Assessment of Climatic and Vegetation Influence on Spatial Distribution of Groundwater Recharge in Humid Subtropical Central Gangetic Plain

Anuradha Karunakalage[1], Ravi Sharma[1], Mohammad Taqi Daqiq[1], and Suresh KANNAUJIYA[2]

[1]Indian Institute of Technology Roorkee
[2]Indian Institute of remote sensing

May 5, 2023

Abstract

Groundwater Recharge (GR) is a crucial part of sustainability studies since it is one of the key factors responsible for making the groundwater resource renewable. An optimum strategy for responding to water level decline is artificial groundwater recharge. Artificial groundwater recharge projects are limited by cost, and the effective area is less. The role of natural factors for groundwater recharge is well defined and recognized in arid regions, whereas it's challenging for humid areas. The current study's main aim is to understand the contribution of the bio-geophysical aspect to groundwater recharge in the subtropical monsoon state of Uttar Pradesh in the Gangetic Plain. However, recharging is also one of the least understood processes because it changes over time and space and is challenging to quantify directly for a larger area. This research applied the 'water and energy transfer among bare soil, vegetation, and atmosphere (WetSpass)' model to estimate direct natural GR for Uttar Pradesh. The model's output and its regression processes with climate, slope, soil type, and vegetation give a comprehensive understanding of natural controlling factors. Among the aforementioned controlling factors, though climate sharpens recharge dominantly, vegetation has shown a significant role in some areas of the state. In contrast to the prevailing view, vegetation cover can enhance groundwater recharge in the state. Thus, planting, and various tree management options, including groundwater-feeding species as a secondary plantation in cropland, can improve groundwater resources.

Hosted file

962360_0_art_file_10945757_rtywfx.docx available at https://authorea.com/users/614856/articles/641532-assessment-of-climatic-and-vegetation-influence-on-spatial-distribution-of-groundwater-recharge-in-humid-subtropical-central-gangetic-plain

Hosted file

962360_0_table_10946148_rtywfx.docx available at https://authorea.com/users/614856/articles/641532-assessment-of-climatic-and-vegetation-influence-on-spatial-distribution-of-groundwater-recharge-in-humid-subtropical-central-gangetic-plain

Assessment of Climatic and Vegetation Influence on Spatial Distribution of Groundwater Recharge in Humid Subtropical Central Gangetic Plain

Anuradha Karunakalage[1], Ravi Sharma[1], Mohammad Taqi Daqiq[1], and Suresh Kannaujiya[2]

[1]Indian Institute of Technology Roorkee, Department of Earth Sciences, Uttarakhand – 247667 India.

[2]Indian Institute of Remote Sensing, Indian Space Research Organization, Uttarakhand –248001 India.

Corresponding author: Ravi Sharma (ravi.sharma@es.iitr.ac.in)

† Assistant professor/professor in charge of Rock and Fluid Multiphysics Laboratory, Department of Earth Sciences, Indian Institute of Technology Roorkee, India.

Key Points:

- The WetSpass model has estimated spatial distribution of groundwater recharge accounting by hydrometeorological and bio-geophysical factors.
- The relationship between simulated recharge and climate strongly correlates with function derived from a global recharge data and climate.
- The climate is the dominant factor influencing the fraction of recharge, causing groundwater recharge.
- Baseflow can be a proxy for groundwater recharge excluding high terrain, small, and artificial surface water dividing catchments.
- Vegetation is the second crucial controlling factor.

Abstract

Groundwater Recharge (GR) is a crucial part of sustainability studies since it is one of the key factors responsible for making the groundwater resource renewable. An optimum strategy for responding to water level decline is artificial groundwater recharge. Artificial groundwater recharge projects are limited by cost, and the effective area is less. The role of natural factors for groundwater recharge is well defined and recognized in arid regions, whereas it's challenging for humid areas. The current study's main aim is to understand the contribution of the bio-geophysical aspect to groundwater recharge in the subtropical monsoon state of Uttar Pradesh in the Gangetic Plain. However, recharging is also one of the least understood processes because it changes over time and space and is challenging to quantify directly for a larger area. This research applied the 'water and energy transfer among bare soil, vegetation, and atmosphere (WetSpass)' model to estimate direct natural GR for Uttar Pradesh. The model's output and its regression processes with climate, slope, soil type, and vegetation give a comprehensive understanding of natural controlling factors. Among the aforementioned controlling factors, though climate sharpens recharge dominantly, vegetation has shown a significant role in some areas of the state. In contrast to the prevailing view, vegetation cover can enhance groundwater recharge in the state. Thus, planting, and various tree management options, including groundwater-feeding species as a secondary plantation in cropland, can improve groundwater resources.

Plain Language Summary

Information on groundwater recharge is essential for groundwater modeling and management but is challenging to monitor and assess across broad areas. This study has shown that openly accessible data provides an important opportunity to examine the spatial distribution of groundwater recharge using the WetSpass model. Discussing the spatial distribution of groundwater recharge along with the natural controlling factors is vital for establishing policies and regulations on proper management for sustainable usage of aquifers. Further, the study has highlighted the contribution of vegetation to the up taking of recharge, which could deviate from the traditional view of artificial groundwater recharge to enhance the availability of supportive natural factors.

1 Introduction

Groundwater is the largest freshwater source in the world and is a nonrenewable source to meet agricultural, domestic, and industrial water requirements, especially for tropical and subtropical semi-arid regions(Owuor et al., 2016). GR is a crucial part of sustainability studies since it is one of the key factors responsible for making the groundwater resource renewable (Alley et al., 2002; Berghuijs et al., 2022; Gleeson et al., 2012)and sustainability of the groundwater-supportive ecological community and inland water (Gleeson et al., 2020). The entry of water into the saturated zone made available at the water table surface and the corresponding flow away from the water table within the saturated zone is referred to as recharge (O. Batelaan & De Smedt, 2007). Groundwater recharge rates vary by magnitude due to the diversity of Earth's landscapes and climates (MacDonald et al., 2021; Moeck et al., 2020; Scanlon et al., 2006). However, recharging is one of the least understood processes because it changes over time and distance with surface properties, morphology, and vegetation (Crosbie et al., 2018; Moeck et al., 2020; de Vries & Simmers, 2002). In general, near-surface conditions greatly impact groundwater recharge in (semi-) arid regions than in more humid locations. Deep percolation in humid environments is primarily governed by the potential surplus of precipitation (rainfall minus potential evapotranspiration), the soil's ability for infiltration, and the subsurface's capacity for storage and transit (de Vries & Simmers, 2002). Thus, it is challenging to quantify directly (Healy & Scanlon, 2010; Moeck et al., 2020; Scanlon et al., 2006; Zomlot et al., 2015), and validation of large-scale simulated recharge is remnant yet (Döll & Fiedler, 2008; de Graaf et al., 2019; Li et al., 2021; Müller Schmied et al., 2021). Though numerous studies/models exist to simulate the long-term behavior of aquifer systems, such outcomes cannot be incorporated into the development of management schemes excluding reliable estimation of spatiotemporal variation of recharge (Sophocleous, 2005).

India ranks first among the world's largest groundwater exploiters, with 25% of the total global abstraction, about 230 km3 per year. Since the beginning of the Green Revolution in the 1980s, many states have been over-pumping groundwater for cash crop cultivation, which is highly water-intensive (Sarkar et al., 2022). One of the world's largest alluvial expanses is the Ganga Plain (Pokharia et al., 2017; Singh, 1996), covering nearly two-thirds of Uttar Pradesh. Uttar Pradesh is the largest producer of cash crops, drawing out billions of liters of groundwater. Besides introducing high-yielding crops, increasing population and industrialism are adverse demands for groundwater in the state. Thus, estimating the groundwater recharge rate and assessing controlling factors is paramount for establishing new policies and regulations for proper aquifer management.

Though various methods have been developed to estimate groundwater recharge (Scanlon et al., 2002; Zomlot et al., 2015), discrepancy and uncertainty are imperious factors in recharge simulations (Scanlon et al., 2002). According to regional scale analysis of Carbonate, landscapes represent the underestimation of GR from frequently applied hydrological models. Also, it is uncertain how common this model bias is, given that the increased recharge rates have been linked to high preferential flows in karst terrain (Hartmann et al., 2017a). Such disparities between models and observations are based on measurements of recharge and groundwater in specific landscapes, the physical properties of aquifers, and climate conditions (O. Batelaan & De Smedt, 2007; Berghuijs et al., 2022). As a result, over the last two decades, more approaches have been developed to assimilate the spatial-temporal variance of recharge in groundwater modeling (Berghuijs et al., 2022; Best & Lowry, 2014; Cooper et al., 2015; Eilers et al., 2007; Hemmings et al., 2015; Hughes et al., 2008; Jyrkama & Sykes, 2007; Markstrom et al., 2008; Minor et al., 2007; Zomlot et al., 2015). Instead of authentical GR estimation, other primary benefits of this advancement are that it will allow researchers to investigate the effects of climate and land-use change on groundwater resources at unprecedented degrees of temporal and geographical variability (Healy & Scanlon, 2010).

WetSpass was designed as a physically based methodology for estimating long-term average, spatial distribution, water balance components, surface runoff, evapotranspiration, and groundwater recharge (O. Batelaan & De Smedt, 2007; Okke Batelaan & De Smedt, 2001). The acronym for Water and Energy Transfer between Soil, Plants, and Atmosphere in a quasi-steady state is referred to as "WetSpa." It was constructed on the foundations of the "WetSpa" time-dependent spatial distributed water balance model (O. Batelaan & De Smedt, 2007; Okke Batelaan & De Smedt, 2001; Zomlot et al., 2015). In recharge estimation, the WetSpass model functions along with the spatial distribution of soil texture, slope, the spatiotemporal distribution of land use, and climatic variables. WetSpass can be iteratively linked to a groundwater model, MODFLOW, which provides the water table location, and WetSpass returns recharge estimates correspondingly (Zomlot et al., 2015). The WetSpass model has been applied for different watersheds in the world with the heeding significance of estimation of the long-term behavior of water balance components and impact of land use land cover changes on recharge such as Nile delta aquifer, Egypt (Armanuos et al., 2016), Beijing China (Zhang et al., 2017), Bilate basin Ethiopia (Dereje & Nedaw, 2019), Birki watershed of Geba river basin Ethiopia (Meresa et al., 2019), Poznan Upland Poland (Graf & Przybyłek, 2018), Flanders region of Northern Belgium (Zomlot et al., 2015), Moulouya basin, Morocco (Amiri et al., 2022), and Southern hill region Bangladesh (Sadeak & Khan, 2021).

Several publications advise estimating recharge using various techniques and contrasting the results (Risser et al., n.d.; Scanlon et al., 2002) due to the challenge of recharge estimations for a larger area. Base flow from stream gauging stations has been utilized in numerous studies to estimate groundwater recharge as a measure of comparison (Arnold et al., 2000; O. Batelaan & De Smedt, 2007; Eckhardt, 2008; Risser et al., n.d.; Zomlot et al., 2015). Base flow is the gradually changing portion of streamflow that results from groundwater storage and other delayed sources, including lakes, wetlands, melting snow, ice, and channel bank storage (Beck, Van Dijk, et al., 2013). Base flow can be defined as the groundwater reservoir's discharge into the rivers related to the subtropical state of Uttar Pradesh (Zomlot et al., 2015). Base flow estimates recharge under the primary presumption that groundwater discharge and recharge are roughly equal. That base flow equals the entire groundwater discharge of a watershed (Piggott et al., 2005). Different watershed characteristics and the reciprocal action of groundwater-surface water drive the relationship between recharge and baseflow. Hence, except for some minor catchments and catchments with silty soil, the base flow might thus be regarded as a proxy for recharge (Zomlot et al., 2015). However, due to the difficulty in directly comparing base flow and recharge because most base flow methodologies find some proxy for groundwater discharge and, thus, for actual recharge, the scientific community does not fully embrace these hypotheses (Rutledge, 2005). Thus, this study has compared the simulated recharge values from WetSpass with simulated recharge values from global groundwater recharge region models. The state of Uttar Pradesh is humid and subtropical; hence, the WetPass results have been compared with model outputs based on climate aridity (Berghuijs et al., 2022). The model based on climate aridity has been compared with scientifically accepted global models such as PCR-GLOB (de Graaf et al., 2019), WATER-GAP (Müller Schmied et al., 2021), and machine learning models (Mohan et al., 2018)based on observed versus model predicted at 5237 sites (Berghuijs et al., 2022) and proved that compared to the aridity based model other models underestimating recharge 50% than actual recharge measurements (Berghuijs et al., 2022).

In WetSpass simulation, accounting wide range of spatial variability of hydrometeorological and bio-geophysical factors are caused for susceptibility to discuss the behavior of controlling factors individually and combinedly. Statistical regression approaches are generally used in hydrologic studies, such as estimating recharge and base flow based on the watershed (Delin et al., 2007; Gebert et al., 2007; Jing et al., 2019; Longobardi & Villani, 2008; Mazvimavi et al., 2005)The previous studies have successfully performed a potential statistical correlation between WetSpass simulated groundwater recharge and base flow for understanding the reliability of functions of WetSpass and assessment of controlling factors for recharge (O. Batelaan & De Smedt, 2007; Okke Batelaan & De Smedt, 2001; Zomlot et al., 2015).

However, the regression approach might encounter significant challenges when the independent variables are associated with one another (Jasim Mohammed Rajab et al., 2012).

Principal Component Analysis (PCA) is hence helpful in reducing the multicollinearity issue (Jasim M. Rajab et al., 2013). Precipitation, evapotranspiration, saturated hydraulic conductivity of soil (Ks), and land-use type are the key watershed features that affect the recharge of the state of Uttar Pradesh. PCA has encountered the significance of controlling factors among those for variation of the spatial distribution of GR.

