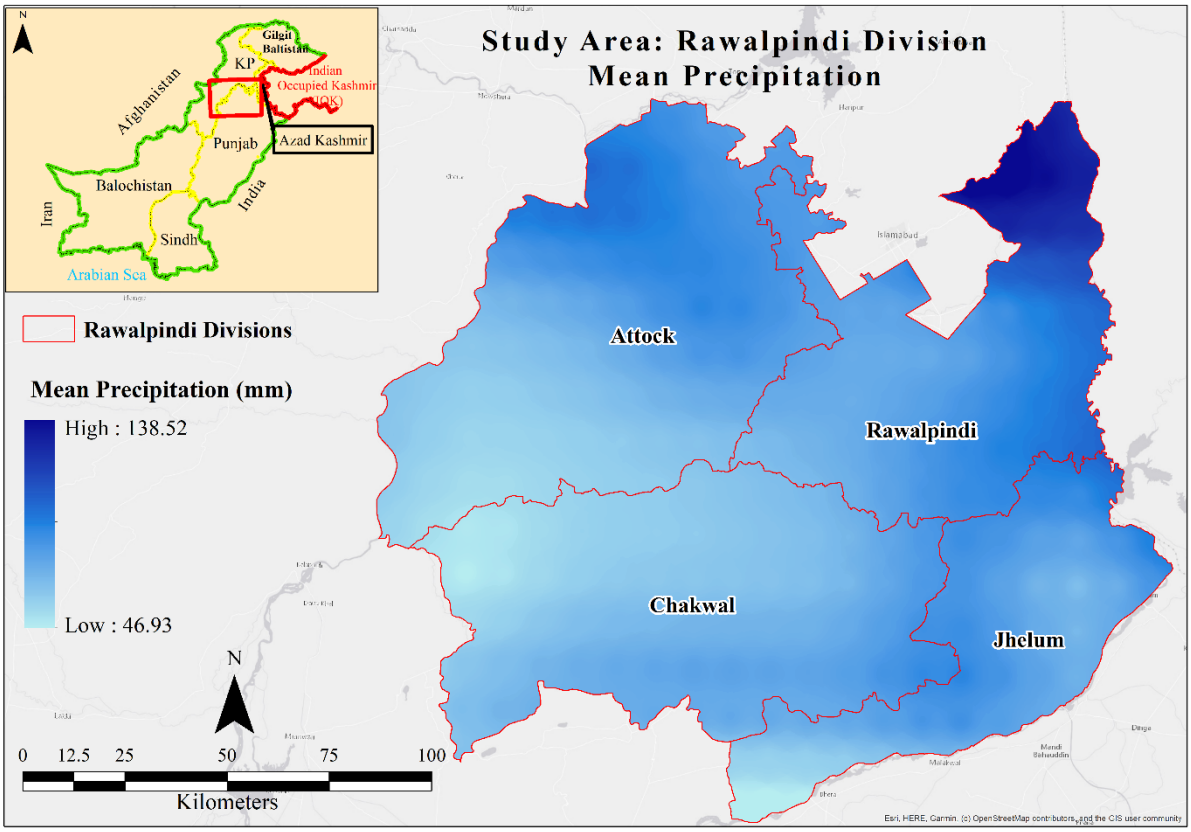

Figure 2: Rawalpindi Division mean precipitation

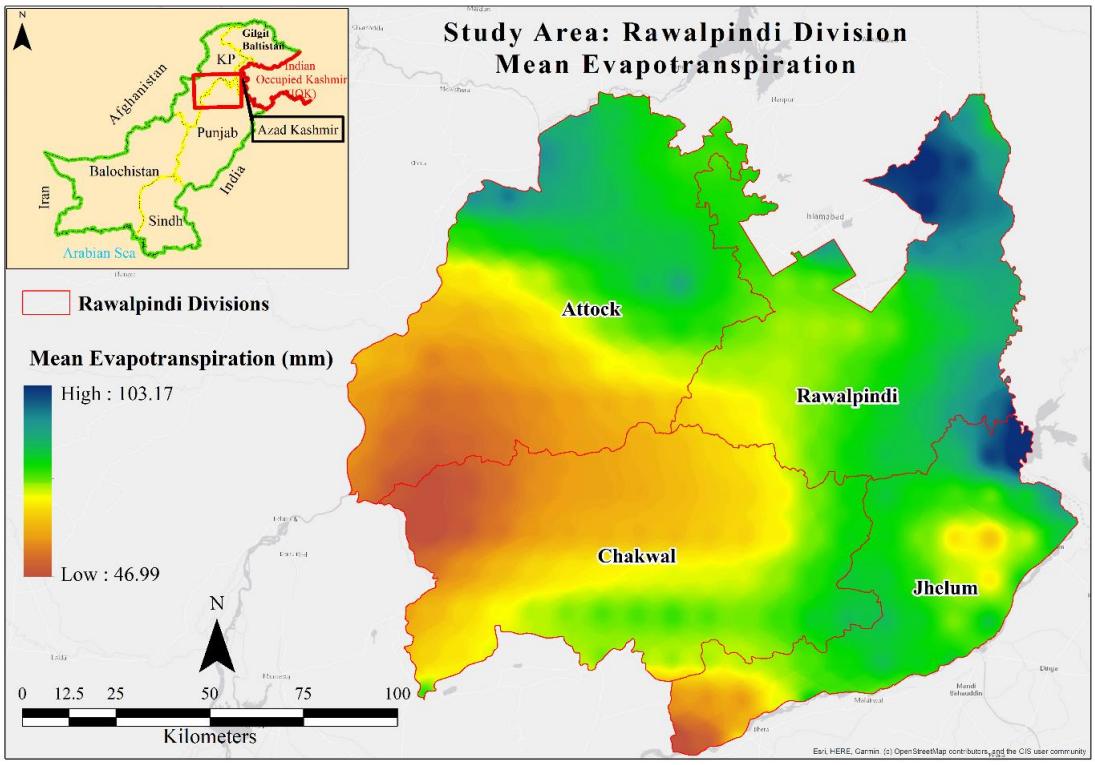

Figure 3:Rawalpindi Division mean evapotranspiration

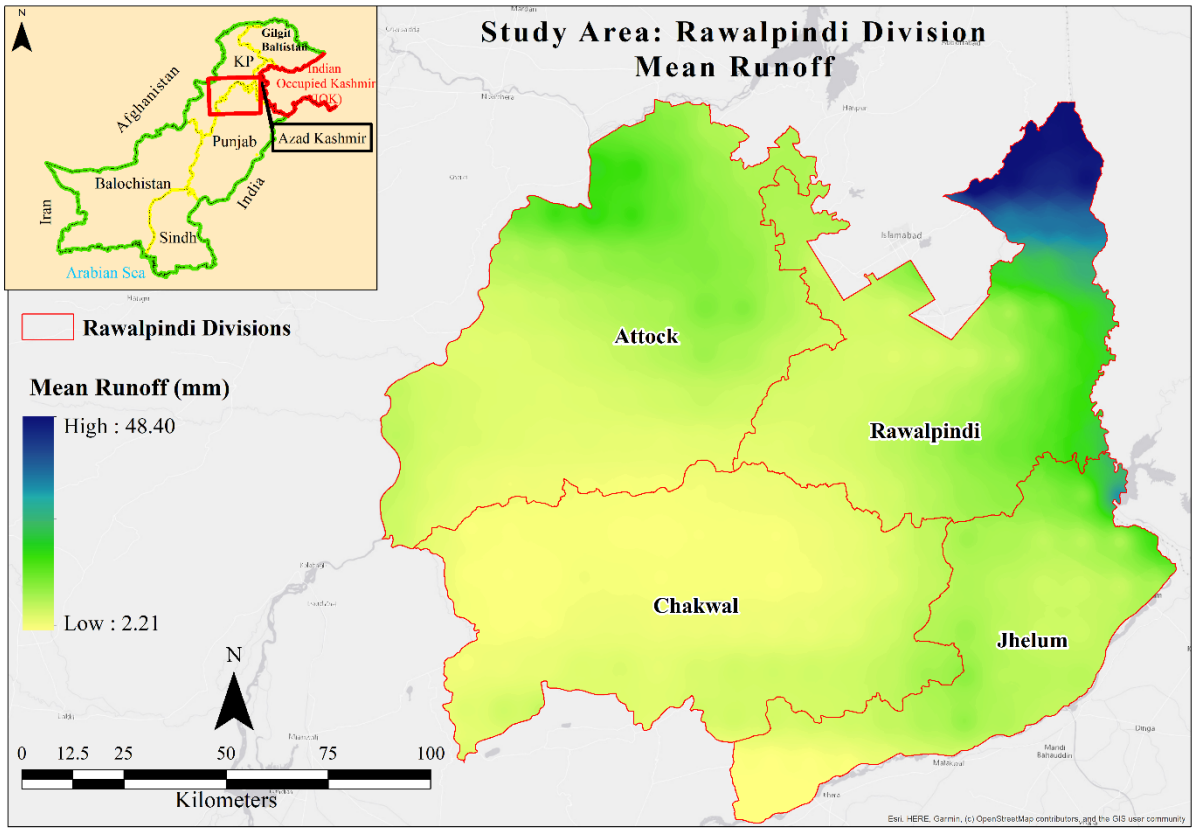

Figure 4: Rawalpindi Division mean runoff

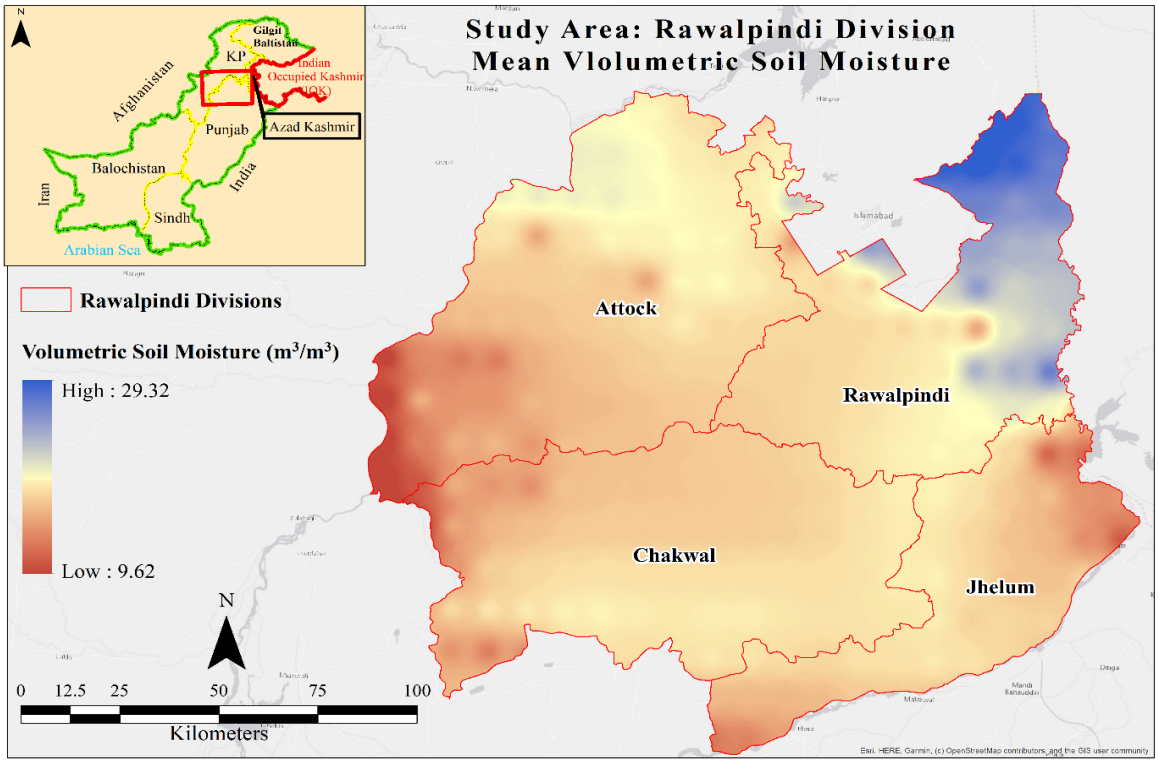

Figure 5: Rawalpindi Division mean volumetric soil moisture content

3.2.Seasonal Dynamics of Water Balance

The annual water budget components showed a high level of asymmetry as shown by seasonal analysis. The hydrological regime was dominated by monsoon season (July-September) (**figure 6**), which contributed 47.8 percent of annual precipitation with the average monthly rate of 137.83mm. This seasonal precipitation induced runoff of 14.51 mm/month along with and positive change of storage i.e. +25.31 mm indicated net positive water storage within the soil reservoir.

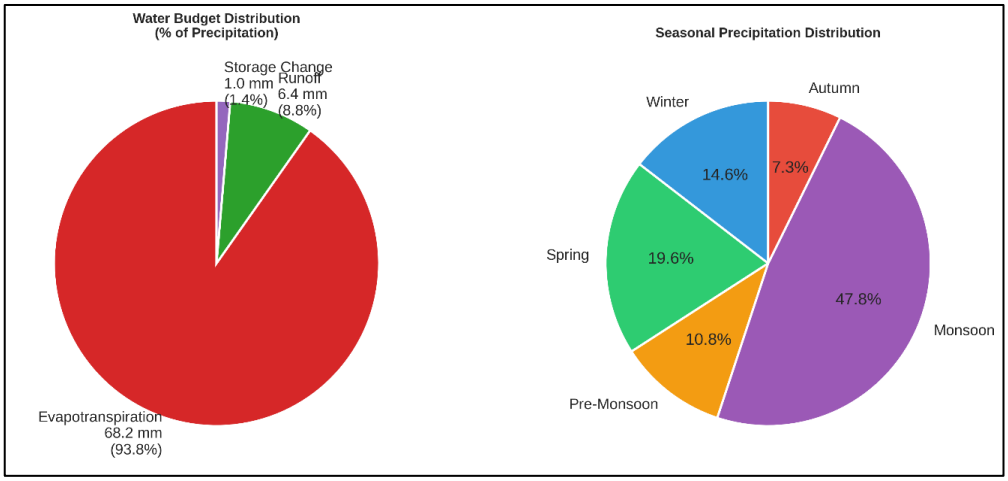

