

Reconstructing high-resolution gridded precipitation data using an improved downscaling approach over the high altitude mountain regions of Upper Indus Basin (UIB)



Arfan Arshad ^{a,b,*}, Wanchang Zhang ^{a,b,*}, Zhijie Zhang ^c, Shuhang Wang ^d, Bo Zhang ^d, Muhammad Jehanzeb Masud Cheema ^e, Masoud Jafari Shalamzari ^{a,b}

^a Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China

^b University of Chinese Academy of Sciences, Beijing 100049, China

^c Department of Geography, University of Connecticut, Storrs CT06269, USA

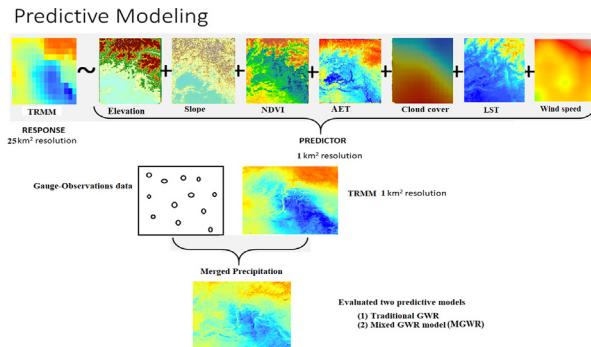
^d National Engineering Laboratory for Lake Pollution Control and Ecological Restoration, Chinese Research Academy of Environmental Science, Beijing 100012, China

^e Faculty of Agricultural Engineering and Technology, PMAS-Arid Agriculture University, Rawalpindi, Pakistan

HIGHLIGHTS

- GVT offers a promising method for filtering out both spatially varying and constant relations.
- Newly proposed MGWR model performed better than the traditional GWR model.
- GRA perform better than the GDA during calibration process.
- Downscaled precipitation at 1 km, demonstrated better accuracy than original-TRMM.
- IDAC approach is suitable to retrieve high-resolution gridded precipitation data.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 12 November 2020

Received in revised form 9 April 2021

Accepted 10 April 2021

Available online 16 April 2021

Editor: A P Dimri

Keywords:

Tropical Rainfall Measuring Mission (TRMM)

Downscaling

Geographical variability test (GVT)

Global variables

Mixed geographically weighted regression

ABSTRACT

Understanding the basin-scale hydrology and the spatiotemporal distribution of regional precipitation requires high precision, as well as high-resolution precipitation data. We have made an attempt to develop an Integrated Downscaling and Calibration (IDAC) framework to generate high-resolution ($1 \text{ km} \times 1 \text{ km}$) gridded precipitation data. Traditionally, GWR (Geographical weighted regression) model has widely been applied to generate high-resolution precipitation data for regional scales. The GWR model generally assumes a spatially varied relationships between precipitation and its associated environmental variables, however, the relationships need to remain constant (fixed) for some variables over space. In this study, a Mixed Geographically Weighted Regression (MGWR) model, capable of dealing with the fixed and spatially varied environmental variables, is proposed to downscale the Original-TRMM precipitation data from a coarse resolution ($0.25^\circ \times 0.25^\circ$) to a high-resolution ($1 \text{ km} \times 1 \text{ km}$) for the period of 2000–2018 over the Upper Indus Basin (UIB). Additionally, accuracy of the down-scaled precipitation data was further improved by merging it with the recorded data from rain gauge stations (RGS) using two calibration approaches such as Geographical Ratio Analysis (GRA) and Geographical Difference Analysis (GDA). We found MGWR to perform better given its higher R^2 and lower RMSE and bias values ($R^2 = 0.96$; RMSE = 56.01 mm, bias = 0.014) in comparison to the GWR model ($R^2 = 0.95$; RMSE = 60.76 mm,

* Corresponding author at: Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China.

E-mail addresses: arfanarshad52@mails.ucas.ac.cn (A. Arshad), zhangwc@radi.ac.cn (W. Zhang), zhijie.zhang@uconn.edu (Z. Zhang), wangsh@craes.org.cn (S. Wang), zhangbo@craes.org.cn (B. Zhang), mjm.cheema@uaar.edu.pk (M.J.M. Cheema), msdjafari@radi.ac.cn (M.J. Shalamzari).

(MGWR)
 Gaussian kernel
 Integrated downscaling and calibration (IDAC)

bias = 0.094). It was observed that the GDA and GRA calibrated-downscaled precipitation datasets were superior to the Original-TRMM, yet GRA outperformed GDA. Annual precipitation from downscaled and calibrated-downscaled datasets was further temporally downscaled to obtain high-resolution monthly and daily precipitations. The results revealed that the monthly-downscaled precipitation ($R^2 = 0.82$, bias = -0.02 and RMSE = 11.93 mm/month) and the calibrated-downscaled ($R^2 = 0.89$, bias = -0.006 and RMSE = 9.19 mm/month) series outperformed the Original-TRMM ($R^2 = 0.72$, bias = 0.14 and RMSE = 19.8 mm/month) as compared to the RGS observations. The results of daily calibrated-downscaled precipitation ($R^2 = 0.79$, bias = 0.001 and RMSE = 1.7 mm/day) were better than the Original-TRMM ($R^2 = 0.64$, bias = -0.12 and RMSE = 6.82 mm/day). In general, the proposed IDAC approach is suitable for retrieving high spatial resolution gridded data for annual, monthly, and daily time scales over the UIB with varying climate and complex topography.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Precipitation is an essential component of the global hydrological cycle, playing vital roles in maintaining the hydro-climate balance and ecosystem processes (Langella et al., 2010; Ma et al., 2018) (Li and Shao, 2010; Shi et al., 2015; Ma et al., 2017), and accurate precipitation datasets are crucial for hydro-meteorological applications, especially in data-sparse and Glaciated Catchments (Shafeeqe et al., 2020; Shafeeqe et al., 2019). However, obtaining accurate high spatiotemporal precipitation data is still a challenging task for researchers (Fang et al., 2013). Traditionally, precipitation was measured at rain gauge stations (RGS), which by no doubt provides accurate measurements. However, the spatial distribution of RGS over mountainous regions, especially over the Upper Indus Basin (UIB), is relatively sparse (Gao et al., 2017; Henn et al., 2018; Shafeeqe et al., 2019). Satellite-based remote sensing and assimilation techniques offer a promising solution to this problem (Darand et al., 2017; Yang et al., 2017) providing considerably finer spatial and temporal resolutions compared to RGS observations (Gebremichael et al., 2010). In recent years, many rainfall estimates have been offered ranging in scale from regional to global while enjoying variant time scales. These include the Global Precipitation Climatology Project (GPCP) (Huffman et al., 1997; Huffman et al., 2009), the Precipitation Estimation from Remotely Sensed Information (Sorooshian et al., 2000), the Global Satellite Mapping of Precipitation (GSMP) project (Kubota et al., 2007), the Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007) and Naval Research Laboratory-Blend satellite precipitation estimates (Kidd et al., 2012). Among these satellite-based precipitation estimates, TRMM data has been widely applied in drought monitoring, hydro-meteorological and land surface process modeling (Yaduvanshi et al., 2015; Wang et al., 2016; Seyoum, 2018; Hadria et al., 2019; Shafeeqe and Yi, 2020). There are still some limitations when it comes to the applications of TRMM data over the small watersheds. Firstly, the spatial resolution of this dataset is still coarse at regional scale (Duan and Bastiaanssen, 2013; Ma et al., 2018) and secondly, it overestimated the RGS observations by >20% due to the topographic factor, retrieval algorithms, clouds and insufficient sensitivity between electromagnetic signals (Adhikary et al., 2015; Tang et al., 2016; Ma et al., 2017). Therefore, downscaling and calibration procedures are required to enhance the spatial information and accuracy of TRMM precipitation (Chen et al., 2018).

Numerous statistical methods have been applied to downscale TRMM precipitation from coarse resolution (0.25°) to fine resolution (1 km) using various high-resolution predictor variables (Duan and Bastiaanssen, 2013; Xu et al., 2015). The most commonly used statistical downscaling models are the exponential regression (Immerzeel et al., 2009), multilinear regression (Zheng and Zhu, 2015; Zhang et al., 2018b; Alexakis and Tsanis, 2016; Zhan et al., 2018), an improved multi nonlinear regression (Wang et al., 2019), random forests model (Shi and Song, 2015; Jing et al., 2016), and Artificial Neural Network Model (ANN) (Zhang et al., 2018b). These methods perform based on the assumption that the relationships between precipitation and predictor variables remains unchanged over space (Ma et al., 2017). On the one hand, it doesn't necessarily hold true and the relationships are

varied at different spatial scales (Foody, 2003; Gao and Li, 2011). On the other hand, geographical weightage regression (GWR) model investigates the non-stationary and scale-dependent characteristics of the association among dependent and independent variables (Foody, 2003). It assumes that the relationship between precipitation and its predictor variables is spatially varied (Chen et al., 2014; Xu et al., 2015) and estimates the regression parameters at each geolocation instead of the whole study area (Fotheringham et al., 2003). Geographical Weightage Regression (GWR) model performs better with simple algorithms for downscaling precipitation data compared to sole regression methods (e.g., exponential, linear, and nonlinear regression models) (Zhang et al., 2018b; Chen et al., 2018; Zhan et al., 2018) as well as ANN (Zhang et al., 2020). Therefore, the present study makes an attempt to use the GWR model to predict the TRMM precipitation at a high spatial resolution using a combination of various environmental variables.

Although many efforts have been made to downscale the TRMM precipitation data at finer resolutions, yet some limitations and research gaps remain to be solved in previous studies. The majority of studies mainly focus on the utilization of local land surface variables (elevation, slope, altitude, normalized difference vegetation index (NDVI), enhanced vegetation index (EVI)), and pay less attention to other climate parameters (wind speed, cloud cover, and evapotranspiration). Research outcomes from previous studies suggested that the variations in precipitation distribution is strongly linked with evapotranspiration (Yang et al., 2016), cloud cover (Mishra, 2019), and wind speed (Back and Bretherton, 2005), and provides the necessary justification in changing trend in precipitation along the mountainous regions. Therefore, it is important to introduce these variables in downscaling model to better represent the spatial variability in precipitation patterns. Secondly, previous studies have used the GWR model by assuming a spatial varying relationship among precipitation and all of its independent environmental variables. These relationships may be constant for some environmental covariates, contrary to the variation assumption embodied in the GWR. Considering all environmental variables to be spatially variant in the GWR model could give rise to random noise (Zeng et al., 2016). Therefore it is necessary to identify and isolate the fixed or constant (global terms) and spatial varying (local) variables to improve the accuracy of model prediction. Great efforts were made in the present study to filter out the fixed and varying environmental variables using geographical variability test (GVT) and reintroduced them into the downscaling framework to propose a mixed geographical weightage regression (MGWR) technique. Since all downscaling techniques predict precipitation based on the Original-TRMM data, over and under-estimations present in TRMM data will eventually find its way into the final downscaled precipitation results. Reliable calibration techniques are still required to significantly improve accuracy in estimating precipitation via integrating down-scaled precipitation data with recorded values at rain gauge stations (RGS) (Cheema and Bastiaanssen, 2012). For example, some studies reported that the integrated downscaling and calibration procedures achieved better results through the use of geographical difference

analysis (GDA) (Duan and Bastiaanssen, 2013) and geographical ratio analysis (GRA) (Chen et al., 2018).

In the present study we have retrieved the 'best' precipitation data at daily, monthly and annual time scales through an Integrated Downscaling and Calibration (IDAC) framework over the Upper Indus Basin (UIB). The objectives of this study are to (a) analyze the response of different environmental variables and isolate them into spatially varied and fixed environmental variables using geographical variability test (GVT), (b) downscale annual TRMM precipitation data from coarse resolution ($0.25^{\circ} \times 0.25^{\circ}$) to ($1 \text{ km} \times 1 \text{ km}$) using GWR and MGWR models, (c) to enhance the accuracy in the downscaled precipitation data by merging it with observations from RGS using Geographical Ratio Analysis (GRA) and Differential Analysis (GDA) approaches, and (d) temporal disaggregation of the annual downscaled precipitation to retrieve a monthly and daily precipitation data. The proposed approach assumes that relationships between precipitation and its environmental variables are mixed in nature, such as spatially variable and constant over space; thus, adding these variables after isolating their spatial fixed and constant nature could improve the model's ability to forecast precipitation precisely at high resolution.