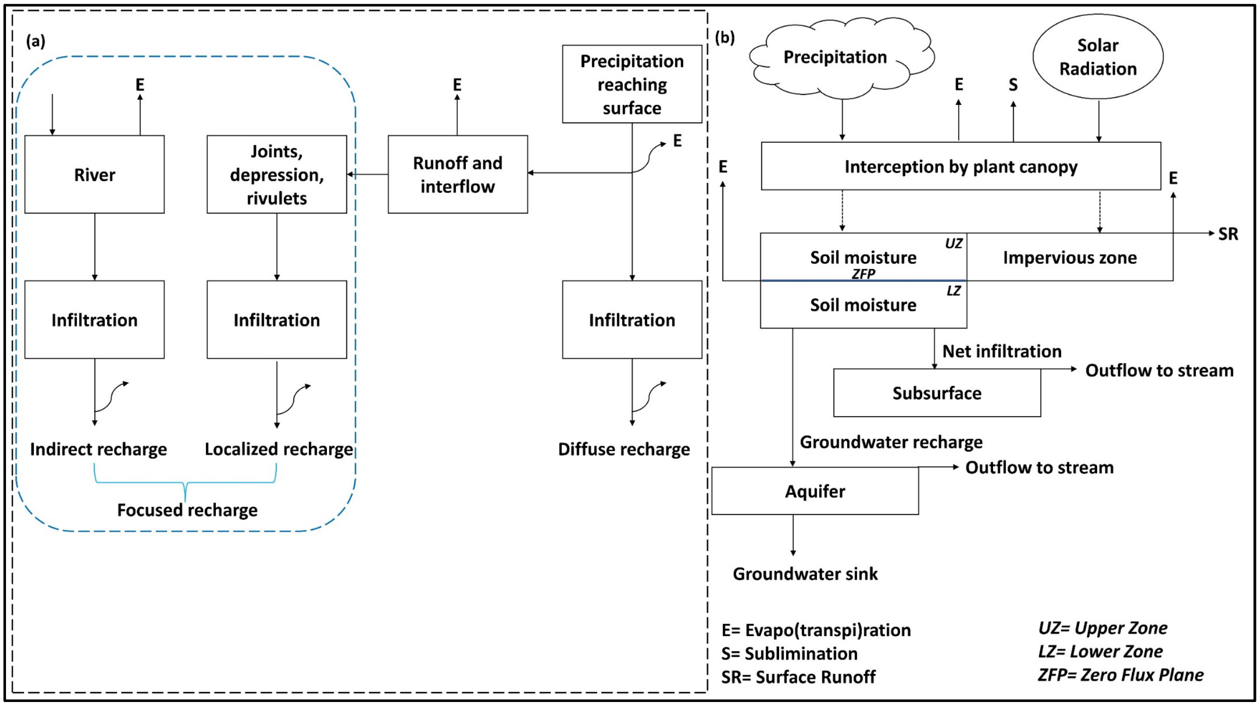


Figure 1. (a) Recharge mechanisms modified after (Healy, 2010), and (b) Diffused groundwater recharging process flow diagram based on the hydrological cycle.

2 Materials and Methods

2.1 Study Area

The study area covers the humid subtropical Central Ganga Plain, which expanded over the state of Uttar Pradesh. The state covers an area of approximately 2,40,928 km^2 and is bordered by latitudes 23°52 to 31°25 N and longitudes 84°39 to 77°03 E. Water stress has been present throughout the state for several decades as a result of the alarming population expansion as the highest populated state of the highly populated nation of the world, rapid urbanization, and quick industrialization (Ansari et al., 2000; Umar, 2006). Agriculture is the mainstay of living, strongly reliant on groundwater(Umar, 2006). The Southwest monsoon significantly impacts the region's climate, which is generally humid. Spring, which lasts from the middle of February to the middle of March, is one of the distinct seasons. Summer is the period from the middle of March to the middle of June when temperatures are high (the mean maximum temperature is roughly 47 ^{0}C), and the wind is strong, hot, and dry. The rainy season begins from the end of June through the end of September. 1020-1140 mm of rain has been recorded as falling in the area annually. The average minimum and maximum temperatures during winter, lasting from November to mid-February, are 7.6 ^{0}C and 21 ^{0}C, respectively. It rarely gets below 0 ^{0}C, and although it occasionally rains in January, the air is generally quite dry (Chauhan et al., 2015; Umar, 2006).

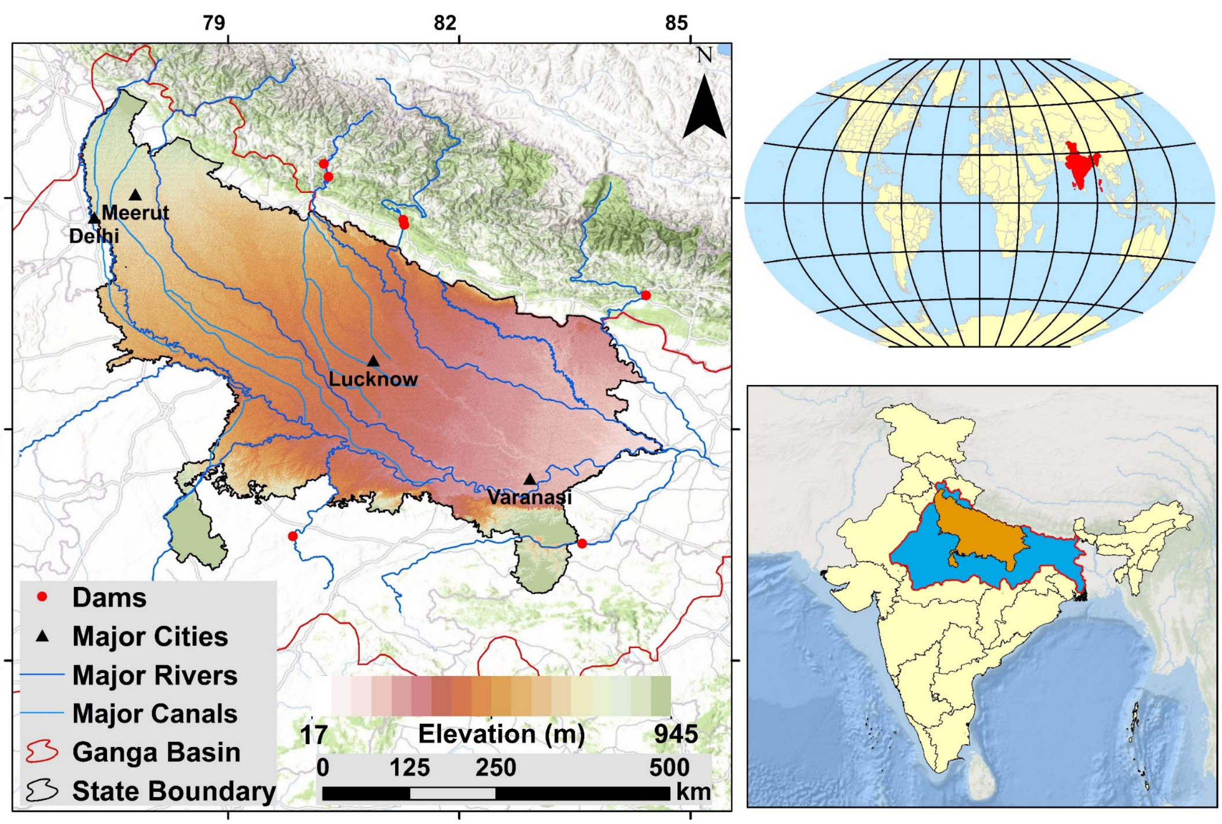


Figure 2. The central Ganga basin covers the state of Uttar Pradesh.

The state comprises many rock types, ranging from the oldest Archean metamorphoses to the most recent Quaternary alluvium. An extensive area of the state is covered by Gangetic Plain alluvium, which is separated from Himalayas and Peninsula India. Archean to Mesozoic period rocks is being exposed in the Southern peninsular portion of Uttar Pradesh. Due to the above-explained geological framework, the state's hydrogeological structure comprises porous and fractured rocks. The primary rivers that represent the drainage of Uttar Pradesh include the Ganga, Yamuna, Ramganga, Gomti, Ken, and Betwa and Ghaghra (Kumar Dinkar et al., 2019).

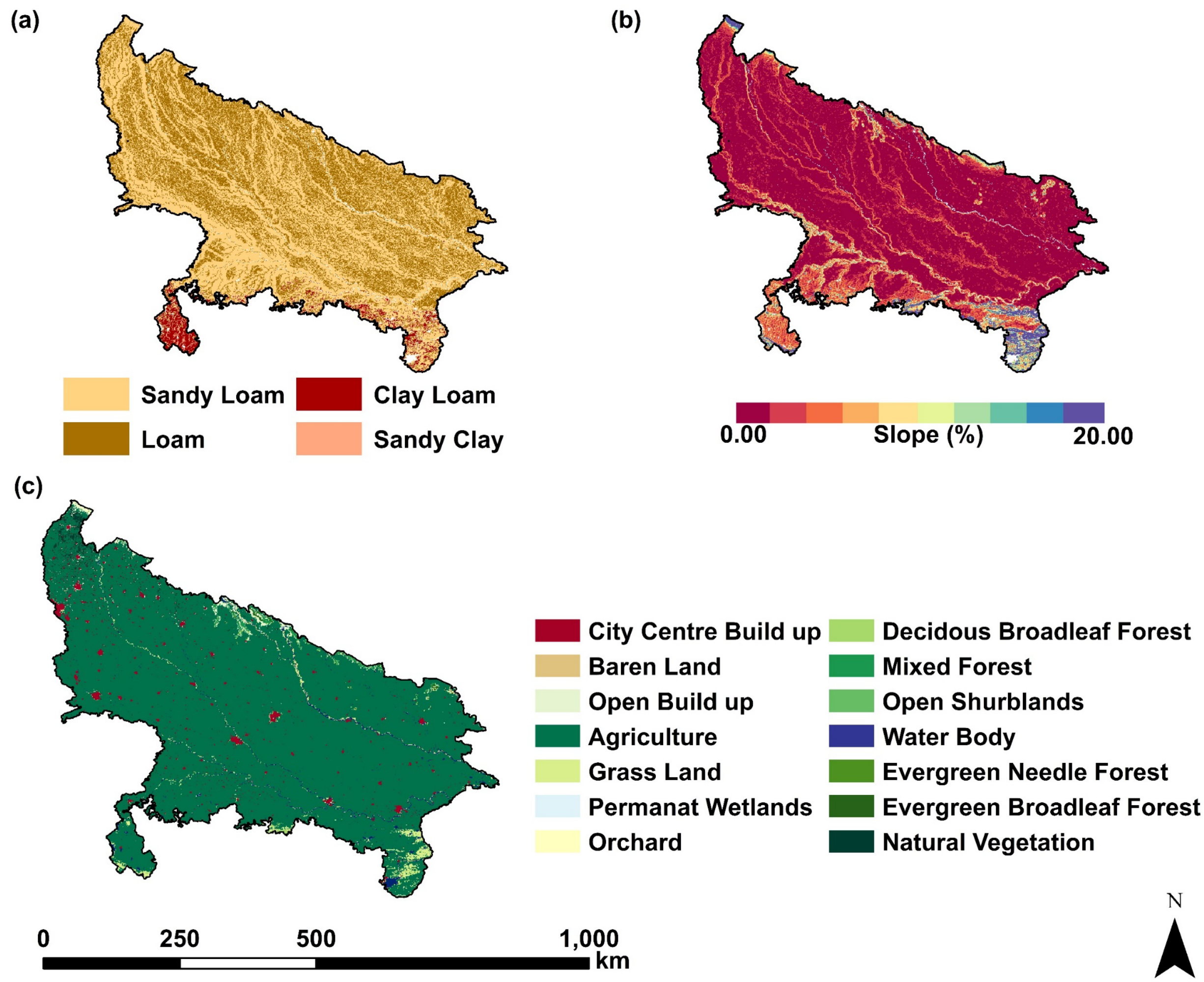


Figure 3. Spatial coverage of (a) Soil texture, (b) change in elevation as slope percentage, and (c) land use of Uttar Pradesh.

2.2 Data for WetSpass Model

The WetSpass model requires spatial distribution of precipitation (mm), potential evapotranspiration (mm), temperature (degree Celsius (^{0}C)), Wind speed (m/s), groundwater depth (m), topography (m), slope (%), and LULC data and spatial resolution should be ideal for all the raster layers. The study is carried out with 1km × 1km spatial resolution, and this fine-scale climatic data has been downloaded from Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA) (Karger et al., 2017). For winter and summer, two sets of raster data are needed to run a year simulation (Park et al., 2014). Winter is the first simulated using WetSpass as a dry season and summer as a wet season. In current study have averaged monthly precipitation, potential evapotranspiration, temperature, and wind speed data to estimate seasonal data. Groundwater depth data has been acquired seasonally as pre-monsoon for winter and post-monsoon for summer. LULC data has been obtained yearly. Instead of raster layers, necessary coefficient values related to LULC type and soil texture, runoff coefficient values, and individual aerial fractions for each land use type have been provided as pre-defined attribute tables based on the literature references.

2.3 Base Flow Index Data

The base flow index (BFI) measures the base flow's contribution to the overall stream flow (Bloomfield et al., 2009; Zomlot et al., 2015). The study has chosen BFI because it has been frequently used in recent literature and proved crucial in identifying how watershed characteristics affect base flow (Jasim M. Rajab et al., 2013; Zomlot et al., 2015). The study has acquired base flow index data of 5km × 5km spatial resolution from the Global Patterns of Base Flow Index by (Beck, van Dijk, et al., 2013). The study of (Beck, van Dijk, et al., 2013)s defined as the ratio between long-term mean base flow and stream flow. The study has used streamflow data from a highly diverse set of 3394 catchments covering over 10,000 km2 globally and widely applicable models for climatic and physiographic data. The BFI of (Beck, van Dijk, et al., 2013) has shown an R square value of 0.82 by performing with watershed characteristics of catchments. In the present study, recharge simulation resolution is 1km ×1km, and due to the level to gentle slope variation of the state of Uttar Pradesh, base flow estimation should be more associated with climate and other physiographic characteristics, including LULC, geology, and soil.

2.4 WetSpass Model

The WetSpass model simulated water and energy transfer between soil, plants, and the atmosphere in a quasi-steady state and was the first novel developed by (Okke Batelaan & De Smedt, 2001)and modified by (O. Batelaan & De Smedt, 2007). On a regional scale, this numerical forecast of the long-term (seasonal or monthly) spatial patterns of surface runoff, evapotranspiration, and groundwater recharge (O. Batelaan & De Smedt, 2007; Zomlot et al., 2015). The model views a basin or area as an arrangement of raster cells. Each raster cell is further separated into four surface models impermeable surface, open water, bare soil, and vegetation. Each grid cell's seasonal water balance is determined (Zomlot et al., 2015). The model uses a general equation for the water balance (Park et al., 2014),

$$P = S + ET + R \tag{1}$$

where P is precipitation (mm), S is surface runoff (mm), ET is evapotranspiration (mm), and R is recharge (mm).

$$S_j = f_{1_j}.P_n \tag{2}$$

where f_1 is the runoff factor for the surface model j that depends on land use type (as vegetated area (v), bare soil area (s), open water area (o), and impervious surface (i)), vegetation characteristics, soil texture, and slope. P_n is net precipitation recharging the subsurface (total precipitation minus interception by the plant canopy) (Yenehun et al., 2021).

$$ET_j = f_{2_j}.PET \tag{3}$$

where f_2 is the evapotranspiration factor for the surface model j that depends on land use type, vegetation characteristics, soil texture, and slope. PET is the potential evapotranspiration of open water (mm) (Yenehun et al., 2021).

$$T_{rv} = c.PET \tag{4}$$

where T_{rv} is transpiration, and c is the vegetation coefficient. The Penman-Monteith equation can be used to calculate the vegetation coefficient (Park et al., 2014).