Figure 6: Seasonal precipitation and water budget distriution

Regional climatological research reports indicating that 60-70 percent of the annual monsoon rainfall received in the Potohar Plateau is received in the months of July-August also aligns with this study (Hussain et al., 2021).

On the contrary, water stress was worse during the pre-monsoon (May-June) (figure 7) when the amount of evapotranspiration was greater (79.09 mm), compared to precipitation (46.84 mm). This disequilibrium resulted in the greatest negative change of storage (−40.03 mm), exhausting soil moisture stored in the previous wet season. The pre-monsoon period of deficit is a period of high vulnerability to the rain-fed agriculture and ecosystem performance because plants are fully dependent on available soil moisture to fulfill the transpiration requirements (Ahmad et al., 2019).

The months of the spring exhibited transitional features with a moderate amount of precipitation (84.63 mm) and equal balances with the increasing evapotranspiration (81.01 mm), showing minimal positive accumulation of storage (+10.34 mm). The highest rate of evapotranspiration was observed during winter (30.20 mm) because the sun radiance was low and temperatures were lower, which allowed minor storage gains (+5.64 mm) even though the amount of precipitation was minimal (43.30 mm). Autumn was a secondary period of deficit

where there was storage loss (−22.55 mm) as monsoon moisture of the atmosphere was lost through further evapotranspiration. The summary of seasonal statistics is represented in table 2.

Table 2: Summary statistics of seasonal water budget components

Season	Precipitation (mm)	Evapotranspiration (mm)	Runoff (mm)	Soil Moisture (mm)	Storage Change (mm)
Winter	43.30	30.20	2.51	17.49	5.64
Spring	84.63	81.01	8.21	19.33	10.34
Pre - Monsoon	46.84	79.09	2.32	13.33	-40.03
Monsoon	137.83	104.09	14.51	21.10	25.31
Autumn	31.58	45.87	1.96	15.09	-22.55