2. Material and methods

2.1. Study area

The Indus Basin is a transboundary basin distributed across four countries (Pakistan, India, China, and Afghanistan) with a total area of

approximately 1.12 million km^2 . The study area, as illustrated in Fig. 1, covers the upper Indus basin (UIB) which is situated in the high mountain ranges of the Hindu Kush, Karakoram, Himalaya, and the Tibetan Plateau with an area of ~4,61,698 km^2 (Fig. 1). It is among the most glaciated areas on Earth, with 22,000 km^2 of glacier surface area. Elevation in the study area varies from 177 m to 8600 m above the mean sea level (m.s.l) (Fig. 1a). Digital elevation model (DEM) from the Shuttle Radar Topography Mission (STRM) at 90 m resolution was used to delineate the study area into 11 sub-basins as shown in Fig. 1b. There are 11 major tributaries to UIB's namely, the upper Indus River, Kharmon, Shigar, Shyok, Hunza, Gilgit, Chitral, Kabul, Jhelum, Chenab, Ravi, and Sutlej. Monsoon rainfalls, glaciers, and runoff from snowmelt are the major sources of flow in UIB. The contribution of these sources differ temporally and spatially. Flow regime in the upper watersheds (Hunza, Shyok, Shigar, and Gilgit) mainly arise from melting glaciers while in lower Punjab (Jhelum, Chenab, and Satluj) and Kabul, monsoon rainfalls and snowmelt dominate the flow regime (Hasson et al., 2014). Fig. 1d illustrates the annual precipitation from 2000 to 2018. We selected three climate years during 2000–2018 (i.e., 2002, 2005, and 2010) based on their average annual precipitations to represent the dry, normal, wet and average conditions in the study area.

2.2. Data used in the study

2.2.1. TRMM precipitation data

The Tropical Rainfall Measuring Mission (TRMM), a joint project held by the National Aeronautics and Space Administration (NASA) of

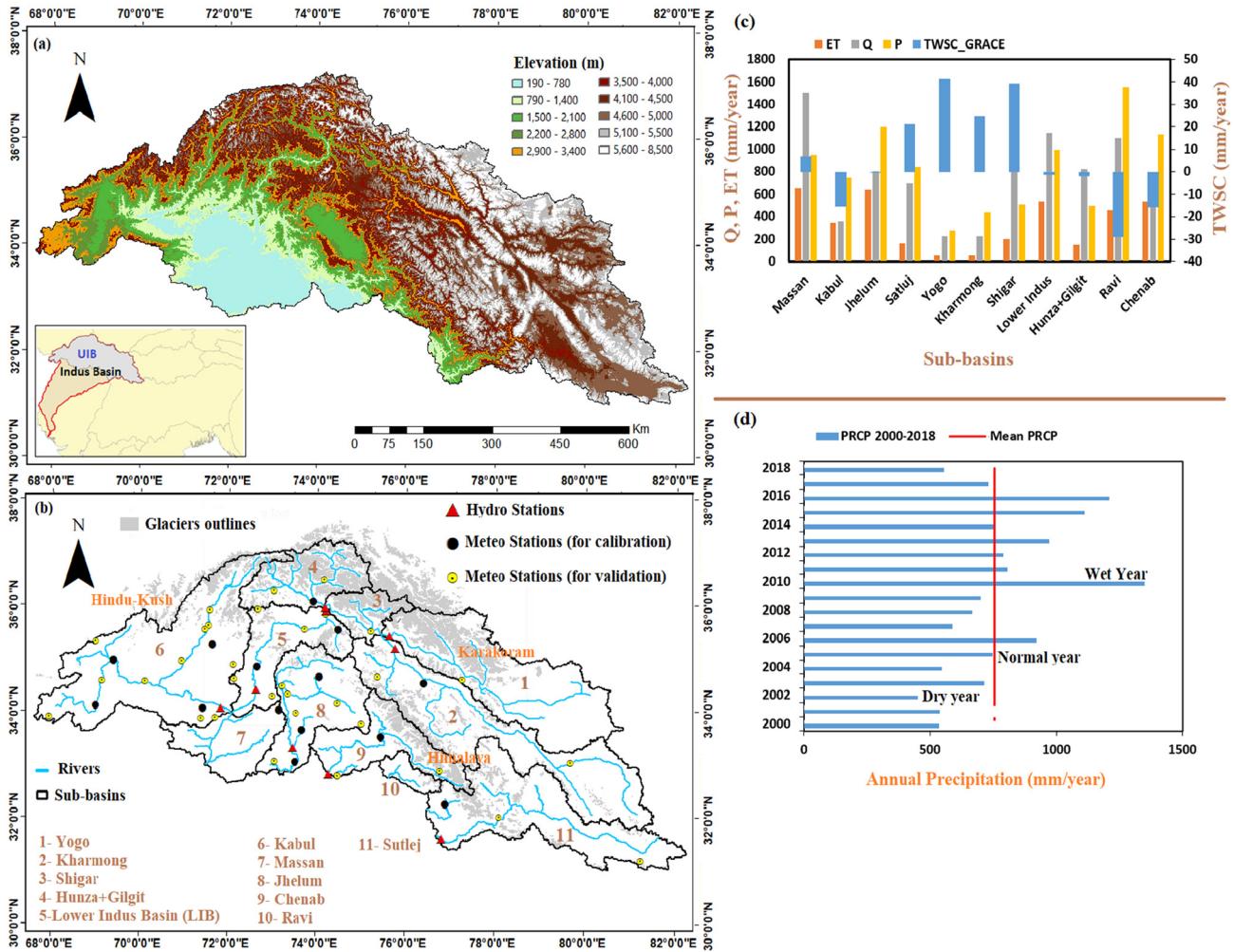


Fig. 1. Map showing features of the study area as (a) topographic condition, (b) mountain ranges, main rivers, sub-basins and locations of hydro-met stations in the high-altitude mountain regions of UIB, (c) average values of water balance components in the 11-sub-basins, and (d) annual precipitation values for different years.

the United States and the Japan Aerospace Exploration Agency (JAXA), was launched in November 1997. In the present study, the daily TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 version 7 at the spatial resolution of $0.25^\circ \times 0.25^\circ$ for the period of 2000–2018 was collected from <http://mirador.gsfc.nasa.gov/>.

2.2.2. Environmental variables

Over the complex topography of the UIB, different environmental variables, such as elevation, slope, normalized difference vegetation index (NDVI), land surface temperature (LST), actual evapotranspiration (AET), wind speed (WDS), and cloud cover, are assumed to provide necessary justification for changing trend in precipitation pattern over the study region. DEM data of SRTM at 90 m resolution was obtained from Jet Propulsion Laboratory <http://www2.jpl.nasa.gov/srtm/> and resampled to 1 km \times 1 km resolution using pixel averaging method. Geolocations (latitude and longitude) and the slope was further derived from DEM in ArcGIS. Monthly NDVI product (MOD13A3) of the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard Terra sensor at a spatial resolution of 1 km \times 1 km was obtained from NASA Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/dataset_discovery/modis). In this study, the NDVI during May-Sep was used as an annual proxy for vegetation growth (Wang et al., 2019). Anomalous pixels of some land cover classes (e.g., snow cover, water bodies, urban areas, wetlands, and farming lands) which mainly influence the vegetation growth were removed from NDVI pixels (Zhou et al., 2017). Land cover data (version MCD12Q1) from MODIS onboard Terra sensor at 500 m \times 500 m spatial resolution was used to remove the anomalous pixels from original NDVI which were then substituted with interpolated values using the IDW kriging approach. LST data of MOD11A2 at 1 km \times 1 km resolution was obtained from NASA Land Processes Distributed Active Archive Center for the period of 2000–2018. This data is composed of 8-day intervals which was finally aggregated to annual average values.

Monthly AET (MOD16) at 1 km \times 1 km resolution was taken from Numerical Terradynamic Simulation Group (<https://www.ntsg.umt.edu/project/modis/mod16.php#data-product>). Monthly reanalyzed data of wind speed (at 10 m height) and cloud cover at a spatial resolution of $0.125^\circ \times 0.125^\circ$ for the period of 2000–2018 was obtained from the European Center for Medium-Range Weather Forecasts (ECMWF) (<http://apps.ecmwf.int/datasets/data/interim-full-modas/>). Monthly wind speed and could cover were further aggregated to an annual scale and spatially resampled to a 1 km \times 1 km resolution using the bilinear interpolation technique (Chao et al., 2018).

2.2.3. Ground observations data from RGS

Recorded precipitation values from 46 RGS in four countries Pakistan (30 stations), Afghanistan (7 stations), India (7 stations), and China (2 stations) were used in present study.

Daily recorded precipitation values from 2000 to 2018 in Pakistan were obtained from Pakistan Metrological Department (PMD) and Pakistan Water and Power Development Authority (WAPDA). Monthly recorded rainfall values for 2008–2018 for the stations located in Afghanistan were provided by the Afghanistan Metrological Department (AMD) as well as NOAA Central Library of the US (https://docs.lib.noaa.gov/rescue/data_rescue_afghanistan.html) and US Geological Survey (<http://edcintl.cr.usgs.gov/downloads/sciweb1/shared/afghan/downloads/documents/>). Moreover, missing values (from 2000 to 2007) were filled with the gridded precipitation data of APHRODITE (Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources). Daily gridded precipitation values from 2000 to 2007 were extracted from APHRODITE version (APHRO_MA_V110) over the station's locations. APHRODITE precipitation dataset has already been utilized as the baseline in the Indus Basin for climate and hydrological studies (Rizwan et al., 2019; Tahir et al., 2011). Daily precipitation data from 2000 to 2018 for the stations located in the Indian Territory was downloaded from India Metrological

Department (www.imd.gov.in), NOAA-NCDC's (<http://www.ncdc.noaa.gov/cdo-web/datasets>) and KNMI Climate Explorer (<https://climexp.knmi.nl>) (Dahri et al., 2016). Daily precipitation data of metrological stations located in China was collected from China Meteorological Data Network (CMA) available at <https://data.cma.cn/en>. Rain gauge stations (RGS). The precipitation record from available RGS was divided into two sample sets: 32 RGS were used to calibrate the precipitation results and the remaining 14 RGS were used for validation.

2.3. Spatial downscaling and calibration process of TRMM data

In this study, an Integrated Downscaling and Calibration (IDAC) framework was developed to generate a high-resolution (1 km \times 1 km) gridded precipitation data. A hybrid downscaling approach using a regression model was proposed to retrieve high-resolution annual TRMM precipitation estimates. The downscaling was performed based on two underlying assumptions: firstly, a spatial relationship exists between precipitation and environmental variables, which can be explained with the established models; and secondly, the models established at low spatial resolution can also be utilized to predict precipitation at finer resolutions with the help of environmental variables (Jing et al., 2016).