The WetSpass model has calculated c as the ratio of reference vegetation transpiration from the Penman-Monteith equation.

$$c = \frac{1+{}^{\gamma}/_{\Delta}}{1+{}^{\gamma}/_{\Delta(1+{}^{\gamma}/_{\Delta})}} \tag{5}$$

where γ/Δ is the Penman coefficient, r_c is canopy resistance (s/m), and r_a is aerodynamic resistance (s/m).

r_a is a function of plant height and wind speed (Amer & Hatfield, 2004).

$$r_a = \frac{1}{k^2 u_a}\left[\ln\left(\frac{z_a-d}{z_o}\right)\right]^2 \tag{6}$$

Where k is Von Karman constant (0.4), u_a is wind speed (m/s), at z_a measurement level (2m), z_o is zero-plane displacement length (m), and d is roughness length for the vegetation or soil (m).

$$I = C_{ip}.P \tag{7}$$

where C_{ip} is the constant percentage of interception by vegetation type.

The interception, transpiration, and evaporation from the bare soil in a grid cell are added to determine the total actual evapotranspiration (Zomlot et al., 2015)

Calculations of surface runoff consider the capacity of the soil for infiltration as well as the quantity and intensity of the precipitation. The surface runoff is calculated in two levels. At the first level, it has been calculated potential surface runoff.

$$S_{v-pot} = C_{Sv}.(P-I) \tag{8}$$

where C_{Sv} is the runoff coefficient that is derived as a function based on slope, soil texture, and vegetation type.

The actual surface runoff, S_v, is estimated for recharge areas in the second stage by considering variations in precipitation intensities connected to soil infiltration capabilities because $S_{v\text{-}pot}$ simulates only groundwater-saturated areas.

$$S_v = C_{Hor}.S_{v-pot} \tag{9}$$

where C_{Hor} is the parameterization coefficient for the seasonal precipitation component of the Hortonian surface runoff. It can be calculated by determining the proportion of seasonal precipitation with an intensity greater than a specific soil type's capacity for infiltration.

Equation (**1**) has been rearranged based on the four different surface models based on the four surface environments where a raster cell has been subdivided.

For vegetated areas,

$$P = I + S_v + T_v + R_v \tag{10}$$

For bare soil areas,

$$P = S_s + E_s + R_s \tag{11}$$

For open water areas

$$P = S_o + E_o + R_o \tag{12}$$

For impervious surfaces

$$P = S_i + E_i + R_i \tag{13}$$

where T_v is actual transpiration (mm), and I is interception (mm).

The residual term from the water balance calculates groundwater recharge (Zomlot et al., 2015). The independent water balances for the various fractions per raster cell are then added to determine each grid cell's water balance.

$$ET_c = a_v . ET_v + a_s . E_s + a_o . E_o + a_i . E_i \quad (14)$$

$$S_c = a_v . S_v + a_s . S_s + a_o . S_o + a_i . S_i \quad (13)$$

$$R_c = a_v . R_v + a_s . R_s + a_i . E_i \quad (14)$$

where *c* and *a* indexes represent cell and aerial fraction of different surface models, respectively. (O. Batelaan & De Smedt, 2007) provide a more detailed explanation, calibration, and validation based on a case study of the WetSpass model for a region of Flanders.

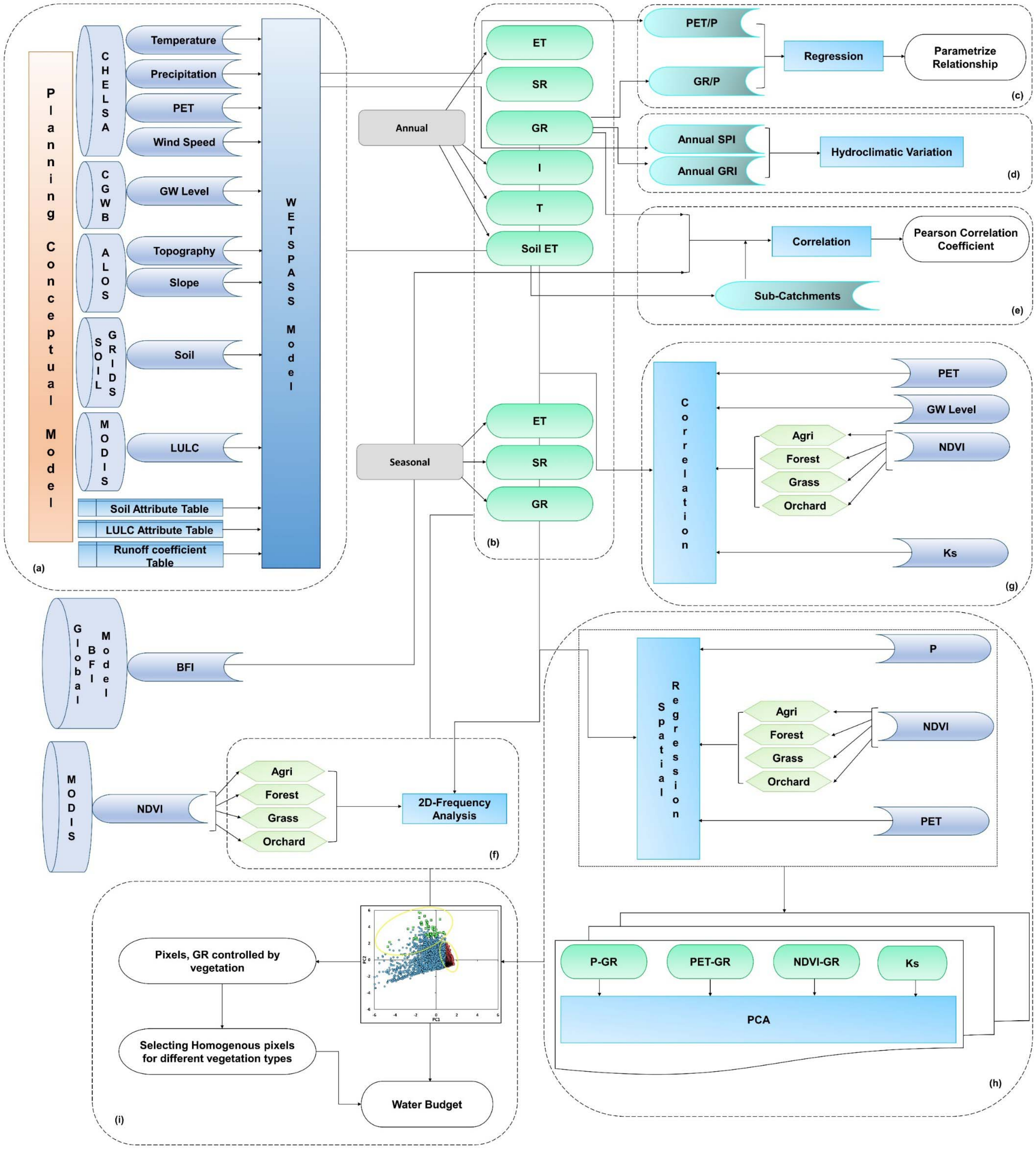


Figure 4. Illustration of workflow and data processing.

2.6 Primarily Identification of Controlling Factors

The study has considered precipitation, potential evapotranspiration, groundwater depth, NDVI, and saturated hydraulic conductivity of soil to be the most significant parameters for recharge in the state after employing the Pearson correlation test. Pearson's correlation coefficients can be used to determine the significance of potential relationships between recharge and different climatic and bio-geophysical watershed characteristics (Zomlot et al., 2015). Understanding the connection between soil texture and recharge is crucial for water resource management because soil texture can affect the water balance and groundwater storage at different scales of inter-annual variability (Keese et al., 2005; Wang et al., 2009). Timing and recharge rates are influenced by land use and land cover. Previous studies have shown the influence of geology and landforms on recharge rates (Moeck et al., 2020). But the state's geology is dominantly unconsolidated sediments and a flat plain in landform, excluding the smaller area peninsular area in the extreme south. According to a global synthesis of recharge estimates, the vegetation type is the second most significant element influencing recharge rates, behind precipitation (Ajami, 2021; Kim & Jackson, 2012). Thus, vegetated type and soil texture are taken into account as the qualitative parameters responsible for the state's spatial distribution of direct natural groundwater recharge variation in the state. Regression analysis is used to find a relationship between the values of two or more variables, at least one of which is subject to random variation, and to determine the statistical significance of such a relationship, whether assumed or calculated (Oosterbaan, 1994). The linear regression analysis has been performed between recharge and precipitation, soil texture (saturated hydraulic conductivity), and vegetation type (NDVI). A statistical relation has not been obtained between recharge and saturated hydraulic conductivity or between recharge and NDVI. Thus, 2D frequency diagrams have been generated to identify the number of grid cells with a higher correlation with recharge, saturated hydraulic conductivity, and NDVI.

2.5 Principal Component Analysis

To compress a huge set of variables into "artificial" variables, known as "principal components," which account for the majority of the variance in the original variables, principal components analysis uncovers hidden structure in the dataset. The high correlation between variables clearly indicates high redundancy in the data. Therefore, the study has employed PCA to reduce redundancy and identify significant variables that account for the majority of the variation in recharge. Since the principal components (PCs) are not independent of the scales in which the original variables are measured, the derivation of the PCs was based on the correlation matrix of standardized data (Jolliffe, 2002; Zomlot et al., 2015). PCA has been conducted in two levels. In the first approach, PCA has been employed separately for different vegetated areas such as agriculture, forest, sparsely vegetated, and orchard areas associating data of P, PET, Normalized Difference Vegetation Index (NDVI) and Ks to identify significant GR controlling factors for each vegetated area. The second PC analysis was carried out with R square values between GR, P, PET, NDVI, and Ks of each grid cell. Based on the PC 1 and PC 2 score values, it has classified extreme positive influence grids for PC 1 and PC 2 to identify areas (raster grids) where GR is controlled by dominantly either precipitation or vegetation type.

3 Results and Discussion

3.1 Spatial Distribution of Water Balance Components

The WetSpass simulation has generated water balance components seasonally and annually. These raster outputs represent the spatial distribution of the water balance components and grided quantification.

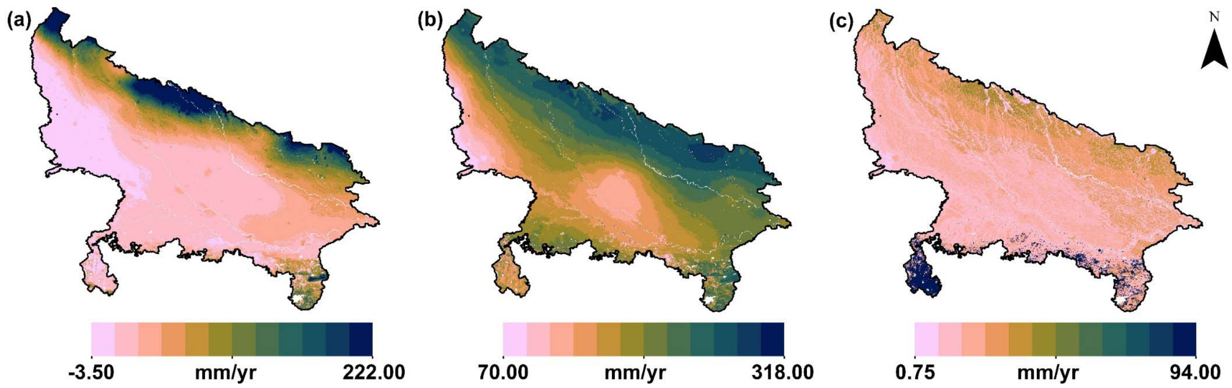


Figure 5. Spatial distribution of the simulated long-term average of (a) groundwater recharge, (b) evapotranspiration, and (c) surface runoff.

The long-term average of simulated recharge over the state shows a wide range of recharge variations between -3.5 to 222 mm. The annual average value per grid is 16.87 mm, and the standard deviation is 21.74 mm. The total aquifer recharge is significantly contributed by the summer season due to associating with the rainy season. The long-term seasonal average recharge values per grid are 16.83 mm and 0.04 mm for summer and winter, respectively. Negative recharge happens when the overall evapotranspiration exceeds the infiltration (Net Precipitation- Runoff). Only areas with shallow groundwater experience this. Plant roots can enter the saturated zone when the water table is close to the Earth's surface, such as valleys, polders, and regions close to lakes and rivers. This enables the plants to transpire water straight from the groundwater system.

The annual average of evapotranspiration spatially varies from 70 to 318 mm and is dominant over the state due to the more extensive coverage of croplands. The average evapotranspiration and standard deviation per grid cell are 150.08 mm and 21.37 mm, respectively. Land usage and vegetation can have significant effects on the recharge processes. Types and densities of vegetation influence evapotranspiration patterns. A land surface covered in vegetation often evaporates at a higher rate, leaving less water available for recharging. The annual mean of surface runoff distribution variation is insignificant compared to recharge and evapotranspiration over the state. The higher and moderate surface runoff is only bounded to extreme southern and northern regions due to higher elevation changes over the areas. The spatial variation of surface runoff varies between 0.75 to 94 mm.

The mean spatial variation of surface runoff per grid is 8.27 mm, and the standard deviation over the area is 9.15 mm. Land surface topography is crucial for both diffuse and focus recharge. High runoff rates and low infiltration rates are typical of steep slopes. Diffuse recharge is dominant in flat terrain environments with low surface drainage and causes floods.