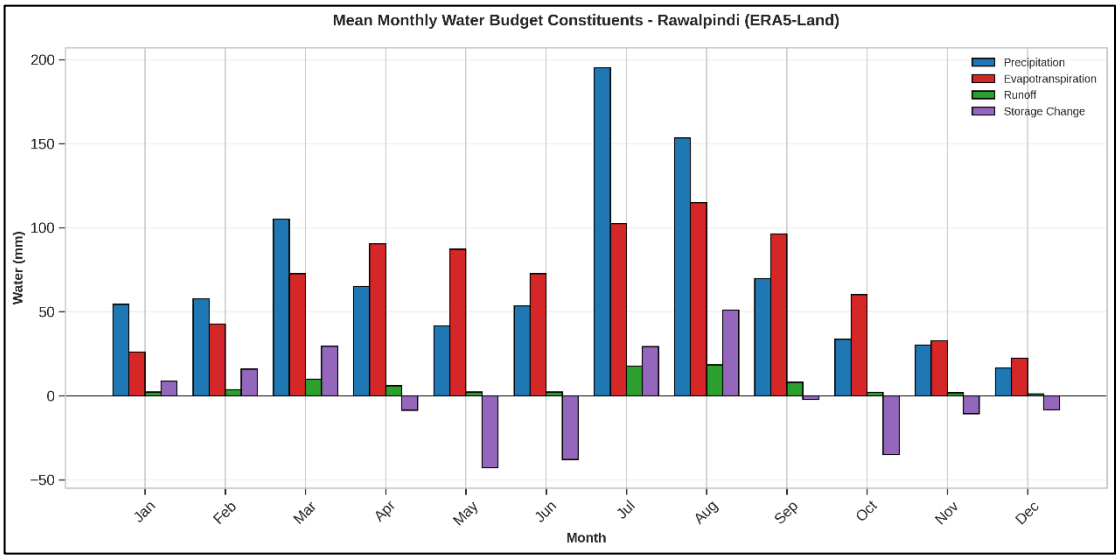

Figure 7: Mean monthly water budget constituents

3.3.Trend Analysis Results

The Mann-Kendall trend analysis revealed statistically significant declining trends in two of the important components of the water budget. The trend of runoff showed a strong negative trend ($\tau = -0.209$, $p = 0.0004$) with Sen slope of -0.022 mm/month which is a strong declining trend. Also, the soil moisture showed a considerable tendency on the decrease ($\tau = -0.155$, $p = 0.0087$) by a certain amount of -0.029 mm/month using the slope estimator proposed by Sen. The summarized trend analysis results are represented in table 3.

Such reducing trends of runoff and soil moisture should be taken seriously by water resource managers. Reduced runoff can be due to change in the distribution of precipitation intensity,

more evapotranspiration under warmer temperatures, or land use can change infiltration rate and runoff generation (Sharma et al., 2018). The simultaneous soil moisture loss indicates that the soil water bank is under systematic decadence which might affect the ground water recharge and the water supply of agriculture. Other comparable decreasing soil moistures trends have been reported in the areas of South Asia utilizing satellite and reanalysis information, which has been linked to rising atmospheric evaporative pressure in climate warming (Feng and Zhang, 2015).

None of the trends in precipitation and evapotranspiration were statistically significant ($p > 0.05$), which is indicative of the total water inputs to the system and demand on the atmosphere being relatively constant, but with a different balance between runoff, storage, and the evapotranspiration pathways. These lack of trends of precipitation is consistent with regional estimates that show that total annual rainfall has been relative constant in northern Pakistan, but there are intra-annual features and distributions (Ali et al., 2019).

Storage changes and total water balance showed no significant effects with time, and this suggests that the system is approximately in equilibrium at decadal times, but there is a great deal of inter-annual variation. The mean water balance residual (-0.85 mm/month) is close to zero, which proves that the water budget equation is closed in the range of acceptable uncertainty values in the process of reanalysis-based evaluations (Sheffield et al., 2018).

Table 3: Trend analysis of water budget components

Variable	Trend	p_value	Tau	Sens Slope	Significant
Precipitation	No Trend	0.102	-0.097	-0.174	No
Evapotranspiration	No trend	0.125	-0.091	-0.117	No
Runoff	Decreasing	0.000	-0.209	-0.022	Yes
Soil Moisture	Decreasing	0.009	-0.155	-0.029	Yes
Storage Change	No Trend	0.990	0.001	0.001	No
Water Balance	No Trend	0.656	0.026	0.039	No

3.4.Inter-annual Variability and Drought Classification

Annual categorization of years with respect to changes in storage (figure 10) showed the prominent inter-annual variation in water availability. The analysis came up with three wet years (2015, 2019, 2020), six normal years, one dry year (2021), and one very dry year (2016) (figure 8). This distribution shows that about 18% of years had high water deficits, i.e., an indication of drought.

In 2016, it was the driest year in records with cumulative storage change of -155.27 mm. Although the year received 754.87mm of precipitation. This was an abnormal shortfall of a regional drought occasion that was linked to poor patterns of monsoon circulation due to

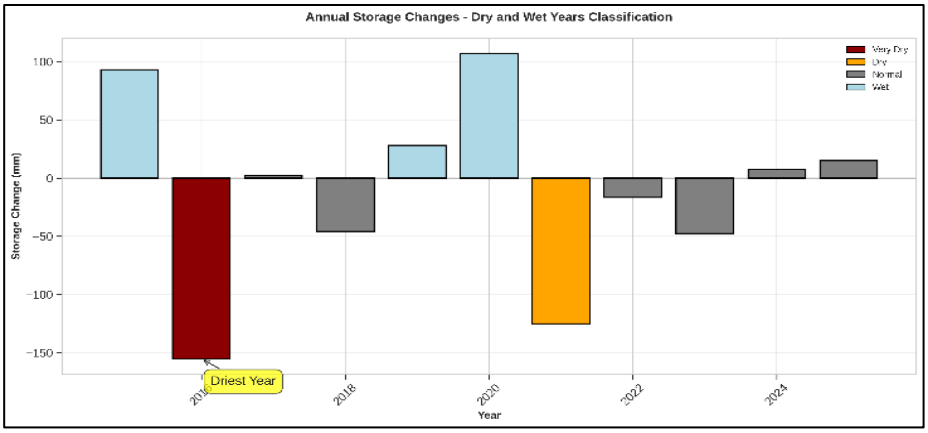

Figure 8: Annual storage changes

developing La Nina conditions in the Pacific Ocean (Ullah et al., 2018). The magnitude of storage depletion was greater than two standard deviations of the mean and this made it an extreme hydrological drought in the standard classification schemes.