The detailed procedure of how the TRMM precipitation is down-scaled is summarized in Fig. 2 and explained as follows:

1. The daily TRMM precipitation data and the environmental variables were temporally up-scaled to the annual timescale where precipitation variations depicted a strong relationship with environmental variables.
2. All environmental variables at the 1 km \times 1 km spatial resolution were resampled to a coarse resolution ($0.25^\circ \times 0.25^\circ$) using pixel averaging to match with the spatial resolution of the Original-TRMM precipitation ($0.25^\circ \times 0.25^\circ$).
3. The relationships between resampled environmental variables and the Original-TRMM precipitation were established using geographical variability test (GVT) to screen out the spatially varying (local variables) and fixed (global variables) environmental variables.
4. Spatially varying and fixed environmental variables at coarse resolutions were introduced in the GWR and MGER models to derive the predicted precipitation at the spatial resolution of $0.25^\circ \times 0.25^\circ$.
5. The difference between the predicted precipitation ($0.25^\circ \times 0.25^\circ$) and the Original-TRMM precipitation ($0.25^\circ \times 0.25^\circ$) was measured to obtain the precipitation residuals which were then interpolated at a 1 km \times 1 km spatial resolution using kriging interpolation technique.
6. Again, the high resolution environmental variables (1 km \times 1 km) were introduced into those models developed in step-4 to obtain the estimated precipitation (1 km \times 1 km) without being corrected with residuals. Finally, precipitation residuals at the 1 km \times 1 km scale obtained in step-5 were added with the estimated precipitation (without residual) to obtain the high resolution downscaled gridded precipitation data at the 1 km \times 1 km scale.
7. Annual downscaled gridded precipitation data was finally calibrated and merged with observations from rain gauges (RGS) using spatial algorithms of geographical ratio analysis (GRA) (Duan and Bastiaanssen, 2013) and geographical difference analysis (GDA) (Cheema and Bastiaanssen, 2012). Eqs. (1) and (2) represent the GRA and the GDA methods applied, respectively.

$$P_{cal_down}(x) = P_{down}(x) \times \sum_{i=1}^n \lambda_i \frac{P_{obs}x_i}{P_{down}x_i} \quad (1)$$

$$P_{cal_down}(x) = P_{down}(x) + \sum_{i=1}^n \lambda_i [(P_{obs}(x_i) - P_{down}(x_i))] \quad (2)$$

where; P_{cal_down} represents the calibrated-downscaled precipitation or finally merged precipitation with ground-based data at the target point

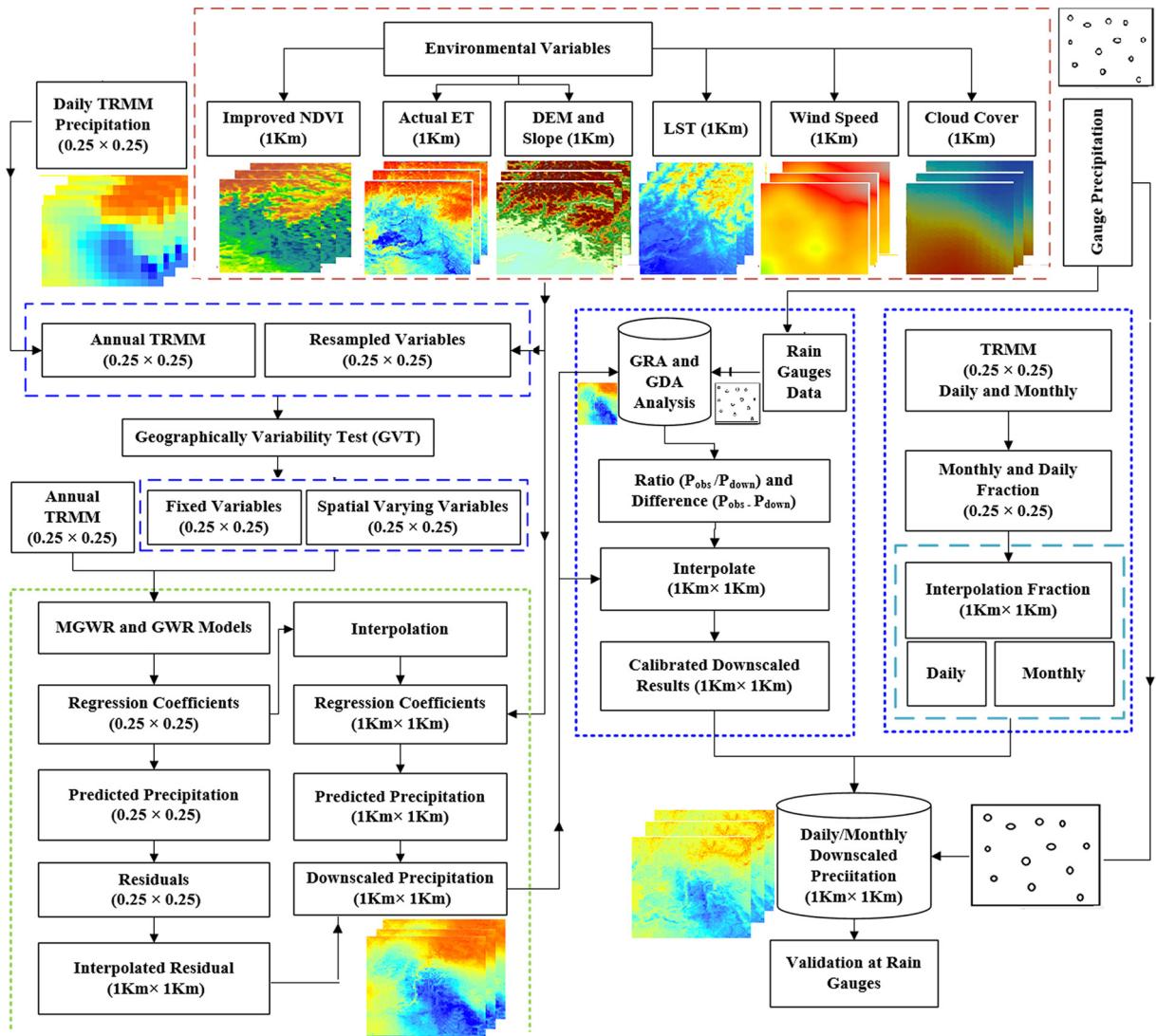


Fig. 2. Comprehensive flow diagram of an Integrated Downscaling and Calibration (IDAC) framework to generate high-resolution ($1 \text{ km} \times 1 \text{ km}$) gridded precipitation data in the present study.

"x". Observed precipitation at point x_i is denoted with $P_{\text{obs}}(x_i)$ while the corresponding downscaled precipitation value is represented by $P_{\text{down}}(x_i)$. λ_i is the weight at each location x_i . The ratio ($P_{\text{obs}} / P_{\text{down}}$) and the residuals ($P_{\text{obs}} - P_{\text{down}}$) between the observed and the downscaled precipitation ($P_{\text{obs}} / P_{\text{down}}$) at target points were computed and interpolated at $1 \text{ km} \times 1 \text{ km}$ spatial resolution using the ordinary kriging method in ArcGIS. Finally, the downscaled precipitation ($1 \text{ km} \times 1 \text{ km}$) data was multiplied and added with the ratio (at a $1 \text{ km} \times 1 \text{ km}$ scale) and the residuals ($1 \text{ km} \times 1 \text{ km}$) respectively to obtain the calibrated-downscaled precipitation or merged precipitation data.

8. In the final step, high resolution annual TRMM precipitation data retrieved in the previous step was temporally downsampled to obtain the high-resolution monthly and daily precipitation estimates using a simple fraction disaggregation method proposed by (Duan and Bastiaanssen, 2013).

$$\begin{aligned} \text{TRMM}_{\text{down}}(\text{monthly and daily})^{1\text{km}} &= \text{Annual TRMM}_{\text{down}}(1\text{km}) \\ &\times \sum_{i=1}^N \lambda_i \left(\frac{\text{TRMM}_j(x_i)^{0.25 \text{ deg}}}{\sum_{j=1}^n \text{TRMM}(x_i)^{0.25 \text{ deg}}} \right) \end{aligned} \quad (3)$$

where the $\text{TRMM}_j(x_i)^{0.25 \text{ deg}}$ represents the Original-TRMM precipitation of the j^{th} day at point x_i ; and the denominator is the annual total precipitation for $n = 365$. λ_i is the weight at each location x_i and N indicates the total number of grid cells ($0.25^{\circ} \times 0.25^{\circ}$). TRMM Precipitation fraction at $0.25^{\circ} \times 0.25^{\circ}$ spatial resolution for each month & day was computed by dividing with annual total precipitation and then interpolated to $1 \text{ km} \times 1 \text{ km}$ spatial resolution by using ordinary kriging. Precipitation fraction at $1 \text{ km} \times 1 \text{ km}$ was finally multiplied with annually downscaled precipitation ($1 \text{ km} \times 1 \text{ km}$) to obtain high resolution monthly and daily precipitation estimates.

2.3.1. GWR and MGWR regression models

The GWR model is an extension of the traditional ordinary least square regression method which describes the relationship between variables at geographical locations instead of a sole regression for the entire study region (Fotheringham et al., 2003; Kumari et al., 2017). The GWR model can be mathematically described as:

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon(u_i, v_i) \quad i=1, 2, 3, \dots, n \quad (4)$$

where Y_i is the i^{th} observation of the dependent variable; x_{ik} is the i^{th} observation of the k^{th} independent variable; the (u_i, v_i) represents the geographical locations of the i^{th} regression point; $\beta_0(u_i, v_i)$ is the

intercept at locations (u_i, v_i) , and $\beta_{ik}(u_i, v_i)$ is the k th regression parameter in the i th regression point, which is a function of geographical location. Furthermore, $\varepsilon(u_i, v_i)$ represents the random error at each location (u_i, v_i) .

Regression parameters were solved by using the following matrix:

$$\beta(u_i, v_i) = (x^T(w(u_i, v_i))x)^{-1} x^T w((u_i, v_i)) y \quad (5)$$

In this equation, $\beta(u_i, v_i)$ is the local regression coefficients to be estimated at the location (u_i, v_i) ; x and y are the vectors of the explanatory and the dependent variables, respectively; and $w(u_i, v_i)$ is the weight matrix. Solving the GWR equations requires primarily estimating the weight matrix which is crucial for evaluating model accuracy (Chao et al., 2018). In this study, two different kernel functions such as BI (Bi-square) and GU (Gaussian) were applied to solve the weight matrix. In the Gauss method (GAU), w_{ij} is computed as a function of d_{ij} according to Fotheringham et al. (1998):

$$w_{ij} = \exp\left[-\frac{d_{ij}^2}{\theta_{i(k)}^2}\right] \quad (6)$$

w_{ij} is the weight of the observation at the location j for estimating the coefficient at the location i ; d_{ij} is the Euclidean distance between i and j and $\theta_{i(k)}$ is an adaptive bandwidth size defined as the k -th nearest neighbor distance. To reduce the computational cost, the above-defined Gauss function can be replaced by a bi-square (BI) function to generate continuous weight values while eliminating those points that are excessively distant (Fotheringham et al., 1998).

$$w_{ij} = \begin{cases} (1-d_{ij}^2/\theta_{i(k)})^2 & d_{ij} < \theta_{i(k)} \\ 0 & d_{ij} > \theta_{i(k)} \end{cases} \quad (7)$$

MGWR model is the extended form of the GWR which handles the mixed coefficients with spatial varying (local variables) and fixed or constant coefficients (global variables) in computation (Nakaya, 2015). In general, MGWR is a generalized GWR method by introducing additional global variables to better represent the spatial variability of precipitation in the regression model (Chao et al., 2018).

$$Y_i = \text{Global variables} \left[\sum_{j=1}^k \alpha_j x_{ij} \right] + \text{Local variables} \left[\sum_{l=k+1}^p \beta_l(u_i, v_i) x_{il+\varepsilon(u_i, v_i)} \right] \quad (8)$$

For $j = 1, 2, \dots, k$, $l = 1, 2, \dots, p$, α_j represents fixed coefficient of global constant parameters while $\beta_l(u_i, v_i)$ is a geographically varying coefficient of the local parameters, and $\varepsilon(u_i, v_i)$ is the residuals at each location. In this study, the geographical variability test (GVT) was performed to identify the spatially varied (local) and fixed (global) environmental variables over the study area. Diff of Criterion (DIF) was used as an indicator to judge the fixed and varied terms in the model. The positive value of DIF indicates no spatial variability in specifying that term as a local variable in the GWR model. We repeated the GVT for each variable in the GWR model to finally confirm the fixed and varied variables, and the final MGWR model thus can be structured.

2.4. Statistical evaluation

In this study, the precipitation data was validated in the following two steps;

1. Annual precipitation estimated with two different models (GWR and MGWR) and two kernel filters (Bi-Square and Gaussian) were compared with the Original-TRMM precipitation at $0.25^\circ \times 0.25^\circ$ spatial resolution using the Akaike information criterion (AICc) and cross

validation (CV). A model with a low AICc and CV values was considered to be more effective at predicting precipitation.