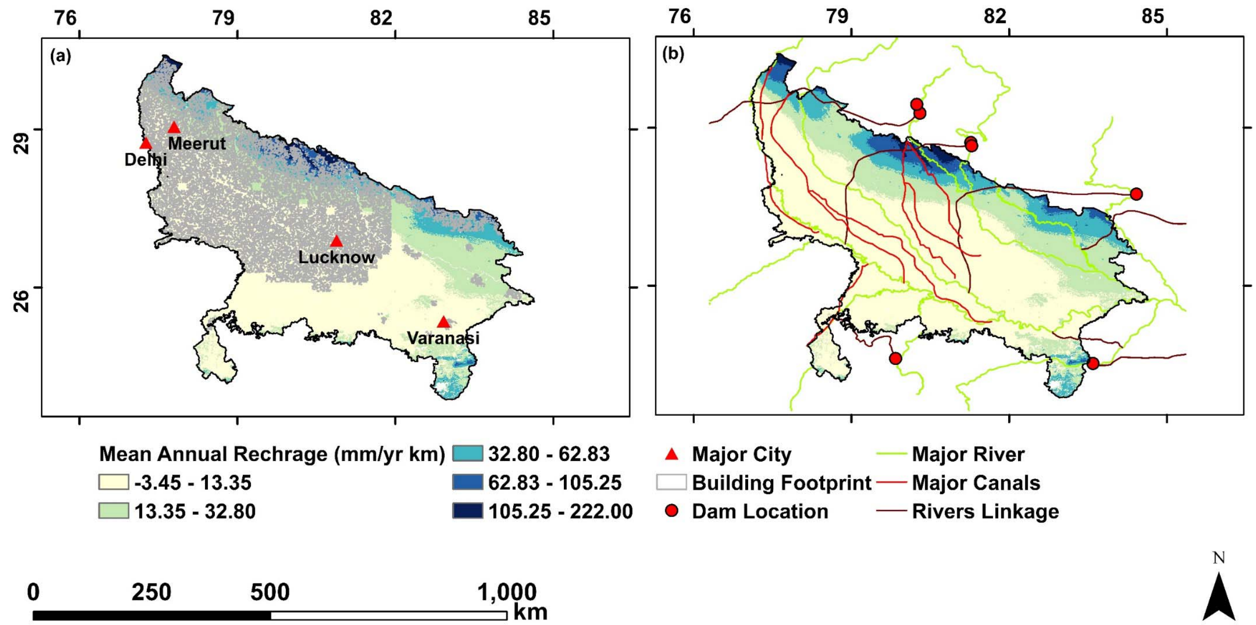


Figure 6. Building footprints (Data source;) (a) and inland hydrology (b) over the state.

The state has shown significantly low groundwater recharge over the East and central part of the state. This can be explained by prominent agricultural lands and urbanization. In comparison to an unvegetated land surface under similar conditions, a vegetated land surface often has a higher rate of evapotranspiration (and, consequently, less water available for recharging) (Healy & Scanlon, 2010). Uttar Pradesh is the highest populated state in India. Many alterations to the land surface that urbanization causes can significantly impact recharging processes due to artificial treatment on the surface and subsurface (Healy & Scanlon, 2010; Price, 2011). Impervious places like parking lots, buildings, and roads can all prevent recharge. Diversions for runoff are typical elements of urban settings. Diversions may lead to infiltration galleries or surface-water bodies. Under the first scenario, the region's overall recharge is decreased. The latter scenario could change the source of recharge from a diffuse to focused recharge. However, it may not necessarily result in a reduction in recharge. The state is dominant with natural rivers and artificial canals. This may be caused by focused recharge. The WetSpass simulation has not been associated with indirect or localized recharge. Thus, net recharge value can be uplifted with the focused recharge process.

3.2 Parameterized Relationship between Climate and Recharge

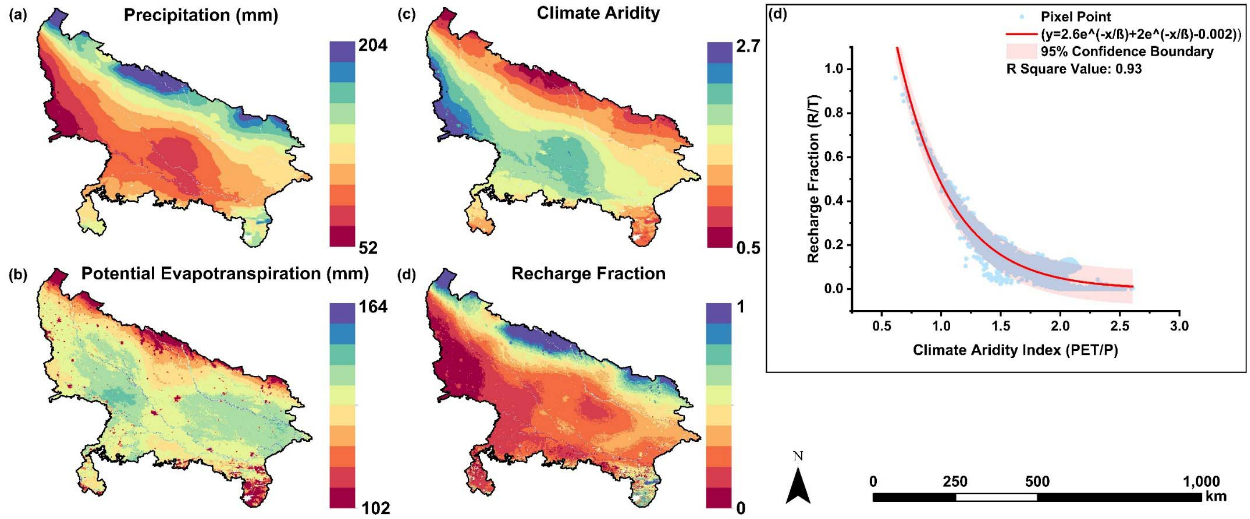


Figure 7. (a) and (b) represent the spatial distribution of hydrometeorological components as mean annual precipitation and potential evapotranspiration, respectively, whereas (c) and (d) shows the response of climatic and hydrogeological components for quantity and distribution of hydrometeorological factors such as climate aridity, and recharge fraction. (d) is the graphical output of the statistical relationship between climate aridity and recharge fraction for 1km ×1km grids over the state. The red line represents the calibrated sigmoidal function.

Systematic and random variations in diffuse and focused recharge rates can be observed in space. Climate patterns are frequently linked to systematic trends, although geology and land use are also significant (Healy & Scanlon, 2010; Moeck et al., 2020). Often, the most significant factor influencing variation in recharge rates is climate fluctuation in humid regions. The primary factor in the water budget for the majority of watersheds is precipitation, which is the source of natural recharge. In Figure 7), it was proved that spatial changes in recharge fraction have been strongly influenced by climate aridity. Generally, larger portions of precipitation recharge groundwater in humid areas. With increasing aridity, this recharge proportion decreases until it frequently drops to virtually zero in highly arid locations. The simulated recharge values of each pixel show substantial variation with the aridity of the same particular pixels, and it is non-linear. The same relationship has been derived from the global scale study of (Berghuijs et al., 2022)using empirical recharge values and the given aridity of the particular location. The pattern is sufficiently monotonous to derive a highly significant correlation between the amount of precipitation that recharges groundwater and the aridity of the climate.

$$\frac{R}{P} = 2.6e^{(-\emptyset/\beta)} + 2e^{(-\emptyset/\beta)} - 0.002 \quad (15)$$

$$R = P.\left[2.6e^{(-\emptyset/\beta)} + 2e^{(-\emptyset/\beta)} - 0.002\right] \quad (16)$$

where R is the mean annual recharge, P is the mean annual precipitation, $\emptyset$ is aridity (ratio between mean annual potential evapotranspiration and precipitation), and β is the characteristic subtractor of the aridity.

Equation (**15**) represents the exponential decrease of recharge fraction due to increased aridity.

3.3 Impact on Hydroclimatic Variation

The most crucial element influencing variation in recharge rates is frequent climate fluctuation. The primary factor in the water budget for most watersheds is precipitation, which is the source of natural recharge. Temporal fluctuation of precipitation is also significant. The frequency, duration, and intensity of specific precipitation events and seasonal, annual, and longer-term precipitation patterns all impact the recharging processes. In some circumstances, the length and intensity of a single precipitation event can significantly impact recharging. When precipitation rates in Uttar Pradesh surpass evapotranspiration rates, the circumstances are best for water to drain through the unsaturated zone to the water table (Figure **8**).

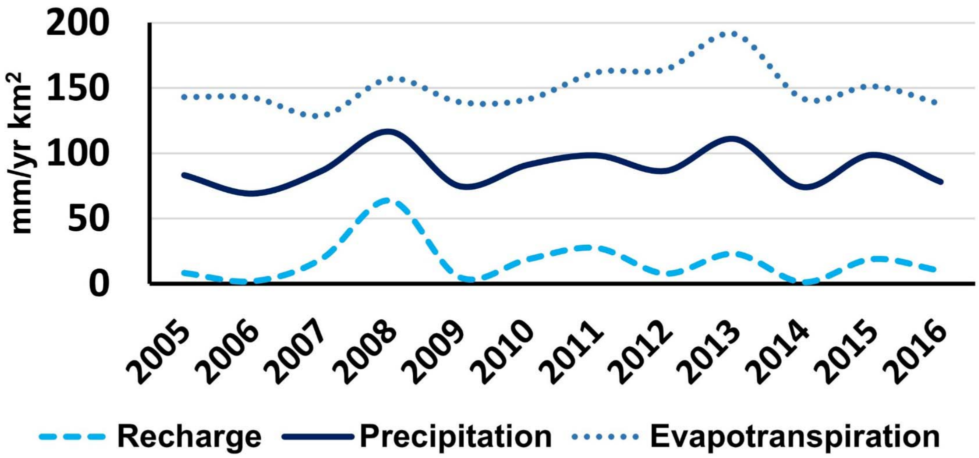


Figure 8. Average annual diffuse recharge, precipitation, and evapotranspiration, for the 12 years from 2005 to 2016 in Uttar Pradesh.

Thus, we have estimated the spatial distribution of the annual SPI and annual groundwater recharge index (GRI). In (Figure **9**), Annual SPI and annual GRI follow the same pattern spatially and quantificationally. 2005, 2006, 2009, 2014, and 2016 are drought years, and GRI varies between moderate and low. Similarly, GRI is higher in wet years such as 2008 and 2013. The drought years have been followed by wet years in the state. Thus, an extreme reduction in groundwater recharge has not been shown.

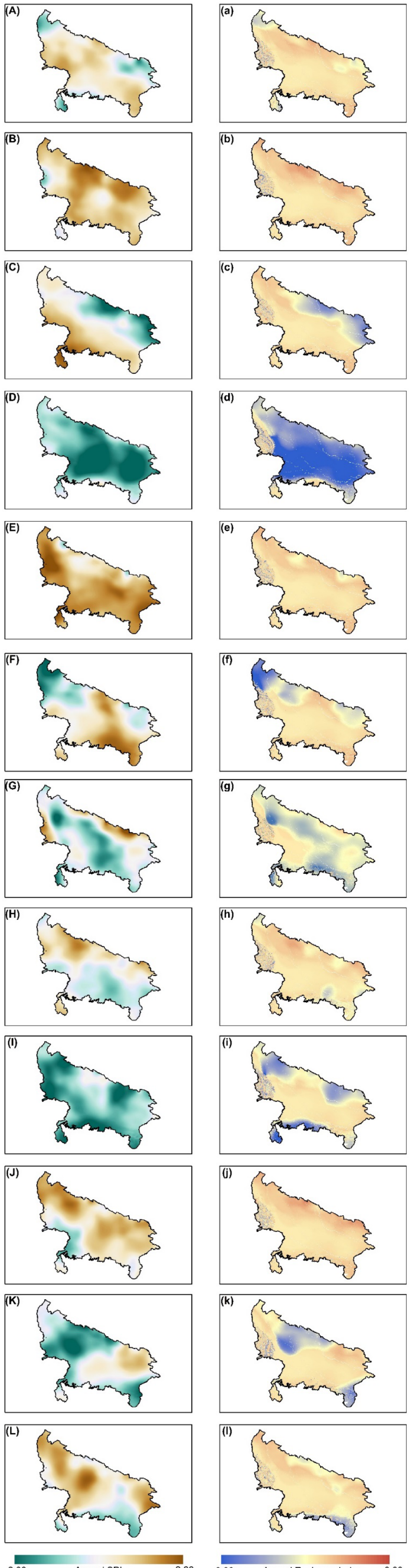
436 
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
(K)
(L)
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
2.00
Annual SPI
-2.00
3.00
Annual Recharge Index
-3.00

Figure 9. Spatial distribution of annual SPI along with annual GRI.

3.4 Comparison with a Recent Global Model of Recharge

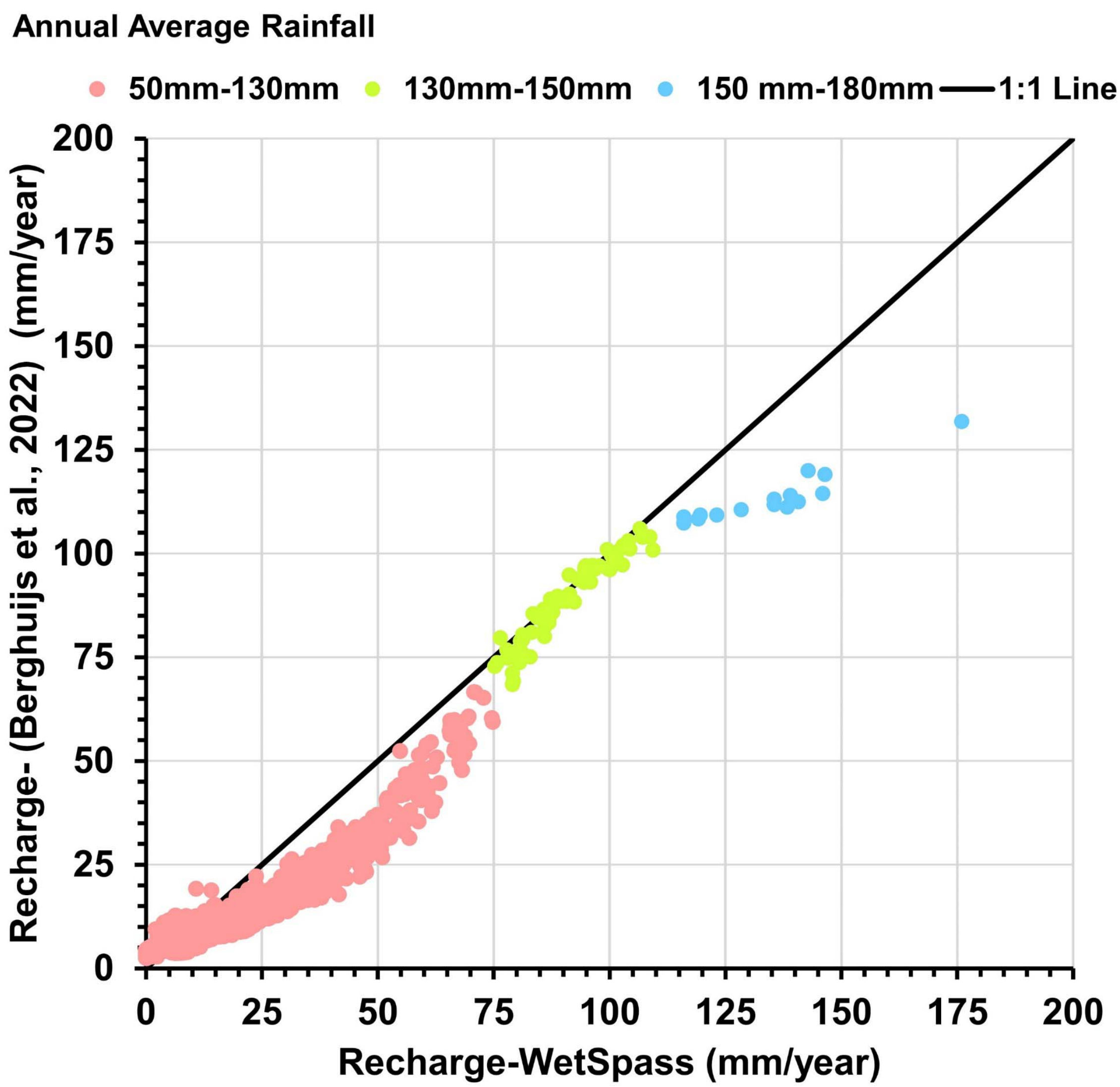


Figure 10. Comparison between simulated recharge from WetSpass and global recharge prediction model of (Berghuijs et al., 2022).