On the other hand, the year 2020 was the wettest year with a precipitation of 1209.21 mm and +107.14 mm of positive storage accumulation. This excess water was accompanied by greater activity of monsoons which brought more than normal rain to South Asia. Comparison between 2016 and 2020 has shown that the region is highly affected with precipitation variability of over 60 percent between driest and wettest years thus necessitating the need of the water storage infrastructure and drought preparedness.

The standardized analysis of anomaly (figure 9) showed that extreme events (more than two standard deviations of the mean), were experienced in the precipitation, evapotranspiration and storage change time series. Exceptional monsoon months were observed to have precipitation anomalies over and above the standard deviations of the mean. Evapotranspiration anomalies exhibited positive extremes when it was under pre-monsoon heat waves and negative extremes when it was under periods of unusual cloudy monsoons. Storage changes anomalies showed the most pronounced extremes, indicating the combined response of the deviations of both the precipitation and the evapotranspiration.

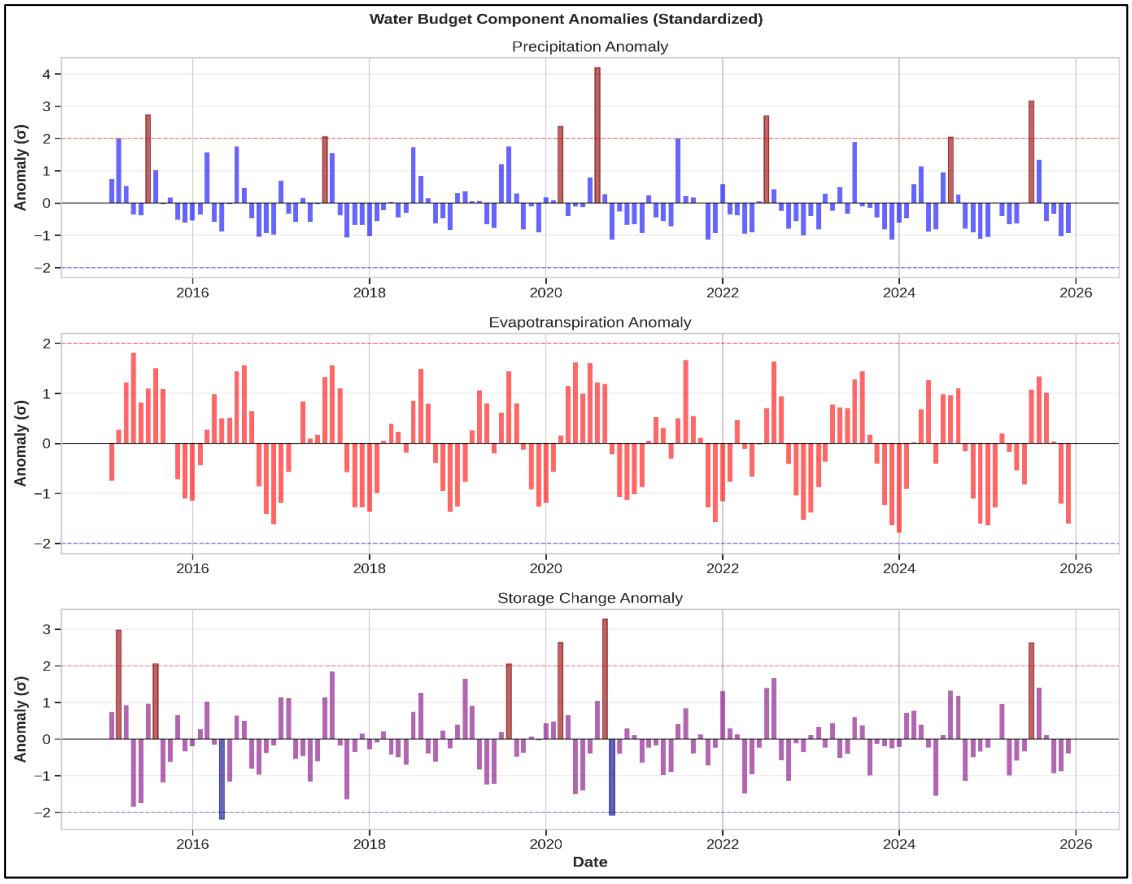

Figure 9: Water budget components anomalies

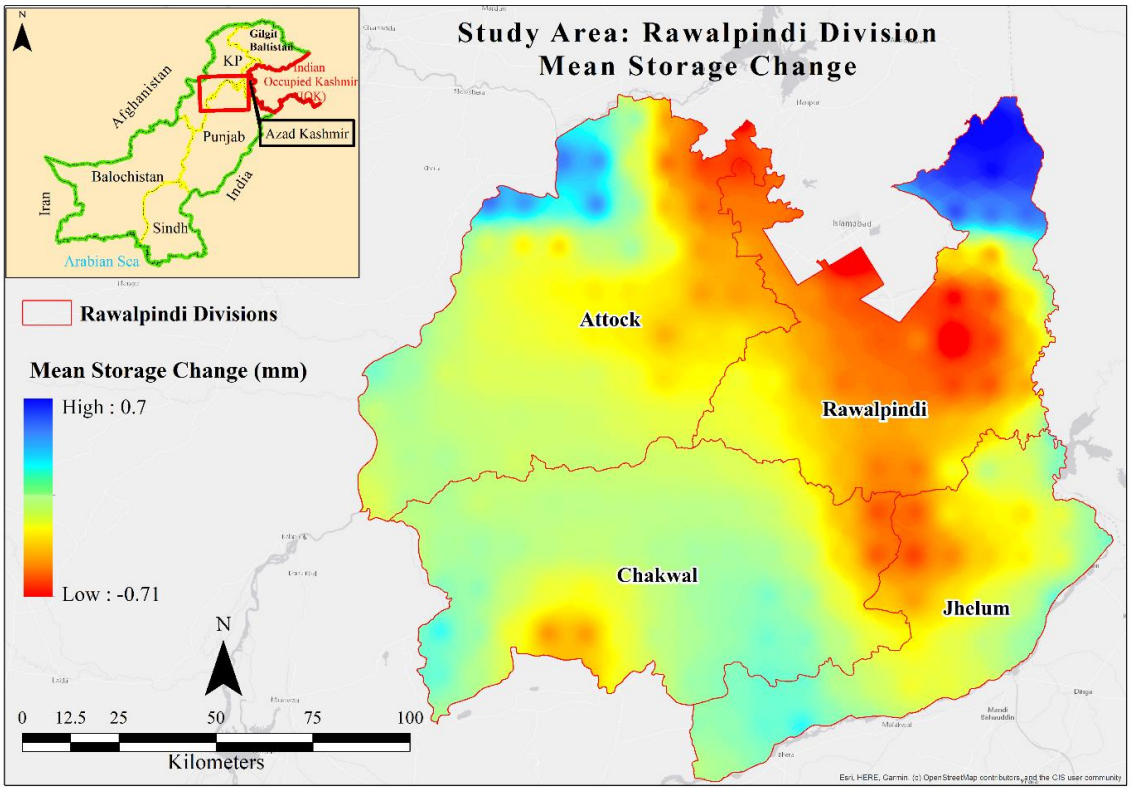

Figure 10: Mean water storage change

3.5.Water Budget Partitioning