2. Accuracy of the model-based downscaled and calibrated-downscaled results were compared with the observations at RGS. Only those gauges were selected in the validation step which were not used previously during the calibration step. Different statistical indicator coefficients of determination (R^2), Root Mean Square Error (RMSE), Bias (B), and mean values were used to evaluate the results as described below.

$$R^2 = \frac{\left[\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P}) \right]^2}{\sum_{i=1}^n (O_i - \bar{O})^2 (P_i - \bar{P})^2} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}} \quad (10)$$

$$Bias = \frac{\sum_{i=1}^n P_i}{\sum_{i=1}^n O_i} \quad (11)$$

where P_i is precipitation estimate (downscaled and calibrated-downscaled), and O_i is the observed precipitation. O and \bar{P} represent the mean values for observed and high-resolution precipitation estimates, and n is the total number of values in the corresponding dataset.

3. Results

3.1. Precipitation response to environmental variables in UIB

To confirm the spatial relationships of the TRMM with its influencing environmental variables in different climatic regions, we selected two climate zones: (1) sub-humid region with low elevation (300–1000 m), and it's mainly comprised of a dense vegetation and (2) semi-arid region with elevation ranging 400–5500 m and its mainly comprised of snow cover, glaciers, scarce vegetation, mountain peaks, and lakes. Fig. S1 illustrates the changes in the TRMM and its influencing variables along transects of 700 km length drawn in (a) sub-humid region, (b) semi-arid region, while (c) shows the correlation coefficient (cc) between TRMM precipitation and its influencing variables over the entire UIB at $0.25^\circ \times 0.25^\circ$ scale. It can be seen that the precipitation profile shows a similar pattern as elevation in the sub-humid region, e.g., precipitation is continuously increasing from 150 mm to 1400 mm corresponding to the elevation rise from 600 m–1000 m, while in the semiarid region the relationships show the opposite trend. Previous studies have also shown that the topographical changes in different climate zones are presumed to provide the necessary justification for changing trend in precipitation (Bookhagen and Burbank, 2006; Mukherjee et al., 2015). In the present study, the changes in precipitation pattern seems to be consistent with NDVI profile in both climate zones but the trend was much clearer in the semiarid region with normal precipitation and sparse vegetation. There is a strong correlation ($cc = 0.73$) between precipitation and the NDVI over the whole study area. The spatial pattern of precipitation is strongly associated with land surface temperature (LST), but the relationship turned to be different in both sub-humid and semiarid regions. In case of evapotranspiration, the precipitation and ET_a variations are consistent with each other. The correlation coefficient between precipitation and ET_a was found to be positive with an average value of 0.70 over the entire study region. In both climate regions, precipitation decreased with an increase in cloud cover. However, no clear trend was found between wind speed and precipitation in the UIB. The other environmental

variables considered in current study also revealed different responses to precipitation in both climate regions. Table S1 summarizes the correlation coefficients between precipitation and environmental variables at $0.25^\circ \times 0.25^\circ$ spatial resolution. Results indicated that NDVI, elevation, and AET are highly correlated with precipitation in the UIB for all climate years. Overall, there was a close relationship between all environmental variables and precipitation, indicating that these variables could be useful in downscaling TRMM precipitation at high spatial resolution (see in supplementary material).

Moreover, we also performed the geographical variability test to evaluate the responses of different environmental variables in terms of spatially varying (local) and fixed (global) variables over the entire study region. According to the diff-Criterion (DIFF) value, only the variables with a higher positive value are considered as the fixed variables. Table 1 describes the summery of statistics of the geographical variability test (GVT) of environmental variables based on DIFF of Criterion. It can be seen that only slope and LST have positive DIFF values (i.e., 47.92 and 28.26). Positive DIFF values for slope and LST suggests that there is no spatial variability in terms of specifying these variables as local terms in the model. Therefore, slope and LST were filtered out to be the fixed variables (global terms) while other variables were introduced as varying parameters (local terms) into the model.

3.2. The comparison between GWR and MGWR predictions

The establishment of regression relationships between the TRMM precipitation and environmental variables is a vital step for downscaling which in turn affects the performance of model prediction. Therefore, the predictive performances of four developed regression models (GWR-BI, GWR-GU, MGWR-BI, and MGWR-GU) were evaluated by comparing the predicted annual precipitation with Original-TRMM precipitation at $0.25^\circ \times 0.25^\circ$ spatial resolution. BI (Bi-square) and GU (Gaussian) indicate the two different kernel weighting functions used in this study. Table 2 summarized the comparison of the two models (GWR and MGWR) based on the two different kernel weighting functions. The selection of the best model was done based on the AICc (Akaike Information Criterion-corrected), R^2 (correlation coefficient), and CV (cross-validation). The model with the lowest AICc and CV values and a higher R^2 value was considered to be the best model. The results indicated that the MGWR model has smaller AICc values as compared to traditional GWR, which indicated that introducing new global variables into the traditional GWR model most probably have reduced the AICc value and improved the model prediction accuracy. Regarding the two kernel weighting functions, the largest AICc and smallest R^2 corresponded to GU while the opposite was found for BI, which indicated that the BI kernel performs better than the GU (Chao et al., 2018). Therefore, models with BI kernel (GWR and MGWR) were selected for further evaluation and downscaling process of precipitation.

Both GWR and MGWR models were applied with BI kernel to predict the TRMM precipitation at $0.25^\circ \times 0.25^\circ$ spatial resolution using different environmental variables. Fig. S2 illustrates the comparison between GWR

Table 2

Performance comparisons of GWR and MGWR model predictions with two different kernels (GU and BI) applied.

| Models | Sub-Models | Auxiliary Variables | R-seq | AIC _C | CV |
|--------|------------|---|-------|------------------|---------|
| GWR | GWR_GU | Local ($x_1, x_2, x_3, x_4, x_5, x_6, x_7$) | 0.90 | 8504.64 | 7549.42 |
| | GWR_BI | Local ($x_1, x_2, x_3, x_4, x_5, x_6, x_7$) | 0.93 | 8408.46 | 6298.42 |
| | MGWR_GU | Local (x_1, x_4, x_5, x_6, x_7) | | | |
| | | Global (x_2, x_3) | 0.95 | 8349.52 | 7332.68 |
| MGWR | MGWR_BI | Local (x_1, x_4, x_5, x_6, x_7) | | | |
| | | Global (x_2, x_3) | 0.97 | 8037.18 | 6104.32 |

x_1 = elevation; x_2 = slope; x_3 = LST; x_4 = AET; x_5 = NDVI; x_6 = wind speed; x_7 = cloud cover.

and MGWR models in terms of accuracy against the Original-TRMM precipitation according to two statistical indicators (a) distribution of R^2 values over 714 grids at a spatial resolution of $0.25^\circ \times 0.25^\circ$ and (b) spatial distribution of the residuals of GWR and MGWR during the period of 2000–2018. Both models showed promising results in predicting precipitation with higher R^2 values (0.25–0.90) over the entire study region (Fig. S2a & b). However, it was observed that MGWR model performed better over most of the grids with R^2 values between 0.6 and 0.9 with only few grids having unsatisfactory model performance ($R^2 < 0.6$). R^2 values in the range of 0.25–0.60 were slightly higher in GWR predictions as compared to the MGWR model, which suggested that MGWR model performed better than the traditional GWR model. Fig. S2c & d illustrate the spatial distribution of residuals in GWR and MGWR models as compared with the Original-TRMM. Obviously, both models overestimated the precipitation values against the Original-TRMM precipitation in the lower sub-humid region, while under-estimated it in the upper semi-arid region. Nevertheless, MGWR performed better than GWR model given it lower residuals (see in supplementary materials).

GWR and MGWR with BI kernel were further tested for different climate years; dry year (2002), normal year (2005), wet year (2010), and an average of entire period (2000–2018). Table 3 summarizes the values of statistical indicators for the evaluation of GWR and MGWR models. Different statistical parameters (R^2 , RMSE, and bias) were estimated to evaluate the predictive performance of GWR and MGWR models. It was found that both models produced satisfactory results with higher R^2 values and RMSE lower values in all climate years. The MGWR model outperformed the GRW model with lower bias and higher R^2 values in the entire period, suggesting that introducing fixed (constants) variables in the global part of MGWR model can reduce the bias and increase accuracy of the model prediction (Zeng et al., 2016).

3.3. Spatial downscaling of annual TRMM precipitation and merging analysis

GWR and MGWR models with BI-square filter were finally applied to generate high-resolution annual precipitation estimates for different climate years 2002, 2005, 2010, and average of entire period 2000–2018. Firstly, the relationship between environmental variables (elevation,

Table 1

Geographical variability test (GVT) of environmental variables based on DIFF of Criterion.

| Variables | Estimate | Min | Max | Range | DIFF of Criterion |
|-------------|----------|---------|---------|---------|-------------------|
| Intercept | 856.46 | 3660.15 | 4674.40 | 8334.64 | -4239.98 |
| Elevation | 0.47 | -0.17 | 0.67 | 0.50 | -1107.22 |
| Slope | -1.52 | -12.96 | 4.90 | 17.92 | 47.92 |
| LST | -5.36 | -59.86 | 28.50 | 88.46 | 28.26 |
| AET | 0.19 | -1.00 | 2.20 | 3.22 | -59.08 |
| NDVI | 43.33 | 1574.52 | 1306.20 | 2880.81 | -9.17 |
| Wind speed | 70.57 | 1208.97 | 1204.50 | 2413.51 | -156.02 |
| Cloud cover | -576.81 | 8107.79 | 5221.80 | 3329.66 | -1909.79 |

Positive value of diff-Criterion (DIFF) suggests that there is no spatial variability in terms of specifying environmental variables as local parameter.

Table 3

Comparisons of GWR and MGWR models with annual precipitation from TRMM 3B43.

| Models | Evaluation parameters | Dry year (2002) | Normal year (2005) | Wet year (2010) | Average (2000–2018) |
|--------|-----------------------|-----------------|--------------------|-----------------|---------------------|
| GWR | R ² | 0.94 | 0.95 | 0.90 | 0.95 |
| | Bias | 0.08 | 0.000 | 0.15 | 0.094 |
| | RMSE | 73.41 | 72.56 | 89.10 | 60.76 |
| | R ² | 0.95 | 0.95 | 0.92 | 0.96 |
| MGWR | Bias | -0.03 | -0.001 | 0.012 | 0.014 |
| | RMSE | 69.78 | 70.55 | 77.34 | 54.01 |

slope, ET, NDVI, LST, wind speed, and cloud cover) and Original-TRMM precipitation was evaluated using both models at coarse spatial resolution ($0.25^\circ \times 0.25^\circ$). Local regression coefficients ($0.25^\circ \times 0.25^\circ$) obtained from GWR and MGWR models were interpolated at $1\text{ km} \times 1\text{ km}$ spatial resolution using the ordinary kriging method. An example of the spatial distribution of local coefficients and R^2 generated by the MGWR model over the UIB for the period of 2000–2018 could be seen in supplementary materials (Fig. S3 see in supplementary material). Local regression coefficients at $1\text{ km} \times 1\text{ km}$ spatial resolution were applied to high resolution environmental variables ($1\text{ km} \times 1\text{ km}$) to predict and downscale the Original-TRMM precipitation (Fig. 3). Fig. 3 a, b show the comparison between model-based predicted precipitation at coarse resolution ($0.25^\circ \times 0.25^\circ$) with the Original-TRMM precipitation during the period of 2000–2018. In general, precipitation estimated by the MGWR model exhibited a similar pattern as to the Original-TRMM precipitation, however over- and underestimations still exist at some locations. Fig. 3c indicates the spatial distribution of residuals (difference between model-based predicted precipitation and Original-TRMM) at coarse resolution ($0.25^\circ \times 0.25^\circ$) which are not explained by the model. Positive residuals values imply precipitation underestimation while negative values indicate model overestimation. The maximum under- and over-estimation in model-based estimated precipitation compared to the Original-TRMM were 84 mm/year and 92 mm/year, respectively. Precipitation estimation (without residuals correction) at high-resolution ($1\text{ km} \times 1\text{ km}$) was carried out by applying high resolution environmental variables to local coefficients obtained from the model in the previous step (Fig. 3e). In the upper reaches of the study area, precipitation was underestimated which could not be explained by the model due to the presence of residuals. Therefore, residuals obtained in the previous step were interpolated in a $1\text{ km} \times 1\text{ km}$ resolution and then added to the estimated precipitation to retrieve high resolution ($1\text{ km} \times 1\text{ km}$) downscale precipitation as shown in Fig. 3f. It was seen that the downscale precipitation data not only improved the spatial information but also retained the spatial pattern when compared with the Original-TRMM data. Precipitation profile along the transect line (~750 km) drawn over the Original-TRMM and downscale precipitation grids shows a very strong similarity in both datasets. However, spatial variations depicted by the downscale precipitation data are obviously showing much clearer and variable trend along the line of transect which could be seen as uniform in the Original-TRMM precipitation (Fig. 3g). A similar methodology was applied to downscale the TRMM precipitation data for each year.