The global recharge prediction model of (Berghuijs et al., 2022)has been derived from the parameterized relationship between empirical recharge observations from 5237 observation sites around the globe and the given global climate aridity (Berghuijs et al., 2022).

$$\frac{R}{P} = \alpha\left(1 - \frac{\ln(\emptyset^{\beta}+1)}{1+\ln(\emptyset^{\beta}+1)}\right) \tag{17}$$

where *R/P* is recharge fraction, ∅ is aridity, α is an equating constant for the fraction of precipitation, which is equal to recharge when ∅ →0.

The R square value 0.415 has been obtained for linear regression analysis between the global model simulated recharge and observed recharge. The widely used global models such as

WATER-GAP, PCR-GLOB, and machine learning have underestimated recharge by 50% compared to absolute values. Compared to such models, the recent global model of (Berghuijs et al., 2022)has shown similar magnitudes for both the simulated and actual recharge estimation due to the sigmoidal relationship (Equation 19) between recharge fraction, and aridity has removed biasing effect.

The global recharge model has been applied to the current study area, and the simulated recharge values have shown a strong correlation as R square 0.93 with simulated recharge values of the WetSpass model. The global model estimated mean annual recharge per 1km^2 over Uttar Pradesh is 13.55 mm/yr, while by WetSpass model is 16.87 mm/yr. The WetSpass and the global model have shown a 1:1 relation for moderate recharge associated with the mean annual precipitation of 130-150mm/yr per 1 km^2 area. For higher and lower recharge, the WetSpass has shown higher estimation than the global model. According to (Berghuijs et al., 2022), the hydrological models underestimate recharge at lower and higher recharge due to biasing. The recharge and discharge zones are both covered by global hydrological models, which mimic hydrological dynamics at several km^2 per grid cell scale (Moeck et al., 2020).

Conversely, most observations will occur in recharge zones, whereas discharge zones typically only cover a small portion of the Earth's surface (Berghuijs et al., 2022; O'Loughlin, 1981). But the WetSpass simulation doesn't associate with discharge estimation. The WetSpass simulation considers the amount of precipitation and the effect of other physiographic and biophysical factors involving infiltration. Therefore, the comparison between the global model and WetSpass has shown that recharge fraction is not a function of the amount of precipitation.

3.5 Correlation between Recharge and Base Flow Index

As aforementioned, though groundwater recharge is necessary for water balance, it is challenging to quantify the magnitude and spatial and temporal variation directly. Large-scale models frequently generalize correlations between the climate and hydrological fluxes and tend to oversimplify processes (Hartmann et al., 2017b). Moreover, the validity of the simulated recharge rates is frequently weak. For instance, despite few recharge measurements for validation, the runoff was categorized into quick surface runoff, slow subsurface runoff, and recharge using a heuristic approach in global modeling research of (Döll & Fiedler, 2008). In previous studies, the estimation of recharge from available streamflow records at gauged basins and the development of a regression equation linking those recharge estimates to the physical and climatic parameters of the gauged basins are two commonly used methods for mapping recharge rates on a statewide scale (Risser et al., n.d.). Hence, the present study has used the Pearson correlation test using the global grided base flow index data of (Beck, Van Dijk, et al., 2013) to test the statistical correlation between simulated recharge and BFI. The strength of the correlation between the two variables is shown by Pearson's correlation coefficient (r). Based on the comparison, base flow and recharge have a strong, substantial Pearson correlation as r is 0.73.

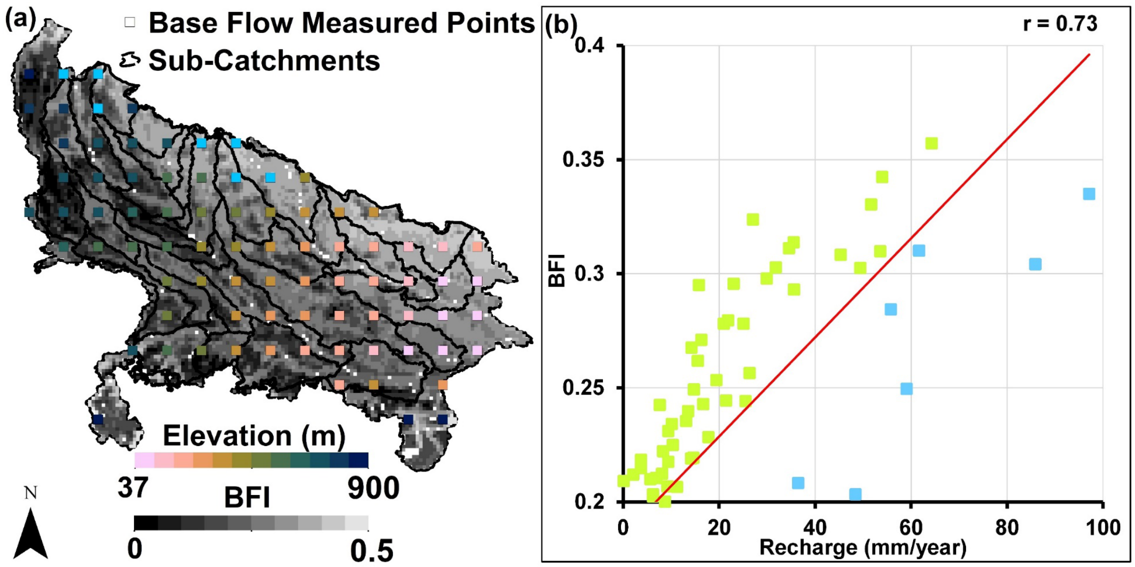


Figure 11. (a) has shown the delineated sub-catchment areas, distribution of BFI over the state, and elevation of selected locations for correlation test between WetSpass estimated recharge and BFI, and the scatter plot (b) shows the relationship between recharge estimates by the WetSpass and BFI. The location and elevation of the cluster (blue) below the 45-degree line (red) have been represented in the same color (blue) in the map (a).

The study has selected 87 locations related to 33 delineated sub-catchments for correlating with BFI. The association between recharge and BFI is correlated for many sub-catchments, as shown in Figure 11(b). However, a significant group (blue points) below the 45-degree line shows higher recharge than BFI. This cluster allows us to study the biased circumstances while comparing WetSpass and BFI. According to the prominent cluster (green) of the scatter plot (Figure 11(b)) has shown significantly moderate to low BFI. The land use of the sub-catchments is dominantly agriculture. Agricultural land use may positively or negatively affect recharge and base flow depending on management practices. In Uttar Pradesh, groundwater irrigation is significant, and this can cause minimal base flow.

The locations of the biased cluster belong to smaller upstream sub-catchments with high topography, thus characterizing deeper groundwater table regions. These characteristics increase the likelihood of surface water dividing into smaller sub-catchments and do not coincide with the deviation of groundwater. Therefore, these sub-catchments have higher recharge than base flow, which indicates that they are net exporters of groundwater. This has been agreed that baseflow is influenced by watershed topography, geomorphology, and climate, according to earlier research (Price, 2011; Zomlot et al., 2015).

3.6 Vegetation and Recharge

Vegetation and Land usage can have a significant impact on recharge processes. Types and densities of vegetation influence evapotranspiration patterns. Plants' efficiency in extracting water from the subsurface depends on how far their roots penetrate the soil. For instance, trees may extract moisture from depths up to several meters (Healy & Scanlon, 2010). In contrast, in Cerrado, Brazil, depending on the extent of the plant roots cause for taking up water from the aquifers during the dry season and similar causing for making way to water back to the aquifer during the rainy season (Tonello & Bramorski, 2021).

Due to the increased surface storage component, vegetation can intercept rainfall with its leaves and branches, which affects evapotranspiration and lengthens the time it takes for the soil to recharge (Jyrkama et al., 2002). The growth cycle of crops is frequently described using long-term variations of vegetation indices (Gorelick et al., 2017), such as the NDVI. The NDVI is an index that measures how green the vegetation (Peng et al., 2011)is and is a good indicator of how the vegetation in one zone has changed over time (Fu & Burgher, 2015). This indicator is generated using the difference between the near-infrared and visible (red) bands (Bulcock & Jewitt, 2010; Otto et al., 2011)and is based on the reflectance of differential that trees, shrubs, and plants exhibit for various sections of the solar radiation spectrum (Mohajane et al., 2018).

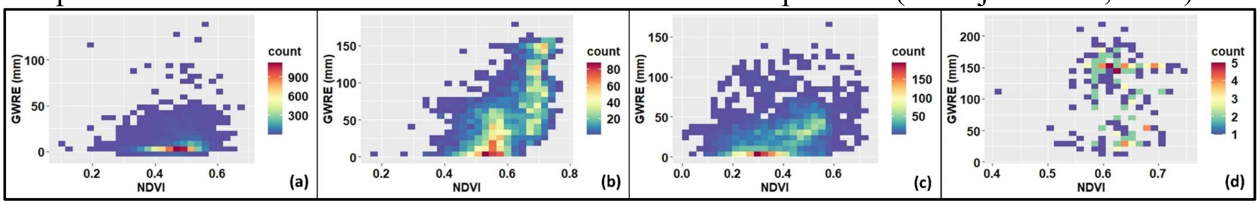


Figure 12. 2D frequency diagram between NDVI and groundwater recharge (GWRE). (a), (b), (c), and (d) are agricultural, forest, grass, and orchard areas, respectively. Annual NDVI and annual GWRE have been considered for (a), (b), and (c), whereas seasonal NDVI and seasonal GWRE has been plotted for (a) due to summer being the rainwater irrigation season of the state.

Agriculture, forest (including intermediate tree cover), and grass areas have shown significant relation between NDVI and GWRE. But orchard areas have shown scattered distribution and not shown significant clusters related to NDVI vs. GWRE. Agricultural lands have shown GWRE varying only from 25-50 mm/km^2 during variation of NDVI. This represents the same crop type that spreads over a large area, such as Paddy in the summer (Karif season). But in the forest (including intermediate tree cover) and grass areas, GWRE has varied over an extended range from 25 to 150 mm/km^2 while the variation of NDVI. Though, in agricultural areas, GWRE has been constant over the increasing of NDVI (<0.4), in forest and grass areas, GWRE is lowering while increasing of NDVI (<0.4). In grass areas is shown a cluster NDVI≥0.6 where GWRE increases. This cluster vegetation can correlate with grasses which have shallow-rooted and cannot access soil water from higher depths. This is typical of semi-arid regions where GWRE rates have been enhanced by vegetation with shallow-root systems (Healy & Scanlon, 2010).

3.7 Identification of Controlling Factors

The 2D frequency analysis followed by the Pearson correlation test between recharge and selected watershed characteristics, including precipitation, potential evapotranspiration, groundwater level, NDVI of agriculture, forest, grass, orchard areas, and saturated K_s. Then it has extended to the PCA due to the single value of the correlation coefficient cannot identify the hidden structure of the spatial correlation between recharge and watershed characteristics.

Table 1. Pearson correlation coefficients between recharge and selected variables involved in controlling recharge.

Variables	Pearson correlation coefficients between recharge and variable
Precipitation	0.79
Potential evapotranspiration	0.56
Groundwater Level	-0.16

Agriculture	0.18
Forest	0.45
Grass	0.34
Orchard	-0.14
Soil	-0.07

In general, we discovered strong correlations across variables, leading us to believe they are redundant. More variables correlating significantly with hydrometeorological factors can be seen in groundwater recharge. The relationship between precipitation and groundwater recharge is strong and positive. This conclusion is a general norm demonstrated in numerous groundwater recharge investigations (Edmunds & Gaye, 1994; Jan et al., 2007; Zomlot et al., 2015). Due to a correlation between the temperature gradient and potential evapotranspiration, recharging was positively associated with potential evapotranspiration (determining PET significantly). In the state, after precipitation, the second significant controlling factor for recharge is vegetation, whereas agriculture, forest, and grass areas have shown contrasting correlations with recharge. As confirmed by the 2D frequency analysis, Orchard areas have not shown a prominent correlation with recharge.

As aforementioned, PCA decreases the redundancy of the data and identifies the significant controlling factors of variance of groundwater recharge by reducing the multidimensional distribution of the data set into a few "artificial variables" such as PC1, PC2, PC3, and PC4, etc. (Zomlot et al., 2015).

The PCA analysis between the R square value between recharge, precipitation, potential evapotranspiration, NDVI, and K_s has shown that the second most important controlling factor for groundwater recharge is vegetation in agriculture, forest, and grass, but not in orchard areas.