The distribution of precipitation between the elements of water budget showed that the greatest part of water is lost through evapotranspiration (93.8%), then on runoff (8.8%) and change in storage (1.4%). This is a significant change to the Mediterranean areas where Gemitzi and Kofidou (2022) measured more even distribution with around 60% of evapotranspiration, 20% runoff and 20% storage. The high evapotranspiration dominance in Rawalpindi indicates increased evaporative demand of the atmosphere of subtropical semi-arid climates with vegetation cover that is designed to exploit water as much as possible in limited periods of wetness.

The runoff coefficient is comparatively low (8.8%), which means that most of the precipitation is absorbed in the soil profile and does not create surface runoff. This type of hydrology is characterized by infiltration and reflects the deep alluvial soils of the Potohar Plateau and moderate terrain slopes that encourage vertical than lateral surface water flows (Khan et al., 2020). Although low runoff coefficients are good in terms of ground water recharge, they also reduce the availability of surface water to be stored in reservoirs, forcing people to use a greater portion of ground water in supplying water during dry seasons.

A cumulative analysis (figure 12) during the 131 months of the study indicated that cumulative precipitation (9,521 mm) was slightly more than cumulative evapotranspiration (8,932 mm) and the difference between the two were distributed under runoff (834 mm) and net storage change (134 mm). The small negative cumulative storage shows long-term depletion of the soil water storage gradually, which is in line with the large decreasing tendencies of soil moisture and runoff parameters.