Fig. 4 shows the comparison between three precipitation datasets (1) Original-TRMM, (2) downscale precipitation from the GWR model, and (3) downscale precipitation from the MGWR model for three different climate years; the dry period (2002), normal period (2005), wet period (2010) and average of entire period (2000–2018) over the UIB. Obviously, both models performed better than the Original-TRMM, and explained more of the spatial variability in precipitation pattern than the Original-TRMM. However, the spatial patterns of downscale precipitation explained by the MGWR model was more consistent with that of the Original-TRMM. For instance, consider the black circles drawn over some regions in Fig. 4. It can be seen that the downscale precipitation estimated by GWR model does not correspond to that of the Original-TRMM while MGWR-based downscale precipitation exhibits a very strong similarity to the Original-TRMM.

Fig. S4 (see in supplementary material) shows the statistical evaluation of different gridded precipitation datasets (1) (Original-TRMM), (2) estimated precipitation (before residual correction), and (3) downscale precipitation (after residual correction) obtained from GWR and MGWR models. Coefficients of determination (R^2), root mean square error (RMSE), mean, and bias values were estimated for each gridded precipitation dataset to evaluate their accuracy against the observed precipitation from 14 RGS (Fig. S4). On average, the Original-TRMM data under- and over-estimated the observed precipitation with the bias values of -0.06% and 0.08%, respectively, with an R^2 of 0.69 when compared with recorded values. The results from GWR and MGWR indicated that downscale precipitation data (after residual correction) was more realistic than the estimated precipitation (without residual correction) with lower bias and RMSE. The comparison between the downscale precipitation obtained from the GWR and MGWR models indicated higher $R^2 = 0.75\text{--}0.83$ and lower RMSE (bias) ~127.08–204.50 mm (-0.05–0.01) for the MGWR model pointing to a better accuracy than the GWR model ($R^2 = 0.72\text{--}0.73$; RMSE = 164.70–228.79 mm, bias = -0.03–0.03). From Fig. S4, it was observed that high-resolution precipitation after residual correction generates more accurate results, therefore residual correction can be taken as an important step to increase the reliability of precipitation estimation in downscaling (see more detail in supplementary material). The downscale TRMM precipitation data is still overestimated, and the integration of the observations from RGS for available locations is necessary to generate more reliable precipitation data over the study area. Since the downscaling techniques used in present study predict the precipitation based on the Original-TRMM data, therefore, over and under-estimations existing in TRMM data will eventually be introduced into the final downscale precipitation results (Immerzeel et al., 2009). A combination of satellite data with ground observation data can improve the accuracy (Cheema and Bastiaanssen, 2012; Li and Shao, 2010). The downscale precipitation obtained from the MGWR model were further merged with the observation at 32 RGS while the remaining 14 RGS were used during the validation step in later stage.

Fig. 5 illustrates the spatial comparison of high-resolutions calibrated-downscale gridded precipitation derived from the GRA and GDA models for three climate years (2002, 2005, and 2010), and an average of the entire period 2000–2018. An interesting result was observed for the dry years (2002) where downscale and calibrated-downscale precipitations showed a similar spatial pattern, with a slight deviation for the Kabul River Basin after merging with the data from rain gauges. It indicates that the downscaling model used in this study was capable of producing the actual ground precipitation with lower uncertainty for the dry and normal years. However for the wet year (2010), gridded precipitation derived from both calibration methods (GRA and GDA) exhibited obvious spatial differences after merging with actual observations especially in upper Indus watersheds such as Hunza, Gilgit and Shigar.

Table 4 indicated the validation results of the calibrated-downscale precipitation after merging with gauge observations using GDA and GRA methods. It was observed that the precipitation estimated with the GDA and GRA models is better than that of the Original-TRMM and downscale precipitation data. However, the GDA method appeared to be slightly less precise as compared to the GRA method in terms of all statistical parameters. Therefore, calibrated-downscale precipitation from GRA was used to retrieve the monthly and daily downscale precipitation. It was seen that, the downscale precipitation data after being calibrated with GDA and GRA models had an improved accuracy with regards to R^2 (ranging 0.81–0.83 using GDA and 0.81–0.88 using GRA). Furthermore, the accuracy of the downscale precipitation merged with gauge observations was significantly better in a dry year (2002) than a wet year (2010). For the dry and wet years, R^2 improved from 0.78 to 0.84, and 0.75 to 0.80, respectively. The RMSE was reduced for all climatic years: the dry year (186.07–124.79 mm), normal

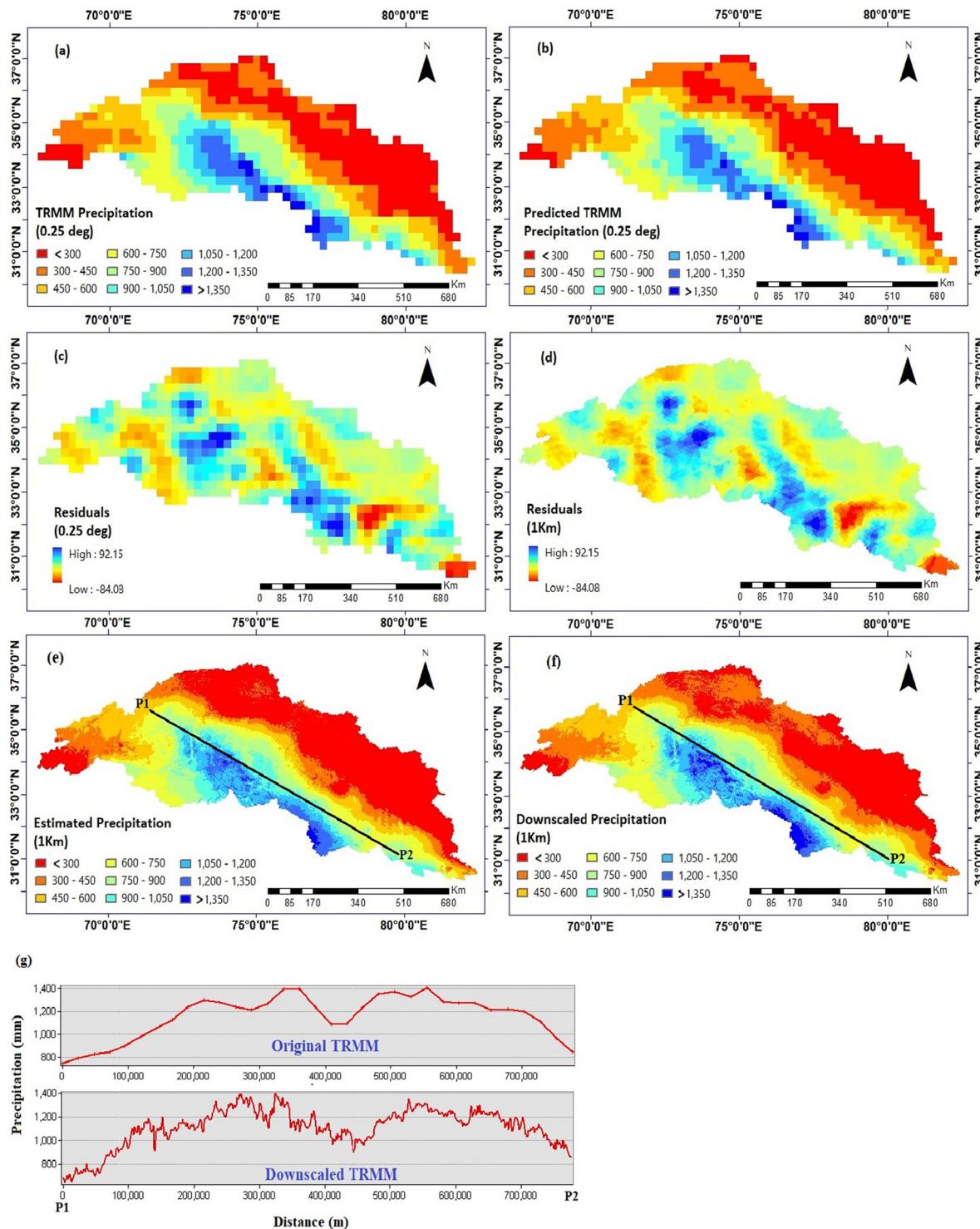


Fig. 3. Showing output maps of MGWR model during the average of entire period (2000–2018) (a) Original-TRMM precipitation at the coarse resolution ($0.25^\circ \times 0.25^\circ$), (b) predicted precipitation by the MGWR model at the coarse resolution ($0.25^\circ \times 0.25^\circ$), (c) residuals at ($0.25^\circ \times 0.25^\circ$), (d) residuals interpolated at the $1\text{ km} \times 1\text{ km}$ resolution, (e) estimated precipitation at the $1\text{ km} \times 1\text{ km}$ resolution without residual correction, (f) final downscaled TRMM precipitation at the $1\text{ km} \times 1\text{ km}$ resolution after residual correction, and (g) precipitation profile along the line of transect (~750 km distance) drawn over the Original-TRMM and downscaled precipitation grids.

year (127.08–81.69 mm), wet year (204.50–134.77 mm), and average of entire period 2000–2018 (127.47–88.94 mm).

3.4. Temporally downscaling of high-resolution annual TRMM precipitation

The annual TRMM precipitation, after being spatial downscaled and merged, was temporally downscaled to obtain the high-resolution monthly and daily precipitation data. For instance, we have took an

example to show the spatial distribution of monthly precipitation derived from the original TRMM, downscaled, and calibrated-downscaled data for January, April, July, and October as shown in Fig. 6. The monthly precipitation from the original TRMM exhibited mosaic-like precipitation distribution due to the coarse resolution. On the other hand, precipitation derived from model-based downscaling procedure not only preserved the spatial pattern in each month but also had an improved spatial information. Monthly precipitation data derived from the calibrated-downscaled