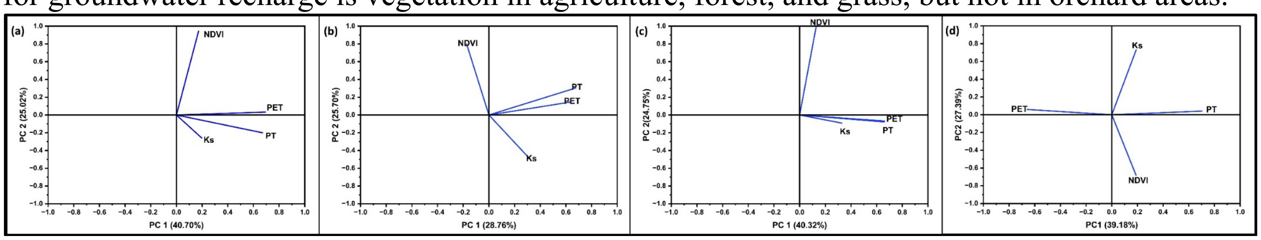


Figure 13. PC1 vs. PC 2 in (a) agriculture, (b) forest, (c) grass, and (d) orchard areas.

3.8 Soil and Recharge

Processes for recharge can be significantly impacted by the permeabilities of surface and subsurface materials. In contrast to places with fine-grained, low-permeability soils, recharge is more likely to occur in areas with coarse-grained, high-permeability soils. The permeability of coarse-grained soils is generally high and may flow water quickly. Hence, water can quickly permeate and drain through the root zone before being extracted by plant roots, the presence of these soils encourages recharge. Even though they are less porous, finer-grained sediments can store more water. Therefore, compared to an area with coarser-grained sediments, one might anticipate less infiltration, improved surface runoff, increased plant extraction of water from the unsaturated zone, and decreased recharge. Permeability is crucial for focused recharge as well. Streambeds with high permeability make it easier for groundwater and surface water to exchange.

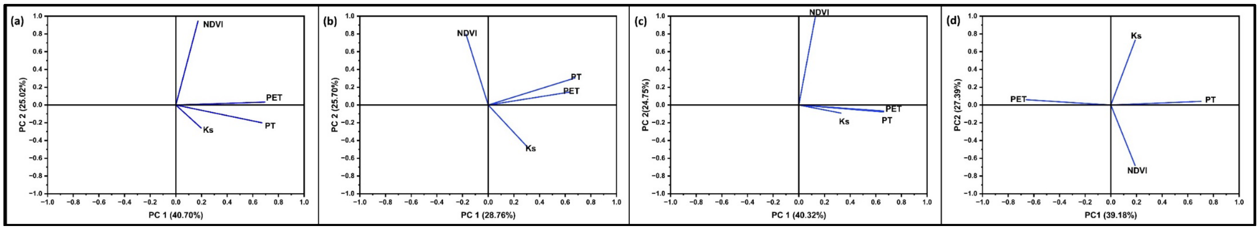


Figure 14. PC 2 vs. PC 3 in (a) agriculture, (b) forest, (c) grass, and (d) orchard areas.

The state is situated on the Central Ganga plain, where drainage conditions and climatic characteristics govern the characteristics of alluvial soil. The typical soil cover texture of the state is Sandy loam and Loam which have a typical permeability range of 300-1800 mm/day and 200-500 mm/day, respectively. Though moderate permeability of soil is involved in easing recharge due to less soil texture variation, the influence on the spatial distribution of recharge is less significant than the effect of vegetation types. This is a recognition that vegetation cover can enhance hydraulic conductivity and minimize the overland flow (Bruijnzeel, 2004; Ilstedt et al., 2016).

3.9 Hidden Clusters in PCA

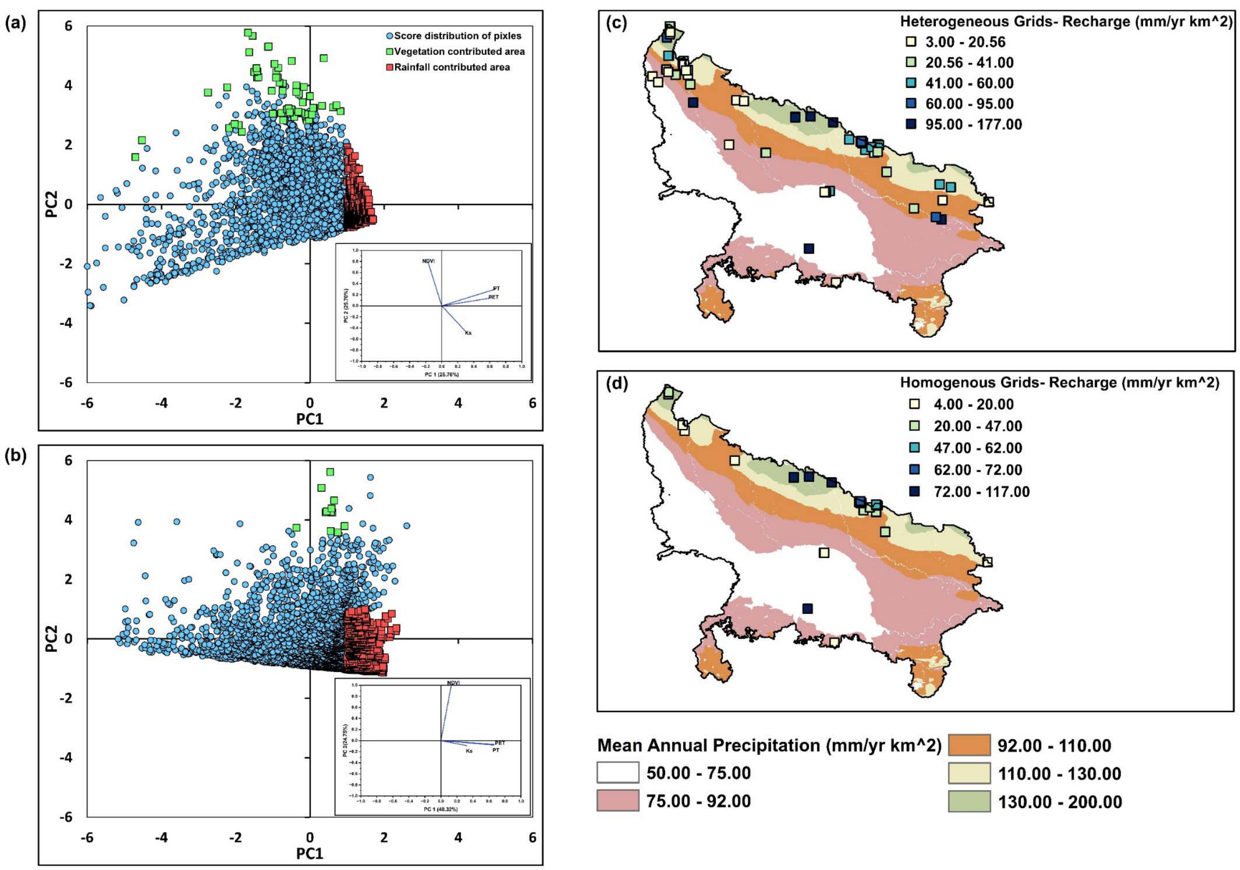


Figure 15. Loading value plots between PC1 vs. PC2 for (a) forest and (b) grass areas. Maps show the selected vegetated contributed areas, (c) and (d) is the distribution of heterogeneous grids, including 74 grids, and homogenous grids, including 28 grids, respectively.

The PCA tries to find a meaningful way to flatten the data by focusing on the parameters with different qualities or influences. PC1 is the axis that spans the most variation, PC2 is the axis that spans the second most variation, and so on. Eigenvectors express the contribution of different parameters to PCs. All the vegetated areas, such as agriculture, forest, grass, and orchard, have significantly span R square values between recharge and precipitation (P-GR) along PC1. The extracted eigenvectors of PC1 for (P-GR) are 0.67, 0.68, 0.66, and 0.7 for agriculture, forest, grass, and vegetated areas, respectively. Excluding orchard areas, agriculture, forest, and grass areas have obtained the high extracted eigenvectors for PC 2 from R square values between recharge and NDVI (NDVI-GR), such as 0.94, 0.80, and 0.99, respectively. We have plotted PC1 versus PC2 accounting for the loading values (x) for each grid in forest and grass areas to identify different clusters with different influences for groundwater recharge. We have excluded agricultural areas from this analysis because WetSpass models have not been simulated with consideration of irrigation water and different crop types in agricultural lands. From the loading value plots of PC1 versus PC2, we have clusters positively influencing clusters for PC1 and PC2. The selected clusters were further refined as vegetation-contributed areas satisfying the conditions of loading value for PC1 is less than 1($x<PC1$), loading value for PC2 is higher than 1 ($x>PC2$), and NDVI-GR is equal or more than 0.5 ($R^2>0.5$). Seventy-four grids from forest and grass areas that satisfy the aforementioned condition have been selected. Among the selected grids, including heterogenous and homogenous grids, dominant vegetated grids, called homogenous grids, have been selected for water budget estimation.

3.10 Simulated Water Budget

The water balance components precipitation, recharge, and evaporation in selected homogenous grids have been extracted from simulated results. Compared to forest areas, intermediate tree cover areas have significantly contributed to groundwater recharge by exceeding evapotranspiration.

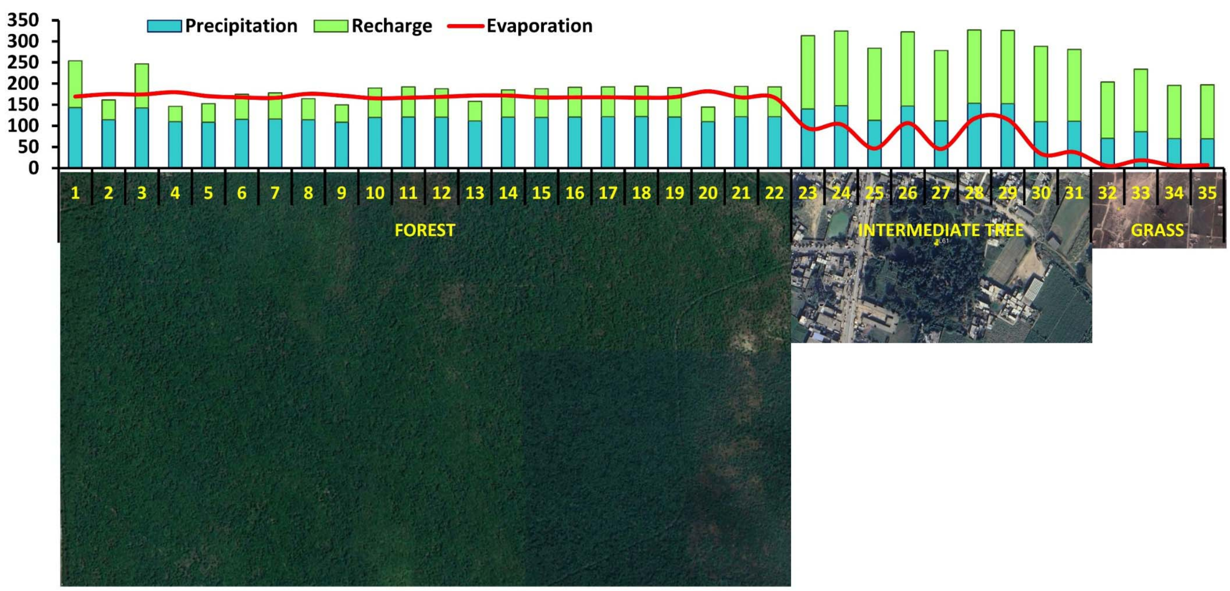


Figure 16. Fractionation of precipitation to recharge and evaporation in different vegetative areas. Satellite images have been acquired from Google Earth Imagery.

This can be explained based on the "optimum tree cover theory," where recharge is enhanced by intermediated tree cover density. Below this optimum tree density, any additional trees' water-percolation benefits outweigh their additional water use, increasing groundwater recharge. In contrast, the opposite happens above the optimum (forest, including dominant and co-dominant)(Ilstedt et al., 2016). The detailed water budget of homogenous grids in forest and intermediate tree cover areas have been considered.

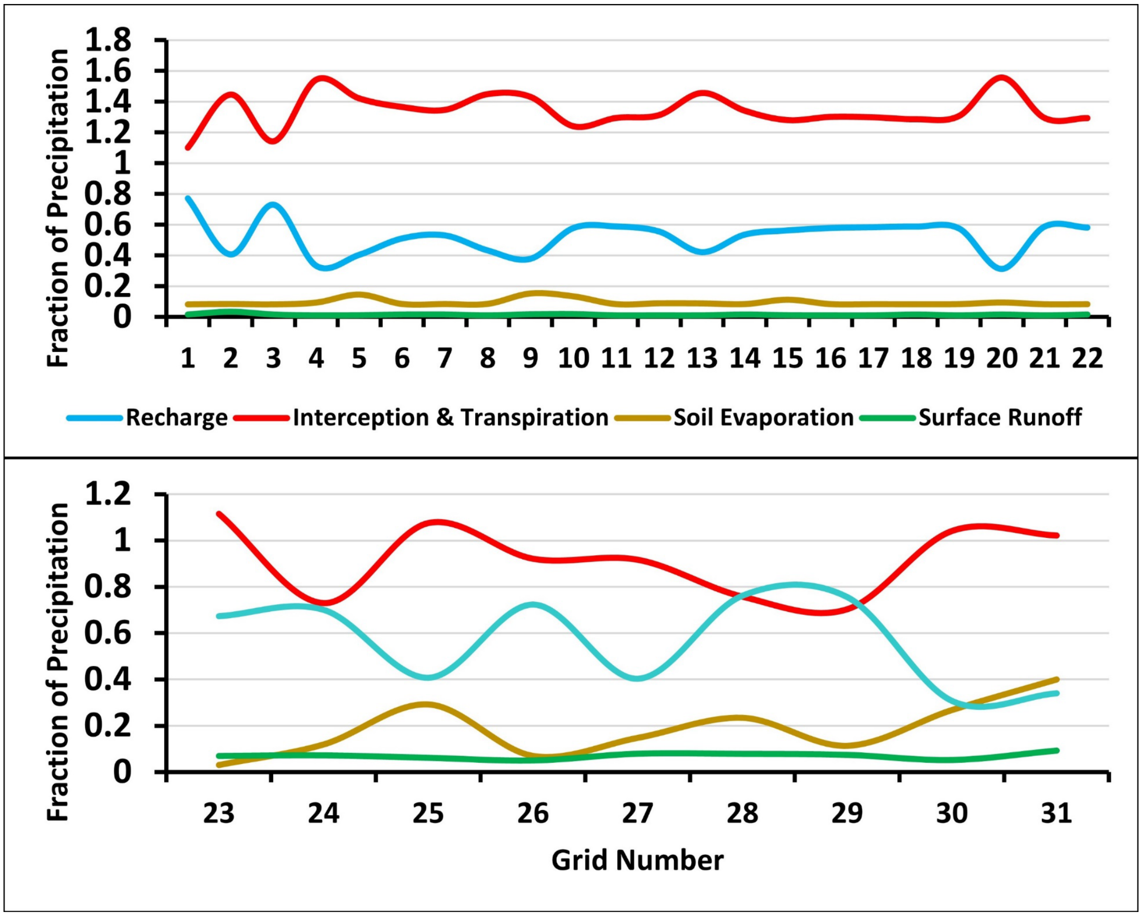


Figure 17. Simulated water budget in (a) Forest areas and (b) intermediate tree cover.