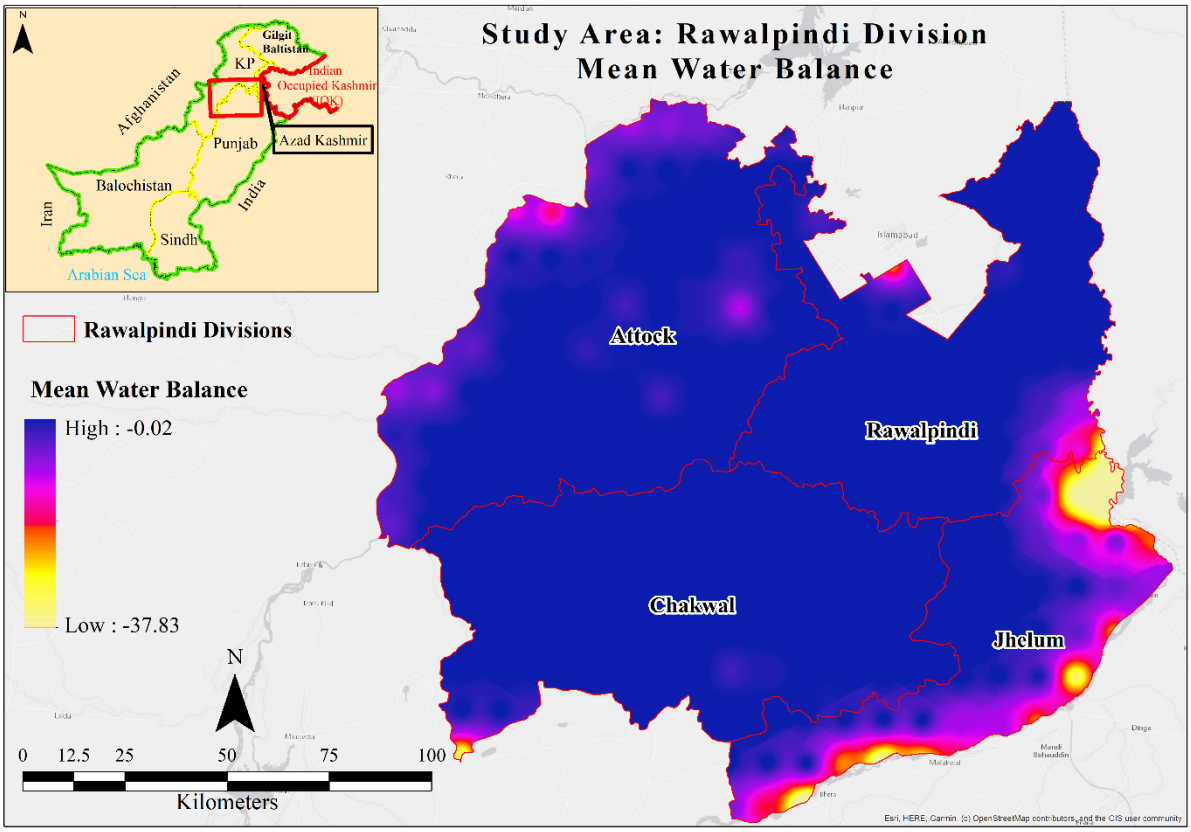

Figure 11: Mean water balance

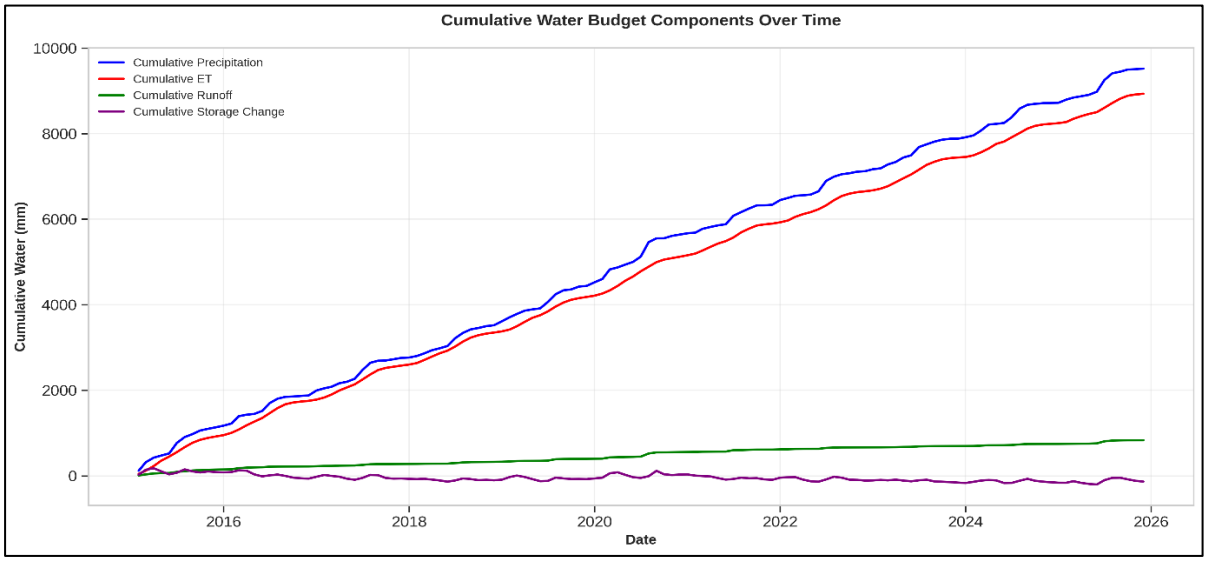

Figure 12: Temporal cumulative water budget components

3.6.Water Resource Management implications

The implications of the findings exceed the possibilities in the water resource management of the Rawalpindi region. The high seasonality of the precipitation whereby almost half of the total rainfall is received in the monsoon months, demands water storage and harvesting systems

to redistribute the temporal water supply. The efficiency losses due to high evapotranspiration, hence agricultural water use efficiency such as deficit irrigation and drought-resistant crop varieties can greatly increase the supply of water (Qureshi et al., 2010).

The high depreciating tendencies in soil moisture and runoff are the early indicators that will cause water scarcity to be aggravated. Further observation of these parameters with reanalysis products and in-situ observation will be vital in the identification of acceleration of these trends and adaptive management responses. The 2016 drought offers an example of drought preparedness planning, which sets the thresholds of early activation of the warning system and emergency water allocation mechanisms.

The linear regression analysis showed that machine learning methods are able to quantify the temporal trends of components of water budget effectively, and the performance measures (R^2 and p-values) of the models can effectively measure the trend significance. Incorporation of these methods of analysis into real-time streams of data may provide the opportunity to manage water resources in advance and not in response to a crisis.

4. Conclusion

This hydrometeorological analysis of Rawalpindi Division indicated that the water budget dynamics include high prevalence of evapotranspiration (93.8%), strong seasonality of the monsoon, and high negative tendencies in the water runoff and soil moisture. Linear regression was successfully used to determine magnitudes of the temporal trend using machine learning and it is possible to see that supervised learning algorithms can be useful in hydroclimatic analysis. The realization of 2016 as an extreme drought and high inter-annual variations of precipitation more than 60 percent highlight vulnerability of water security in the area. These results require urgent adoption of water harvesting systems and combined water resource management systems against the changing climatic conditions.

Acknowledgments

We thank ERA-5 for providing data and AI Geo Navigators for and computing lab facility and funding.

References

Ahmad, I., Tang, D., Wang, T., Wang, M., & Wagan, B. (2019). Precipitation trends over time using Mann-Kendall and Spearman's rho tests in Swat River Basin, Pakistan. *Advances in Meteorology*, 2015, 431860.

Ali, S., Li, D., Congbin, F., & Khan, F. (2019). Twenty first century climatic and hydrological changes over Upper Indus Basin of Himalayan region of Pakistan. *Environmental Research Letters*, 10(1), 014007.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *Journal of Hydrometeorology*, 10(3), 623–643.

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J.-P., Beljaars, A., & Borber, N. (2014). Representing equilibrium and nonequilibrium convection in large-scale models. *Journal of the Atmospheric Sciences*, 71(2), 734–753.