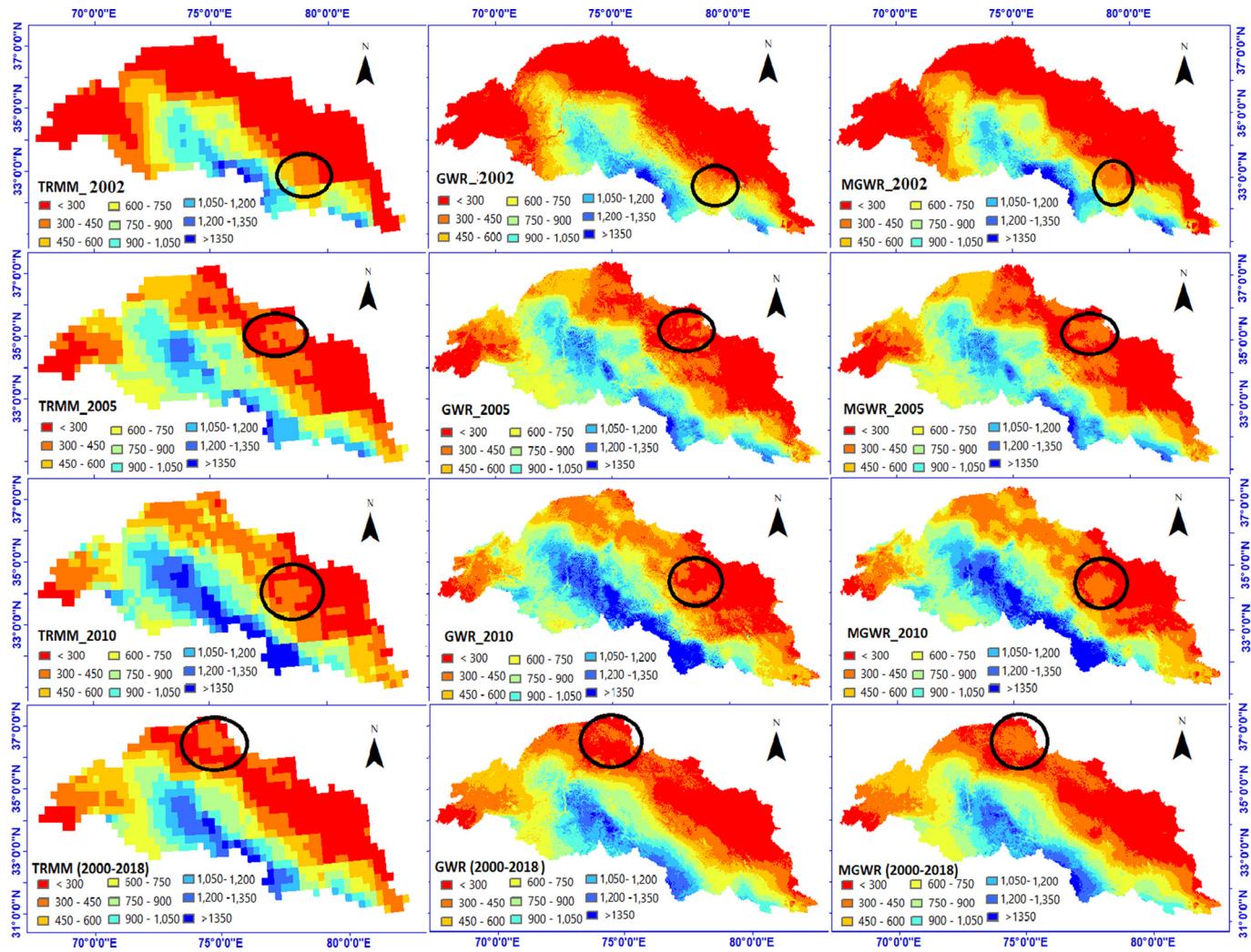


Fig. 4. The comparison between the spatial precipitation patterns obtained from different datasets for the dry period (2002), normal period (2005), wet period (2010) and average of entire period (2000–2018). First, second and third columns represent the precipitation obtained from the Original-TRMM, down-scaled data by the GWR model and down-scaled data by the MGWR model respectively.

procedure were corrected, resulting in a meaningful difference between both down-scaled and calibrated-downscaled precipitation estimates especially in high-altitude regions (including Yugo, Hunza, and Gilgit river basins). The significant difference between both down-scaled and calibrated-downscaled precipitation estimates is associated with the adjusted amount of precipitation after merging with the actual precipitation values. Calibration procedure adjusted the monthly precipitation amount in the down-scaled-TRMM results and eventually reduced the over- and underestimations. For example, the adjusted precipitation showed a significant increase in the Shigar river basin, while a remarkable decrease in the Satluj river basin in the downstream region (Fig. 6).

We examined the capabilities of three types of monthly precipitation datasets (Original-TRMM, down-scaled-TRMM and calibrated-downscaled TRMM) to capture the variations in monthly precipitation at the gauge locations. Monthly observed precipitation data from 14 RGS was compared with precipitation from Original-TRMM, down-scaled-TRMM, and calibrated-downscaled TRMM for three climatic years 2002, 2005 and 2010 (Fig. 7a). Among three precipitation datasets, monthly precipitation time series obtained from down-scaled, and calibrated-downscaled datasets showed quite similar variations as observed precipitation over all gauge locations, however, the Original-TRMM precipitation under-and over-estimated the observed precipitation in some months. Additionally, we calculated four statistical metrics (R^2 , RMSE, bias and mean) for the Original-

TRMM, down-scaled-TRMM, and calibrated-downscaled TRMM compared to observed precipitation averaged over 14 rain gauges stations (see Fig. 7b). Results indicated that the RMSE values of the down-scaled and calibrated-downscaled datasets were close to those of the Original-TRMM in the dry season (from October to April) when the monthly precipitation is slightly lower than ~50 mm/month. However, in the wet season (from May to September) when the monthly precipitation amount is larger than 50 mm/month, the RMSE values of the down-scaled and calibrated-downscaled results were lower than those of the Original-TRMM data. The lower RMSE values in the wet season suggested that the monthly precipitation retrieved from annual down-scaled and calibrated-downscaled data had substantially reduced the seasonal bias in TRMM data (Fig. 7b).

As for R^2 , the down-scaled and calibrated-downscaled datasets outperformed the Original-TRMM in each month; difference is especially more obvious from May–Sep in all climatic years (Fig. 7b). Likewise, it was found that these datasets have unstable performance over time, particularly with lower R^2 for the dry season (as $R^2 = 0.30\text{--}0.50$), and higher $R^2 (> 0.89)$ for the wet season. It was observed that the Original-TRMM data over-and underestimated precipitation against the observations, yet bias in the down-scaled and calibrated-downscaled results are obviously lower. Accordingly, the monthly precipitation from down-scaled and calibrated-downscaled datasets was largely consistent with the actual data. It was also observed that bias

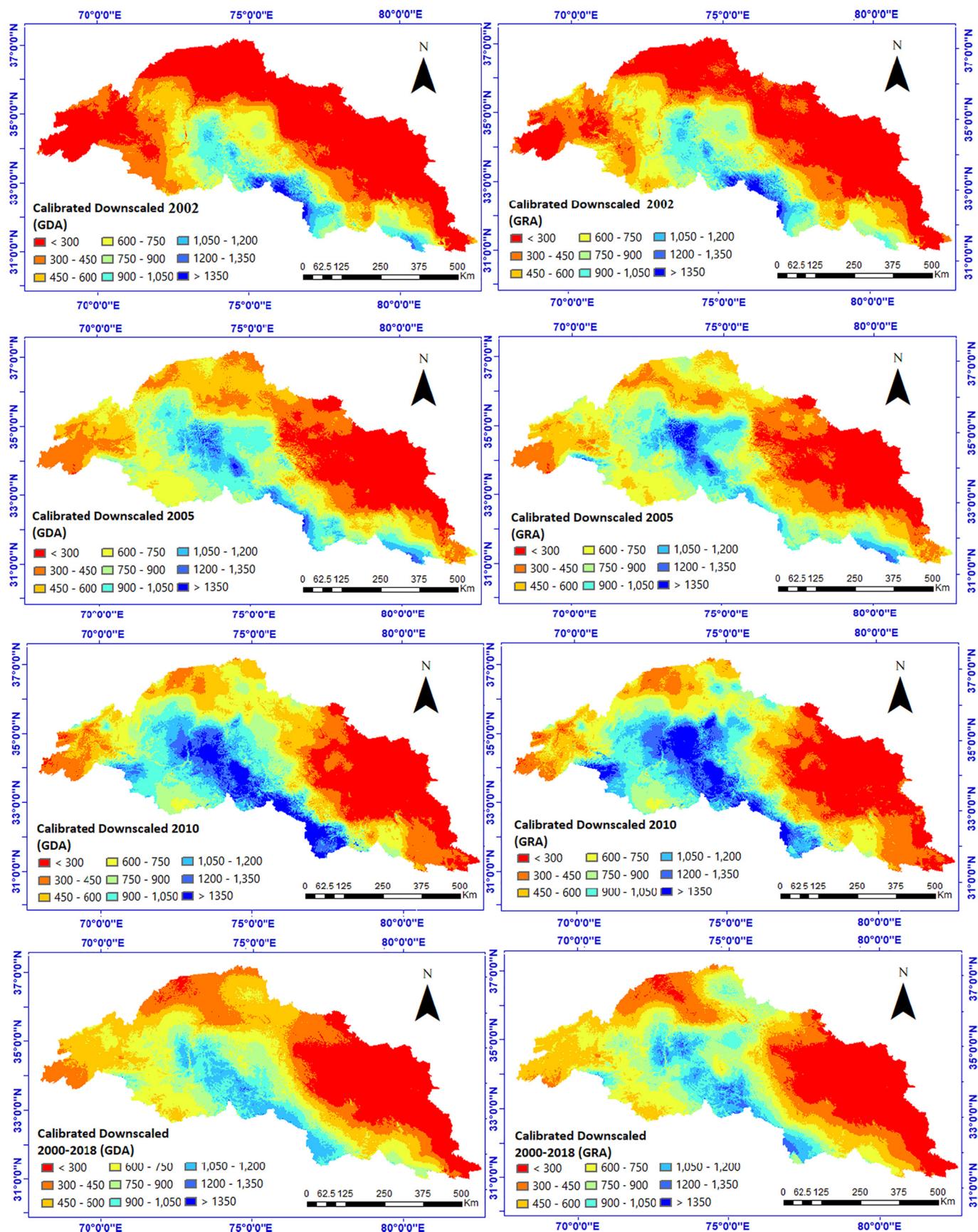


Fig. 5. Spatial comparison between high-resolution calibrated-downscaled precipitation retrieved from GDA and GRA calibration techniques for three reference years (2002, 2005 and 2010) and the average of the entire period 2000–2018.

Table 4

Coefficients of determination (R^2), root mean square error (RMSE), mean, and bias for GDA and GRA calibrated precipitation, compared with precipitation from RGS.

| Datasets | Indicators | Dry Year (2002) | Normal Year (2005) | Wet Year (2010) | Avg. (2000–2018) |
|----------------|-----------------|--------------------|-----------------------|--------------------|---------------------|
| Calibrated GDA | Mean (Observed) | 577.98 | 672.37 | 887.88 | 694.24 |
| | Mean | 528.91 | 692.63 | 895.65 | 674.03 |
| | Bias | 0.03 | -0.02 | 0.05 | 0.00 |
| | RMSE | 147.88 | 105.36 | 159.18 | 103.12 |
| Calibrated GRA | R^2 | 0.81 | 0.84 | 0.79 | 0.83 |
| | Mean | 584.13 | 705.75 | 899.34 | 683.86 |
| | Bias | -0.02 | 0.00 | 0.01 | 0.00 |
| | RMSE | 124.79 | 81.69 | 134.77 | 88.94 |
| | R^2 | 0.84 | 0.88 | 0.80 | 0.88 |

in downscaled and calibrated-downscaled precipitation was slightly higher in the dry season (from Oct-Apr) as compared to the wet season (May-Sep) (also see supplementary Fig. S5). The larger bias corresponds to higher RMSE and lower R^2 values. Moreover, the current findings are also quite similar to those reported by (Duan and Bastiaanssen, 2013). Monthly precipitation retrieved from downscaled and calibrated-downscaled data showed an improvement in prediction accuracy given the higher R^2 and lower RMSE and bias. However, the improvement was unstable over time, mainly for certain months when no changes were observed. The current findings revealed that the integrated spatial downscaling approach along with the calibration procedure could reasonably increase the ability of satellite-based precipitation to capture the temporal variations on the monthly scale.