Forest and intermediate tree cover areas typically have less soil evaporation and surface runoff. Without trees, considerable soil and surface runoff causes limited groundwater recharge despite low transpiration. In closed productive forests, overall transpiration and interception are high despite limited surface runoff and soil evaporation, which again results in low groundwater recharge. Low surface runoff, evaporation, and intermediate transpiration rates at an intermediate canopy cover maximize groundwater recharge.

4 Conclusions

Based on the work accomplished in this thesis, it can be concluded that the parameters such as geology, hydrology, climate, soils, slope, vegetation, and land usage in any region play a significant role in controlling the recharge processes in that region. The WetSpass model is a reliable model as it can incorporate the aforementioned hydrological, physiographic, and biophysical factors into long-term direct natural groundwater recharge estimation. Although, the availability of the essentially required inputs (P, PET, GWL, etc.) influencing the groundwater data in the WetSpass model is challenging, the Remote Sensing data, could very well be used with high reliability for prediction of groundwater recharge, and with validation of the same predictions available from Global Model Data. This approach to apply the modelled output in the data-scarce regions and concerning larger areas, and minimize the research gap, has turned out to be highly encouraging. Our analysis of the correlations between the above-mentioned principal components demonstrates that the most significant predictors of groundwater recharge rates are climatic forcing factors, namely, the 'annual precipitation' and the 'potential evapotranspiration'. The magnitude of recharging rates is very well understood to be strongly influenced by the quantity of precipitation or substantial cyclicity in the climatic driving functions. Therefore, the strong correlation and dependence of recharge rates on the above-mentioned climatic forcing factors indicate that groundwater recharge would be highly susceptible to the anticipated change in climate, limited to the exposure from the variation in physiographic and biophysical factors. Vegetation, in general, is showing up as the second most significant parameter for the spatial distribution of groundwater recharge in Uttar Pradesh. However, in some smaller patches, soil texture has become the second most significant controlling factor for groundwater recharge. Hence, the impact of climatic forcing factors on groundwater recharge can vary greatly, depending on the site and therefore, the correlation coefficient between recharge estimations and precipitation, as used in Global Models for one geological setting, may lead to inaccurate prediction in new settings. Our study helps to identify the order of significance of the controlling parameters for groundwater recharge and their overall influence on the spatial distribution of water balance components. In this work, we highlighted the areas with a scarcity of data on groundwater recharge and lacking understanding of the processes influencing groundwater recharge due to knowledge gaps. We hope that some future work will focus on the open-access models and data to close these gaps, improve the global models, share knowledge, and release new recharge data. Also, this study suggests promoting natural recharge-controlling factors, such as establishing particular vegetation species in suitable locations, which can benefit larger communities whose lives depend upon groundwater footprints.

Data Availability Statement

The listed sources provide access to all the data utilized in this study. Precipitation, temperature, potential evapotranspiration, and wind speed data are available at https://chelsa-climate.org/. Soil data is available at https://daac.ornl.gov/. Land use land cover and NDVI data are available at https://ladsweb.modaps.eosdis.nasa.gov/. Global base flow data is downloaded from Global Streamflow Characteristics Dataset (GSDC), available at http://www.gloh2o.org/gscd/. Digital Elevation Model (DEM) is downloaded from https://asf.alaska.edu/. Central Groundwater Board (CGWB), India https://cgwb.gov.in/ is acknowledged for providing groundwater level data. Building footprint data accessed from https://github.com/microsoft/GlobalMLBuildingFootprints

References

Ajami, H. (2021). Geohydrology: Groundwater. In *Encyclopedia of Geology* (pp. 408–415). Elsevier. https://doi.org/10.1016/b978-0-12-409548-9.12388-7

Alley, W. M., Healy, R. W., LaBaugh, J. W., & Reilly, T. E. (2002). Flow and Storage in Groundwater Systems. *Science*, *296*(5575), 1985–1990. https://doi.org/10.1126/science.1067123

Amer, K. H., & Hatfield, J. L. (2004). *Canopy Resistance as Affected by Soil and Meteorological Canopy Resistance as Affected by Soil and Meteorological Factors in Potato Factors in Potato*. Retrieved from https://digitalcommons.unl.edu/usdaarsfacpub/1342

Amiri, M., Salem, A., & Ghzal, M. (2022). Spatial-Temporal Water Balance Components Estimation Using Integrated GIS-Based Wetspass-M Model in Moulouya Basin, Morocco. *ISPRS International Journal of Geo-Information*, *11*(2), 139. https://doi.org/10.3390/ijgi11020139

Ansari, A. A., Singh, I. B., & Tobschall, H. J. (2000). Role of monsoon rain on concentrations and dispersion patterns of metal pollutants in sediments and soils of the Ganga Plain, India. *Environmental Geology*, *39*(3–4), 221–237. https://doi.org/10.1007/s002540050003

Armanuos, A. M., Negm, A., Yoshimura, C., & Valeriano, O. C. S. (2016). Application of WetSpass model to estimate groundwater recharge variability in the Nile Delta aquifer. *Arabian Journal of Geosciences*, *9*(10). https://doi.org/10.1007/s12517-016-2580-x

Arnold, J. G., Muttiah, R. S., Srinivasan, R., & Allen, P. M. (2000). Regional estimation of base flow and groundwater recharge in the Upper Mississippi river basin. *Journal of Hydrology*, *227*(1–4), 21–40. https://doi.org/10.1016/S0022-1694(99)00139-0

Batelaan, O., & De Smedt, F. (2007). GIS-based recharge estimation by coupling surface-subsurface water balances. *Journal of Hydrology*, *337*(3–4), 337–355. https://doi.org/10.1016/j.jhydrol.2007.02.001

Batelaan, Okke, & De Smedt, F. (2001). *WetSpass: a flexible, GIS based, distributed recharge methodology for regional groundwater modelling*. IAHS Publ.

Beck, H. E., Van Dijk, A. I. J. M., Miralles, D. G., De Jeu, R. A. M., Bruijnzeel, L. A., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, *49*(12), 7843–7863. https://doi.org/10.1002/2013WR013918

Beck, H. E., van Dijk, A. I. J. M., Miralles, D. G., de Jeu, R. A. M., Sampurno Bruijnzeel, L. A., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, *49*(12), 7843–7863. https://doi.org/10.1002/2013WR013918

Berghuijs, W. R., Luijendijk, E., Moeck, C., van der Velde, Y., & Allen, S. T. (2022). Global Recharge Data Set Indicates Strengthened Groundwater Connection to Surface Fluxes. *Geophysical Research Letters*, *49*(23). https://doi.org/10.1029/2022GL099010

Best, L. C., & Lowry, C. S. (2014). Quantifying the potential effects of high-volume water extractions on water resources during natural gas development: Marcellus Shale, NY. *Journal of Hydrology: Regional Studies*, *1*, 1–16. https://doi.org/10.1016/j.ejrh.2014.05.001

Bloomfield, J. P., Allen, D. J., & Griffiths, K. J. (2009). Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK. *Journal of Hydrology*, *373*(1–2), 164–176. https://doi.org/10.1016/j.jhydrol.2009.04.025

Bruijnzeel, L. A. (2004). Hydrological functions of tropical forests: not seeing the soil for the trees? *Agriculture, Ecosystems & Environment*, *104*(1), 185–228. https://doi.org/10.1016/j.agee.2004.01.015

Bulcock, H. H., & Jewitt, G. P. W. (2010). Spatial mapping of leaf area index using hyperspectral remote sensing for hydrological applications with a particular focus on canopy interception. *Hydrology and Earth System Sciences*, *14*(2), 383–392. https://doi.org/10.5194/hess-14-383-2010

Chauhan, M. S., Pokharia, A. K., & Srivastava, R. K. (2015). Late Quaternary vegetation history, climatic variability and human activity in the Central Ganga Plain, deduced by pollen proxy records from Karela Jheel, India. *Quaternary International*, *371*, 144–156. https://doi.org/10.1016/j.quaint.2015.03.025

Cooper, D. J., Wolf, E. C., Ronayne, M. J., & Roche, J. W. (2015). Effects of groundwater pumping on the sustainability of a mountain wetland complex, Yosemite National Park, California. *Journal of Hydrology: Regional Studies*, *3*, 87–105. https://doi.org/10.1016/j.ejrh.2014.10.002

Crosbie, R. S., Peeters, L. J. M., Herron, N., McVicar, T. R., & Herr, A. (2018). Estimating groundwater recharge and its associated uncertainty: Use of regression kriging and the chloride mass balance method. *Journal of Hydrology*, *561*, 1063–1080. https://doi.org/10.1016/j.jhydrol.2017.08.003

Delin, G. N., Healy, R. W., Lorenz, D. L., & Nimmo, J. R. (2007). Comparison of local- to regional-scale estimates of groundwater recharge in Minnesota, USA. *Journal of Hydrology*, *334*(1–2), 231–249. https://doi.org/10.1016/j.jhydrol.2006.10.010

Dereje, B., & Nedaw, D. (2019). Groundwater Recharge Estimation Using WetSpass Modeling in Upper Bilate Catchment, Southern Ethiopia. *Momona Ethiopian Journal of Science*, *11*(1), 37. https://doi.org/10.4314/mejs.v11i1.3

Döll, P., & Fiedler, K. (2008). *Hydrology and Earth System Sciences Global-scale modeling of groundwater recharge*. *Hydrol. Earth Syst. Sci* (Vol. 12). Retrieved from www.hydrol-earth-syst-sci.net/12/863/2008/

Eckhardt, K. (2008). A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. *Journal of Hydrology*, *352*(1–2), 168–173. https://doi.org/10.1016/j.jhydrol.2008.01.005

Edmunds, W. M., & Gaye, C. B. (1994). Estimating the spatial variability of groundwater recharge in the Sahel using chloride. *Journal of Hydrology*, *156*(1–4), 47–59. https://doi.org/10.1016/0022-1694(94)90070-1

Eilers, V. H. M., Carter, R. C., & Rushton, K. R. (2007). A single layer soil water balance model for estimating deep drainage (potential recharge): An application to cropped land in semi-arid North-east Nigeria. *Geoderma*, *140*(1–2), 119–131. https://doi.org/10.1016/j.geoderma.2007.03.011

Fu, B., & Burgher, I. (2015). Riparian vegetation NDVI dynamics and its relationship with climate, surface water and groundwater. *Journal of Arid Environments*, *113*, 59–68. https://doi.org/10.1016/j.jaridenv.2014.09.010

Gebert, W. A., Radloff, M. J., Considine, E. J., & Kennedy, J. L. (2007). Use of Streamflow Data to Estimate Base Flow/Groundwater Recharge For Wisconsin1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 220–236. https://doi.org/10.1111/j.1752-1688.2007.00018.x

Gleeson, T., Wada, Y., Bierkens, M. F. P., & Van Beek, L. P. H. (2012). Water balance of global aquifers revealed by groundwater footprint. *Nature*, *488*(7410), 197–200. https://doi.org/10.1038/nature11295

Gleeson, T., Cuthbert, M., Ferguson, G., & Perrone, D. (2020). Global Groundwater Sustainability, Resources, and Systems in the Anthropocene. *Annual Review of Earth and Planetary Sciences*, *48*(1), 431–463. https://doi.org/10.1146/annurev-earth-071719-055251

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

de Graaf, I. E. M., Gleeson, T., (Rens) van Beek, L. P. H., Sutanudjaja, E. H., & Bierkens, M. F. P. (2019). Environmental flow limits to global groundwater pumping. *Nature*, *574*(7776), 90–94. https://doi.org/10.1038/s41586-019-1594-4

Graf, R., & Przybyłek, J. (2018). Application of the WetSpass simulation model for determining conditions governing the recharge of shallow groundwater in the Poznań Upland, Poland. *Geologos*, *24*(3), 189–205. https://doi.org/10.2478/logos-2018-0020

Hartmann, A., Gleeson, T., Wada, Y., & Wagener, T. (2017a). Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity. *Proceedings of the National Academy of Sciences*, *114*(11), 2842–2847. https://doi.org/10.1073/pnas.1614941114

Hartmann, A., Gleeson, T., Wada, Y., & Wagener, T. (2017b). Enhanced groundwater recharge rates and altered recharge sensitivity to climate variability through subsurface heterogeneity. *Proceedings of the National Academy of Sciences*, *114*(11), 2842–2847. https://doi.org/10.1073/pnas.1614941114

Healy, R. W., & Scanlon, B. R. (2010). *Estimating Groundwater Recharge*. Cambridge University Press. https://doi.org/10.1017/CBO9780511780745

Hemmings, B., Whitaker, F., Gottsmann, J., & Hughes, A. (2015). Hydrogeology of Montserrat review and new insights. *Journal of Hydrology: Regional Studies*, *3*, 1–30. https://doi.org/10.1016/j.ejrh.2014.08.008

Hughes, A. G., Mansour, M. M., & Robins, N. S. (2008). Evaluation of distributed recharge in an upland semi-arid karst system: the West Bank Mountain Aquifer, Middle East. *Hydrogeology Journal*, *16*(5), 845–854. https://doi.org/10.1007/s10040-008-0273-6

Ilstedt, U., Bargués Tobella, A., Bazié, H. R., Bayala, J., Verbeeten, E., Nyberg, G., et al. (2016). Intermediate tree cover can maximize groundwater recharge in the seasonally dry tropics. *Scientific Reports*, *6*(1), 21930. https://doi.org/10.1038/srep21930

Jan, C.-D., Chen, T.-H., & Lo, W.-C. (2007). Effect of rainfall intensity and distribution on groundwater level fluctuations. *Journal of Hydrology*, *332*(3–4), 348–360. https://doi.org/10.1016/j.jhydrol.2006.07.010

Jing, W., Zhang, P., & Zhao, X. (2019). A comparison of different GRACE solutions in terrestrial water storage trend estimation over Tibetan Plateau. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-018-38337-1

Jolliffe, I. T. (2002). *PCA_book* (Vol. Springer New York). Retrieved from https://link.springer.com/book/10.1007/b98835