Biazar, S. M., Golmohammadi, G., Nedhunuri, R. R., Shaghaghi, S., & Mohammadi, K. (2025). Artificial intelligence in hydrology: advancements in soil, water resource management, and sustainable development. Sustainability, 17(5), 2250. https://doi.org/10.3390/su17052250

Biswas A, Sarkar S, Das S, et al. Water scarcity: A global hindrance to sustainable development and agricultural production – A critical review of the impacts and adaptation strategies. Cambridge Prisms: Water. 2025;3:e4. doi:10.1017/wat.2024.16 https://doi.org/10.1017/wat.2024.16

Copernicus Climate Change Service. (2024). ERA5 post-processed daily statistics on single levels from 1940 to present [Dataset]. ECMWF. DOI: 10.24381/cds.4991cf48

ECMWF. (2024). ERA5: Data documentation. Copernicus Knowledge Base.

Feng, H., & Zhang, M. (2015). Global land moisture trends: Drying in summer but wetting in winter. *Water Resources Research*, 51(5), 3618–3636.

Forbes, R. M., Tompkins, A. M., & Untch, A. (2011). A new prognostic bulk microphysics scheme for the IFS. *ECMWF Technical Memorandum*, 649, 1–28.

Gemitzi, A., & Kofidou, M. (2022). A Google Earth Engine tool to assess water budget and its individual components. *Global NEST Journal*, 24(2), 331–336.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.

Helsel, D. R., & Hirsch, R. M. (2002). *Statistical Methods in Water Resources*. U.S. Geological Survey Techniques of Water-Resources Investigations, Book 4, Chapter A3.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049.

Hillel, D. (1998). *Environmental Soil Physics*. Academic Press.

Hussain, M., Mahmood, A., & Yar, P. (2021). Precipitation variability over Potohar Plateau and its impact on rainfed agriculture. *Pakistan Journal of Meteorology*, 17(34), 45–58.

Immerzeel, W. W., van Beek, L. P. H., & Bierkens, M. F. P. (2010). Climate change will affect the Asian water towers. *Science*, 328(5984), 1382–1385.

Jarvis, P. G. (1976). The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London B*, 273(927), 593–610.

Kendall, M. G. (1975). *Rank Correlation Methods* (4th ed.). Charles Griffin.

Khan, A. J., Koch, M., & Chinchilla, K. M. (2020). Evaluation of gridded multi-satellite precipitation estimation (TRMM-3B42-V7) performance in the Upper Indus Basin. *Climate*, 6(3), 76.

Lavers, D. A., Simmons, A., Vamborg, F., & Rodwell, M. J. (2022). An evaluation of ERA5 precipitation for climate monitoring. *Quarterly Journal of the Royal Meteorological Society*, 148(748), 3152–3165.

Liu W, Fu Z, van Vliet MTH, Davis KF, Ciais P, Bao Y, Bai Y, Du T, Kang S, Yin Z, Fang Y, Wada Y. Global overlooked multidimensional water scarcity. Proc Natl Acad Sci U S A. 2025 Jul;122(26): e2413541122. doi: 10.1073/pnas.2413541122. Epub 2025 Jun 23. PMID: 40549919; PMCID: PMC12232667. https://pmc.ncbi.nlm.nih.gov/articles/PMC12232667

Liyin He, Lorenzo Rosa, Solutions to agricultural green water scarcity under climate change, PNAS Nexus, Volume 2, Issue 4, April 2023, pgad117, https://doi.org/10.1093/pnasnexus/pgad117

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3), 245–259.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis* (5th ed.). Wiley.

Qureshi, A. S., McCornick, P. G., Sarwar, A., & Sharma, B. R. (2010). Challenges and prospects of sustainable groundwater management in the Indus Basin, Pakistan. *Water Resources Management*, 24(8), 1551–1569.

Rabie, A. B., Elhag, M., & Subyani, A. (2025). Remote Sensing, GIS, and Machine Learning in Water Resources Management for Arid Agricultural Regions: A Review. Water, 17(21), 3125. https://doi.org/10.3390/w17213125

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., ... & Toll, D. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85(3), 381–394.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, 63(324), 1379–1389.

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If precipitation extremes are increasing, why aren't floods? *Water Resources Research*, 54(11), 8545–8551.

Sharma, B. R., & Sharma, D. (2008). Impact of climate change on water resources and glacier melt and potential adaptations for Indian agriculture. New Dehli: International Water Management Institute

Sheffield, J., Wood, E. F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., & Verbist, K. (2018). Satellite remote sensing for water resources management: Potential for supporting sustainable development in data-poor regions. *Water Resources Research*, 54(12), 9724–9758.

Shemer, H., Wald, S., & Semiat, R. (2023). Challenges and solutions for global water scarcity. Membranes, 13(6), 612. https://doi.org/10.3390/membranes13060612

Somasundaram, D., Zhang, F., Ediriweera, S., Wang, S., Li, J., & Zhang, B. (2020). Spatial and temporal changes in surface water area of Sri Lanka over a 30-year period. Remote Sensing, 12(22), 3701.https://doi.org/10.3390/rs12223701

Soulis, K. X. (2021). Soil Conservation Service Curve Number method: Current applications, remaining challenges, and future perspectives. *Water*, 13(2), 192.

Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., ... & Seneviratne, S. I. (2009). A regional perspective on trends in continental evaporation. *Geophysical Research Letters*, 36(2), L02404.

Tzanakakis, V. A., Paranychianakis, N. V., & Angelakis, A. N. (2020). Water supply and water scarcity. Water, 12(9), 2347.https://doi.org/10.3390/w12092347

Ullah, S., You, Q., Ullah, W., & Ali, A. (2018). Observed changes in precipitation in China-Pakistan Economic Corridor during 1980–2016. *Atmospheric Research*, 210, 1–14.

USDA-NRCS. (2004). *National Engineering Handbook*, Chapter 10. USDA.

Yue, S., Pilon, P., Phinney, B., & Cavadias, G. (2002). The influence of autocorrelation on the ability to detect trends in hydrological series. *Hydrological Processes*, 16(9), 1807–1829.