In the final stage, monthly downscaled and calibrated-downscaled precipitation datasets were further temporally downscaled (using disaggregation method) to retrieve the high-resolution daily precipitation data over the UIB. For instance, we have taken the days 37, 175 and 272

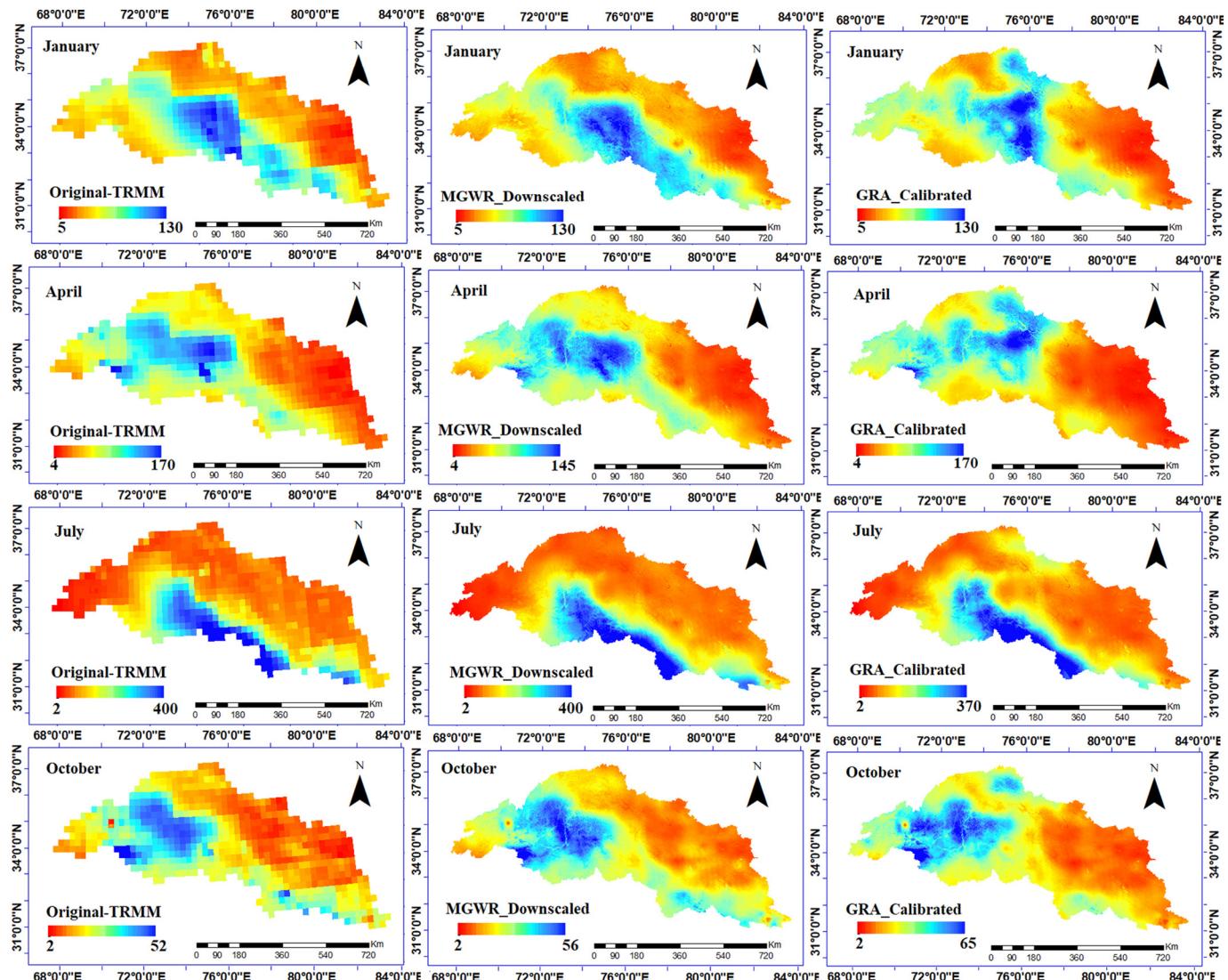
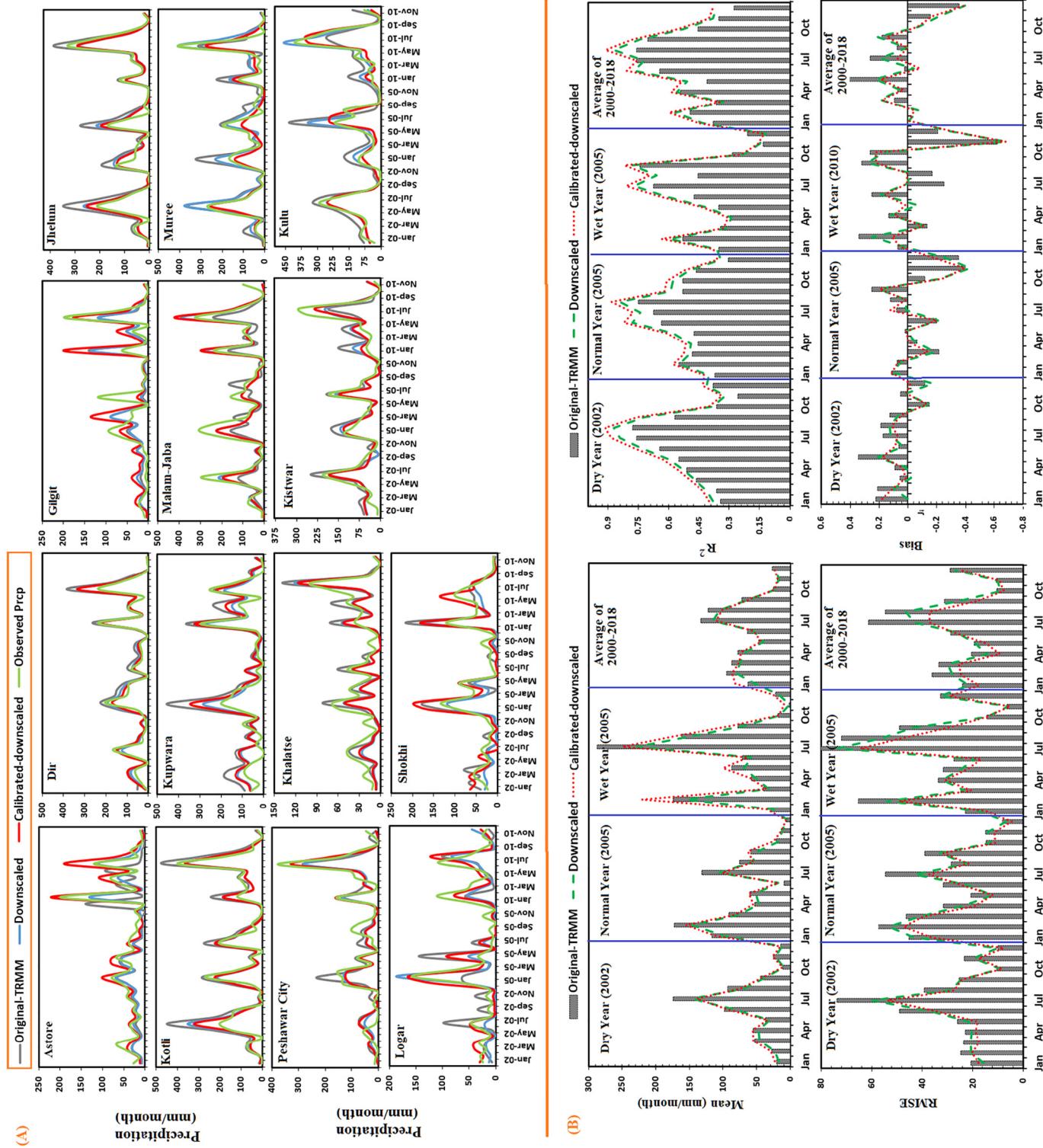


Fig. 6. Spatial representation of monthly precipitation derived from (a) Original-TRMM, (b) MGWR based downscaled, and (c) calibrated-downscaled precipitation data for average of entire period (2000–2018). Columns 1, 2, and 3 represent the Original TRMM, downscaled, and calibrated-downscaled precipitation datasets respectively for different months.

Fig. 7. Time series comparison between mean monthly precipitation from three gridded precipitation estimates (Original-TRMM, downscaled and calibrated-downscaled) with observed precipitation data (a) over 14 rain gauge stations, and (b) monthly statistical evaluation of mean, R^2 , RMSE and bias for Original-TRMM, downscaled, and calibrated-downscaled data compared with the observed precipitation averaged over 14 rain gauge stations.



for the year 2010, to show the spatial distribution of precipitation from Original-TRMM, downscaled, and calibrated-downscaled datasets is shown in Fig. 8a. Accordingly, the daily precipitation from downscaled and calibrated-downscaled datasets revealed more spatial details as compared to the Original-TRMM. Daily precipitation derived from the

calibrated-downscaled dataset was noticeably improved, and exhibited a significant difference in spatial pattern from the downscaled dataset in some regions. However, in most of the study region, precipitation from both downscaled and calibrated-downscaled datasets showed a similar pattern with a slight difference. These finding revealed that the

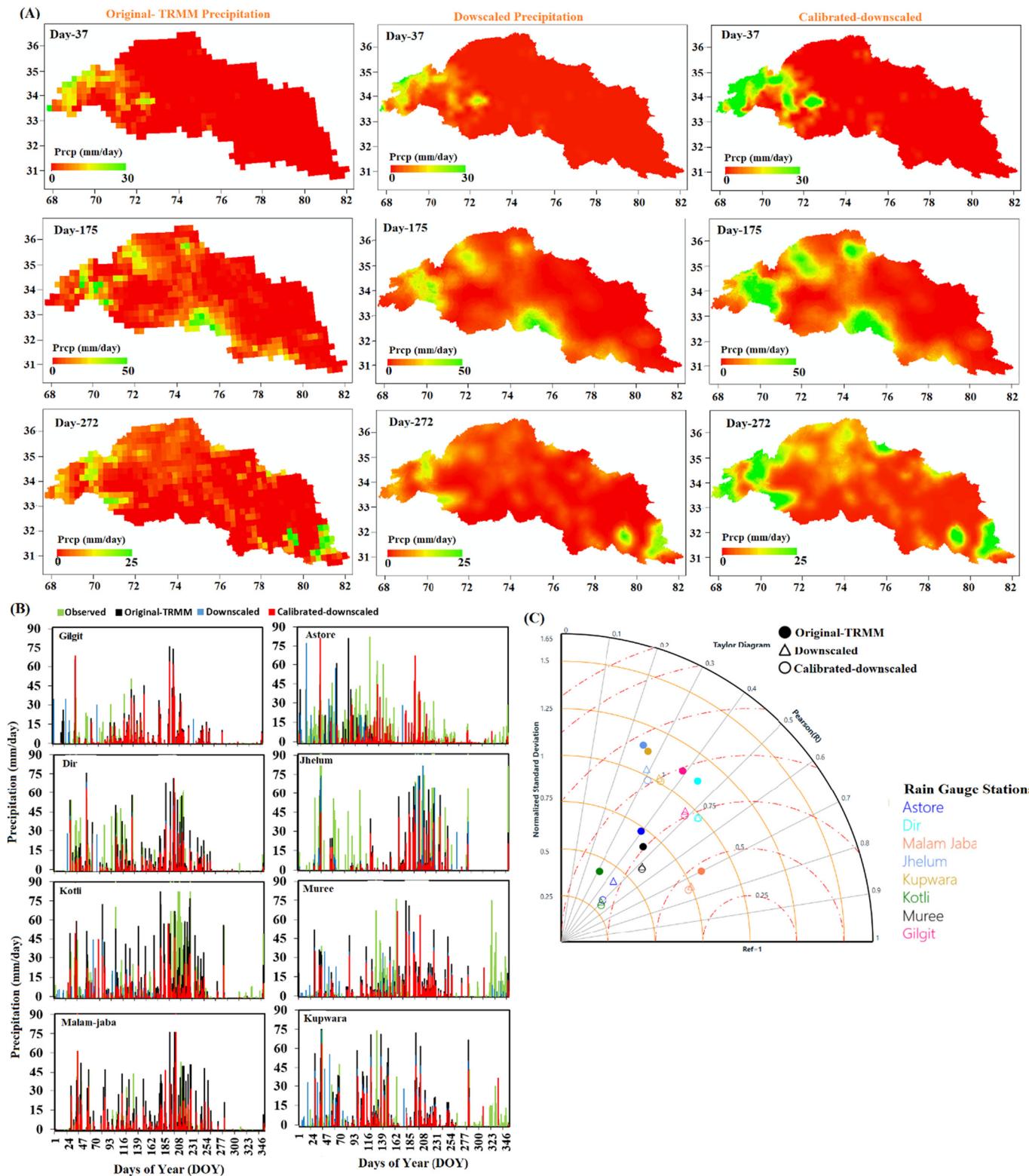


Fig. 8. (a) Spatial representation of daily precipitation derived from three gridded precipitation datasets Original-TRMM, MGWR-based downscaled, and calibrated-downscaled on days, 37, 175 and 272 of year 2010, (b) Time series comparison for performance evaluation of gridded precipitation datasets with reference to observed data over 8 RGS, and (c) Taylor diagram comparing three precipitation dataset with observed precipitation.

temporal decomposition of monthly downscaled and calibrated-downscaled precipitation to retrieve the daily precipitation does not largely influence the final results.