Jyrkama, M. I., & Sykes, J. F. (2007). The impact of climate change on spatially varying groundwater recharge in the grand river watershed (Ontario). *Journal of Hydrology*, *338*(3–4), 237–250. https://doi.org/10.1016/j.jhydrol.2007.02.036

Jyrkama, M. I., Sykes, J. F., & Normani, S. D. (2002). Recharge Estimation for Transient Ground Water Modeling. *Ground Water*, *40*(6), 638–648. https://doi.org/10.1111/j.1745-6584.2002.tb02550.x

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., et al. (2017). Climatologies at high resolution for the Earth's land surface areas. *Scientific Data*, *4*(1), 170122. https://doi.org/10.1038/sdata.2017.122

Keese, K. E., Scanlon, B. R., & Reedy, R. C. (2005). Assessing controls on diffuse groundwater recharge using unsaturated flow modeling. *Water Resources Research*, *41*(6). https://doi.org/10.1029/2004WR003841

Kim, J. H., & Jackson, R. B. (2012). A Global Analysis of Groundwater Recharge for Vegetation, Climate, and Soils. *Vadose Zone Journal*, *11*(1). https://doi.org/10.2136/vzj2011.0021ra

Kumar Dinkar, G., Singh, V., Verma, A. K., Dinkar, G. K., Farooqui, S. A., Singh, V. K., & Prabhat, P. (2019). Geology of South and Southwest part of Uttar Pradesh and its Mineral Significance. *Geoscience,Engineering, Environment, and Technology*, *4*. https://doi.org/10.25299/jgeet.2019.4.2-2.2441

Li, B., Rodell, M., Peters-Lidard, C., Erlingis, J., Kumar, S., & Mocko, D. (2021). Groundwater Recharge Estimated by Land Surface Models: An Evaluation in the Conterminous United States. *Journal of Hydrometeorology*, *22*(2), 499–522. https://doi.org/10.1175/JHM-D-20-0130.1

Longobardi, A., & Villani, P. (2008). Baseflow index regionalization analysis in a mediterranean area and data scarcity context: Role of the catchment permeability index. *Journal of Hydrology*, *355*(1–4), 63–75. https://doi.org/10.1016/j.jhydrol.2008.03.011

MacDonald, A. M., Lark, R. M., Taylor, R. G., Abiye, T., Fallas, H. C., Favreau, G., et al. (2021). Mapping groundwater recharge in Africa from ground observations and implications for water security. *Environmental Research Letters*, *16*(3). https://doi.org/10.1088/1748-9326/abd661

Markstrom, S. L., Niswonger Richard G., Regan, R. S., Prudic, D. E., & Barlow, P. M. (2008). *GSFLOW - Coupled Ground-Water and Surface-Water Flow Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005)*.

Mazvimavi, D., Meijerink, A. M. J., Savenije, H. H. G., & Stein, A. (2005). Prediction of flow characteristics using multiple regression and neural networks: A case study in Zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, *30*(11–16), 639–647. https://doi.org/10.1016/j.pce.2005.08.003

Meresa, E., Girmay, A., & Gebremedhin, A. (2019). Water Balance Estimation Using Integrated GIS-Based WetSpass Model in the Birki Watershed, Eastern Tigray, Northern Ethiopia. *Physical Science International Journal*, 1–17. https://doi.org/10.9734/psij/2019/v22i330133

Minor, T. B., Russell, C. E., & Mizell, S. A. (2007). Development of a GIS-based model for extrapolating mesoscale groundwater recharge estimates using integrated geospatial data sets. *Hydrogeology Journal*, *15*(1), 183–195. https://doi.org/10.1007/s10040-006-0109-1

Moeck, C., Grech-Cumbo, N., Podgorski, J., Bretzler, A., Gurdak, J. J., Berg, M., & Schirmer, M. (2020). A global-scale dataset of direct natural groundwater recharge rates: A review of variables, processes and relationships. *Science of the Total Environment*, *717*. https://doi.org/10.1016/j.scitotenv.2020.137042

Mohajane, M., Essahlaoui, A., Oudija, F., El Hafyani, M., Hmaidi, A. El, El Ouali, A., et al. (2018). Land Use/Land Cover (LULC) Using Landsat Data Series (MSS, TM, ETM+ and OLI) in Azrou Forest, in the Central Middle Atlas of Morocco. *Environments*, *5*(12), 131. https://doi.org/10.3390/environments5120131

Mohan, C., Western, A. W., Wei, Y., & Saft, M. (2018). Predicting groundwater recharge for varying land cover and climate conditions – a global meta-study. *Hydrology and Earth System Sciences*, *22*(5), 2689–2703. https://doi.org/10.5194/hess-22-2689-2018

Müller Schmied, H., Caceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., et al. (2021). The global water resources and use model WaterGAP v2.2d: Model description and evaluation. *Geoscientific Model Development*, *14*(2), 1037–1079. https://doi.org/10.5194/gmd-14-1037-2021

O'Loughlin, E. M. (1981). Saturation regions in catchments and their relations to soil and topographic properties. *Journal of Hydrology*, *53*(3–4), 229–246. https://doi.org/10.1016/0022-1694(81)90003-2

Oosterbaan, R. J. (1994). *6 FREQUENCY AND REGRESSION ANALYSIS OF HYDROLOGIC DATA*. (H. P. Ritzema, Ed.) (3rd ed.). Wageningen, The Netherlands. Retrieved from www.waterlog.info

Otto, M., Scherer, D., & Richters, J. (2011). Hydrological differentiation and spatial distribution of high altitude wetlands in a semi-arid Andean region derived from satellite data. *Hydrology and Earth System Sciences*, *15*(5), 1713–1727. https://doi.org/10.5194/hess-15-1713-2011

Owuor, S. O., Butterbach-Bahl, K., Guzha, A. C., Rufino, M. C., Pelster, D. E., Díaz-Pinés, E., & Breuer, L. (2016, December 1). Groundwater recharge rates and surface runoff response to land use and land cover changes in semi-arid environments. *Ecological Processes*. Springer Verlag. https://doi.org/10.1186/s13717-016-0060-6

Park, C., Seo, J., Lee, J., Ha, K., & Koo, M.-H. (2014). A distributed water balance approach to groundwater recharge estimation for Jeju volcanic island, Korea. *Geosciences Journal*, *18*(2), 193–207. https://doi.org/10.1007/s12303-013-0063-6

Peng, S., Chen, A., Xu, L., Cao, C., Fang, J., Myneni, R. B., et al. (2011). Recent change of vegetation growth trend in China. *Environmental Research Letters*, *6*(4), 044027. https://doi.org/10.1088/1748-9326/6/4/044027

Piggott, A. R., Moin, S., & Southam, C. (2005). A revised approach to the UKIH method for the calculation of baseflow / Une approche améliorée de la méthode de l'UKIH pour le calcul de l'écoulement de base. *Hydrological Sciences Journal*, *50*(5). https://doi.org/10.1623/hysj.2005.50.5.911

Pokharia, A. K., Sharma, S., Tripathi, D., Mishra, N., Pal, J. N., Vinay, R., & Srivastava, A. (2017). Neolithic−Early historic (2500–200 BC) plant use: The archaeobotany of Ganga Plain, India. *Quaternary International*, *443*, 223–237. https://doi.org/10.1016/j.quaint.2016.09.018

Price, K. (2011). Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review. *Progress in Physical Geography: Earth and Environment*, *35*(4), 465–492. https://doi.org/10.1177/0309133311402714

Rajab, Jasim M., MatJafri, M. Z., & Lim, H. S. (2013). Combining multiple regression and principal component analysis for accurate predictions for column ozone in Peninsular Malaysia. *Atmospheric Environment*, *71*, 36–43. https://doi.org/10.1016/j.atmosenv.2013.01.019

Rajab, Jasim Mohammed, Jafri, Mohd. Z. M., Lim, H. S., & Abdullah, K. (2012). Regression analysis in modeling of air surface temperature and factors affecting its value in Peninsular Malaysia. *Optical Engineering*, *51*(10), 101702. https://doi.org/10.1117/1.OE.51.10.101702

Risser, D. W., Thompson, R. E., & Stuckey, M. H. (n.d.). *Regression Method for Estimating Long-Term Mean Annual Groundwater Recharge Rates from Base Flow in Pennsylvania Prepared in cooperation with the Pennsylvania Department of Conservation and Natural*

Resources, Bureau of Topographic and Geologic Survey Scientific Investigations Report 2008-5185. Retrieved from http://www.usgs.gov/pubprod

Rutledge, A. T. (2005). The appropriate use of the Rorabaugh model to estimate ground water recharge. *Ground Water*, *43*(3), 292–293. https://doi.org/10.1111/j.1745-6584.2005.0022.x

Sadeak, S., & Khan, M. (2021). Sustainable Groundwater Management by Geological Modeling and Recharge Estimation to Optimize Sudden Stress on the Geologically Complex Aquifer System in the South-Eastern Hilly Region of Bangladesh. New Orleans, LA: AGU Fall Meeting 2021.

Sarkar, T., Karunakalage, A., Kannaujiya, S., & Chaganti, C. (2022). Quantification of groundwater storage variation in Himalayan & Peninsular River basins correlating with land deformation effects observed at different Indian cities. *Contributions to Geophysics and Geodesy*, *52*(1), 1–56. https://doi.org/10.31577/CONGEO.2022.52.1.1

Scanlon, B. R., Healy, R. W., & Cook, P. G. (2002). Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, *10*(1), 18–39. https://doi.org/10.1007/s10040-001-0176-2

Scanlon, B. R., Keese, K. E., Flint, A. L., Flint, L. E., Gaye, C. B., Edmunds, W. M., & Simmers, I. (2006). Global synthesis of groundwater recharge in semiarid and arid regions. *Hydrological Processes*, *20*(15), 3335–3370. https://doi.org/10.1002/hyp.6335

Singh, I. B. (1996). GEOLOGICAL EVOLUTION OF GANGA PLAIN : AN OVERVIEW. *Journal of the Palaeontological Society of India*, *41*, 99–137.

Sophocleous, M. (2005). Groundwater recharge and sustainability in the High Plains aquifer in Kansas, USA. *Hydrogeology Journal*, *13*(2), 351–365. https://doi.org/10.1007/s10040-004-0385-6

Tonello, K. C., & Bramorski, J. (2021). Can Trees Help Put Water Back Into the Soil? *Frontiers for Young Minds*, *9*. https://doi.org/10.3389/frym.2021.589362

Umar, R. (2006). Hydrogeological environment and groundwater occurrences of the alluvial aquifers in parts of the Central Ganga Plain, Uttar Pradesh, India. *Hydrogeology Journal*, *14*(6), 969–978. https://doi.org/10.1007/s10040-005-0019-7

de Vries, J. J., & Simmers, I. (2002). Groundwater recharge: An overview of process and challenges. *Hydrogeology Journal*, *10*(1), 5–17. https://doi.org/10.1007/s10040-001-0171-7

Wang, T., Istanbulluoglu, E., Lenters, J., & Scott, D. (2009). On the role of groundwater and soil texture in the regional water balance: An investigation of the Nebraska Sand Hills, USA. *Water Resources Research*, *45*(10). https://doi.org/10.1029/2009WR007733

Yenehun, A., Dessie, M., Nigate, F., Belay, S., Azeze, M., Van Camp, M., et al. (2021). Spatial and temporal simulation of groundwater recharge and cross-1 validation with point measurements in volcanic aquifers with variable 2 topography 3. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-2021-527

Zhang, Y., Liu, S., Cheng, F., & Shen, Z. (2017). WetSpass-based study of the effects of urbanization on the water balance components at regional and quadrat scales in Beijing, China. *Water (Switzerland)*, *10*(1). https://doi.org/10.3390/w10010005

Zomlot, Z., Verbeiren, B., Huysmans, M., & Batelaan, O. (2015). Spatial distribution of groundwater recharge and base flow: Assessment of controlling factors. *Journal of Hydrology: Regional Studies*, *4*, 349–368. https://doi.org/10.1016/j.ejrh.2015.07.005

S-1. Climatic and physiographic data requirements for WetSpass model.

Dataset	Units (SI)	Data Source	Data Model/Version	Initial Spatial Resolution	Spatial Resolution for WetSpass Model	Initial Temporal Resolution	Temporal Resolution for WetSpass Model
Precipitation	mm	Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA-Climate)	CHELSA V2.1	830m×830m	1km × 1km	Monthly	Seasonal; Winter & Summer
Potential evapotranspiration	mm	Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA-Climate)	CHELSA V2.1	830m×830m	1km × 1km	Monthly	Seasonal; Winter & Summer
Temperature	C^{0}	Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA-Climate)	CHELSA V2.1	830m×830m	1km × 1km	Monthly	Seasonal; Winter & Summer
Wind Speed	m/s	Climatologies at High Resolution for the Earth's Land Surface Areas (CHELSA-Climate)	CHELSA V2.1	830m×830m	1km × 1km	Monthly	Seasonal; Winter & Summer
Groundwater depth	m	Central Groundwater Board	NA	NA (Point data)	1km × 1km	Seasonal	Seasonal; Winter & Summer
Topography	m	Alaska Satellite Facility	ALOS World 3D - 30m (AW3D30)	27m×27m	1km × 1km	NA	NA
Slope	%	Alaska Satellite Facility	ALOS World 3D - 30m (AW3D30)	27m×27m	1km × 1km	NA	NA
Soil texture	NA	NASA Distributed Active Archive Centre for Biochemical Dynamics (ORNL DAAC)	NA	1km × 1km	1km × 1km	NA	NA
Land use land cover	NA	Level-1 and Atmosphere Archive & Distribution system Distributed Active Archive Centre (LAADS DAAC)	Level-1	500m×500m	1km × 1km	Yearly	Yearly

S-2. Attribute data required for the WetSpass model.

Table Name	Data Columns	
Soil attributes	Field capacity	
	Wilting point	
	Plant-available water content	
	Residual water content	
	Bare soil evaporation depth	
	Tension saturated height	
	Fraction of summer precipitation contributing to Hortonian runoff	
	Fraction of winter precipitation contributing to Hortonian runoff	
Land use land cover attributes	Aerial fractions for each land use type	Vegetated area
		Bare soil area
		Open water area
		Impervious surface
	Root depth	
	Leaf area index	
	Minimum stomatal opening	
	Interception percentage	
	Vegetation height	
Runoff coefficient attributes	Bare soil runoff coefficient	
	Runoff coefficient for impervious land use type	

S-3. γ / Δ is a function of temperature, as given in the table below.

T (^{0}C)	-20	-10	0	5	10	15	20	25	30	35	40
γ / Δ	5.86	2.83	1.46	1.07	0.76	0.59	0.45	0.35	0.27	0.25	0.17