This might be due to three main reasons: first, only 32 sparsely-distributed stations were used in the merging process; density of which is not enough according to extent of study region. Second, most of the gauges are located in downstream humid region where precipitation is higher, while upstream glaciers and the semi-arid region has only few stations where satellite observation are probably highly subjected to uncertainty and errors. Third, improvement in precipitation estimation at daily timescale is offset to some extend when decomposing annual time series to monthly and monthly to daily time scale. Hence, directly disaggregating daily precipitation from annual downscaled and calibrated-downscaled dataset for wet days [e.g. Day-175 and Day-272 during monsoon season when precipitation is higher] remarkably improved the spatial pattern of the daily precipitation which is not clear for the dry days [e.g. Day-37] (Fig. 8a).

Additionally, we have also investigated the performance of the daily precipitation estimated from Original-TRMM, downscaled and calibrated-downscaled datasets to capture the daily precipitation variations at the stations. Daily observation from 8 rain gauges were compared with the original TRMM, model-based downscaled, and calibrated-downscaled datasets for 365 days of the year 2010 (Fig. 8b). Daily time series of precipitation from the downscaled and calibrated-downscaled datasets displayed similar variations in most of the days when compared with the observed precipitation over all selected stations. However, daily precipitation from the Original-TRMM has relatively more over- and underestimation which suggested that the current integrated downscaling and calibration procedures could slightly improve the capability of satellite-based precipitation data to capture temporal changes on the daily scale. Moreover, the performance of the daily precipitation with three types of gridded datasets was summarized with the Taylor's diagram in terms of the correlation coefficient, normalized standard deviation, and root mean-square error over the 8 selected stations (Fig. 8c). Among these datasets, downscaled and calibrated-downscaled results outperformed the Original-TRMM with a higher correlation coefficient (CC) and lower standard deviation over most of the selected stations (Fig. 8c). However, the performance of these datasets was still lower at the stations located upstream in the semiarid regions (such as Gilgit, Astore, Kupwara and Dir) where CC values were < 0.5 as compared to downstream stations.

4. Discussion

Retrieval of high-resolution and reliable gridded precipitation estimates depends on the selection of appropriate environmental variables, spatial downscaling models, and methodological framework that is used to merge model predictions with observations from RGS.

Precipitation is a naturally complex phenomenon that is influenced by macro-geographical factors, land surface characteristics, and sources of water vapor. For example, precipitation distribution in the mountainous regions is greatly influenced by topography (Guan et al., 2009; Yin et al., 2008), slope direction (Meersmans et al., 2016), vegetation (Vicente-Serrano et al., 2013; Jia et al., 2011; Liu et al., 2018) surface temperature (Trenberth and Shea, 2005; LeMone et al., 2003) and distance from the sea (Ezzine et al., 2017). Previous studies have used various high-resolution environmental variables across the globe i.e., normalized difference vegetation index (NDVI), normalized difference water index (NDWI), enhanced vegetation index (EVI), elevation, slope, and land surface temperature (LST) (Shi and Song, 2015; Shi et al., 2015; Meersmans et al., 2016; Ma et al., 2017; Ezzine et al., 2017; Chen et al., 2018). Various combinations of these environmental variables have been introduced into the downscaling frameworks to predict the precipitation at a high-resolution (López et al., 2018). Precipitation response to these environmental variables is rather complex (Zhang et al., 2018a) therefore, the selection of appropriate

environmental variables plays an important role in determining model performance (Chen et al., 2020). We introduced seven environmental variables including AET, elevation, slope, NDVI, LST, cloud cover, and wind speed into the spatial downscaling framework. The precipitation response to each of the mentioned environmental variables varies over different climate regions in the study area. In mountainous regions, precipitation increases with elevation followed by a decrease upon exceeding the elevation threshold (Zhang et al., 2018a). Over the entire study region, elevation was found negatively linked with precipitation in general, and spatial pattern of precipitation was largely related to topographic fluctuations i.e., slope and aspects (Jing et al., 2016). According to Sokol and Blížnák (2009), higher elevations induce a higher relative humidity of the air mass as the result of its expansion and cooling, which could be translated into an increased precipitation. In the dry and sub-humid-regions of the study region, a strong linear correlation exists between precipitation and NDVI (precipitation <1200 mm/year). This relationship is slightly weaker in humid-regions where annual precipitation is much above 1200 mm/year (Zhang et al., 2018a; Davenport and Nicholson, 1993). The positive correlation between precipitation and vegetation growth has extensively been studied at the annual scale (Duan and Bastiaanssen, 2013; Zhang et al., 2018a; Wang et al., 2019). The relationship between precipitation and cloud cover varies upon cloud types and geo-locations according to Didier (2015) and Mishra (2019). It should be noted that other environmental variables including actual ET, wind speed, and land surface temperature also affect the precipitation spatial distribution, and utilizing these variables has provided more feasible prediction of the TRMM precipitation at high-resolution over the UIB.

Numerous studies have used the geographically weightage regression (GWR) model to downscale the TRMM precipitation from a coarse resolution to a finer resolution. Traditionally, this model has been applied based on the assumption that the association between the predictor variable (precipitation) and predictands (its environmental factors) varies over space (Zhang et al., 2018b; Chen et al., 2018; Zhan et al., 2018). The relationships might also be fixed (constant over space) for some environmental factors (Zeng et al., 2016). Therefore, retrieving a high-resolution precipitation estimates based on a GWR model has certain limitations especially in UIB where environmental variables are rather complex, diverse, and with different responses to precipitation in different climatic zones. Therefore, in the present study, a GVT was performed for filtering out the spatially fixed (global) and varying (local) variables and a new mixed GWR (commonly known as MGWR) model was proposed to predict precipitation. Similar findings by Zeng et al. (2016) suggest that introducing fixed (global) variables into the traditional GWR model improves prediction accuracy. However, the determination of which characteristics of environmental variables to be involved in the global or local part of the downscaling model is a challenging step. If the variable does not explain precipitation spatial variability it should not be considered as a spatial varying component as used in traditional GWR model. In the present study, we screened out LST and slope as global parameters (spatially constant) and regarded the other five variables as local parameters (spatially varying). In the first attempt, we introduced all these variables into the MGWR model as spatially varying, and in the second attempt, we introduced two listed variables as spatially constant and the remaining five as spatially varying. Based on the results, the MGWR model performed better than the traditional GWR model. The lower performance of the GWR model in predicting precipitation might pertain to the fact that this model takes all the environmental variables into account as spatially varying, while the relationship between precipitation and some environmental factors might be constant. Attempting to fit varying relations in the GWR model, when they are not present will be fitting random noise and thus result in poorer models (Zeng et al., 2016).

Additionally, we adopted two calibration procedures (namely GDA and GRA) to merge the residuals in the downscaled precipitation, as compared to the ground-based measurements, to enhance the reliability

of precipitation prediction in UIB. Merging observations from RGS with model predictions is an essential step for the downscaling-calibration procedure as there may be some inherent errors in TRMM precipitation when downscaled (Immerzeel et al., 2009) which later can be reduced by calibration procedures (Cheema and Bastiaanssen, 2012; Li and Shao, 2010). Findings of the current study revealed that both GRA and GDA methods have proven very simple and effective techniques to compensate for the under-and over-estimation in model-based downscaled data (Shi and Song, 2015; Zhang et al., 2017). The GRA performed better than the GDA in merging ground-based observations with downscaled precipitation data and similar findings were suggested by (Zhang et al. (2018b)). The sparse distribution of rain gauge stations in UIB especially over the Yogo and Kharlong watersheds, limits the capability of calibrating and merging techniques to enhance the reliability of downscaled precipitation data in those regions.

The present study investigated the physical realism of each dataset (including Original-TRMM, model-based downscaled and calibrated-downscaled datasets) in identifying the over- and under-estimation using point-based observations from the rain gauge stations. Precipitation at each grid was extracted from the corresponding station. It should be noted that the model-based downscaled and calibrated-downscaled precipitation datasets at 1 km resolution outperform the Original-TRMM precipitation ($0.25^{\circ} \times 0.25^{\circ}$) over each rain gauge station in the UIB. The high inconsistency between the Original-TRMM precipitation and observations from RGS might be the reason for the availability of a single rain gauge over a satellite pixel of $0.25^{\circ} \times 0.25$ km area which may consist of both rain and non-rain areas and cause random errors (Gebremichael and Krajewski, 2004; Harmsen et al., 2008). Harmsen et al. (2008) reported that rainfall within a single satellite pixel could vary considerably by 38% between two gauges located within 4 km \times 4 km. It illustrates that rainfall measured at a station could deviate by ~19% from the mean of all gauges in the same pixel of an area of 4 km \times 4 km. The deviation between TRMM precipitation and single gauge observation is due to the discrepancy of scale (Gebremichael and Krajewski, 2004) and could be reduced by increasing the validation stations or rescaling the TRMM precipitation to a finer resolution (Immerzeel et al., 2009). (Cheema and Bastiaanssen, 2012) also analyzed the average deviation of TRMM precipitation from gauge observations at different spatial scales in the Indus Basin. They found that the average deviation over a single pixel was much higher than when computed with a cluster of pixels at a smaller scale, which supports our findings.

4.1. Limitations and future directions

In general, the current research primarily focused on mainly improving the spatial (horizontal) information of precipitation from coarse resolution satellite observation over the high-altitude UIB; however, further work is needed to investigate the vertical distribution of precipitation, which is important in hydrological modeling studies over the upper glaciated catchments. Secondly, in most of the upper catchments with complex topography and sparse rain gauge stations, the observed precipitation is unreliable or unavailable; hence, poses a limitation to the application of remote sensing precipitation datasets in such regions. Therefore, the orographic correction of precipitation based on the vertical gradients along with glacier mass balance is required to retrieve an accurate precipitation dataset in high-altitude mountain regions.

5. Conclusions

This paper investigated the performance of an Integrated Downscaling and Calibration (IDAC) framework to generate high-resolution (1 km \times 1 km) gridded precipitation data at annual, monthly, and daily scales over the UIB. A mixed geographically weighted regression (MGWR) model was proposed for downscaling and the results were

compared with the traditional GWR model. Overall findings of the current study are;

- Based on GVT, slope and LST were filtered out as fixed variables while elevation, AET, NDVI, wind speed, and cloud cover were considered to be spatially varying variables.
- Precipitation predicted by the MGWR model had a better accuracy when compared with the GWR model taking the Original-TRMM and ground observation data as the reference.
- It was seen that introducing all environmental variables as spatially varying in the traditional GWR model generates random noise and thus results in an unsatisfactory performance. Model accuracy can be improved by introducing mixed coefficients (varying and fixed) according to the environmental factors variability as implemented in MGWR compared with the GWR which consider only spatially varying coefficients.
- Downscaled results estimated by the MGWR model not only preserve precipitation spatial pattern but also provide more information at a finer scale over the UIB.
- The accuracy of the downscaling results was significantly improved after merging with ground observations using both GDA and GRA calibration procedures; the GRA outperformed the GDA method in terms of the validation metrics calculated.

In general, the proposed IDAC approach is suitable for retrieving gridded data at annual, monthly, and daily scales over the UIB with varying climate and complex topography.

CRediT authorship contribution statement

Arfan Arshad: Conceptualization, Methodology, Software, Data curation, Formal Analysis, Writing- Original draft preparation. **Wanchang Zhang:** Visualization, Investigation, Supervision; **Arfan Arshad and Wanchang Zhang:** re-arranging the data formulation during major revision; **Arfan Arshad, Wanchang Zhang, Zhijie Zhang, Shuhang Wang, Bo Zhang, Muhammad Jehanzeb Masud Cheema, and Masoud Jafari Shalamzari:** Reviewing, Editing and Proofreading.

Declaration of competing interest

Authors declare no conflicts of interest.

Acknowledgment

The authors are grateful to the Water and Power Development Authority (WAPDA), Pakistan Meteorological Department (PMD), and China Metrological Department (CMD) for providing observed hydro-meteorological data in the upper Indus basin. This study was jointly financed by the National Key R & D Program of China [Grant No. 2016YFA0602302] and the Key R & D and Transformation Program of Qinghai Province [Grant No. 2020-SF-C37]. We appreciate the comments by two anonymous reviewers and Editors to improve the quality of manuscript's content.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2021.147140>.

References

- Adhikary, S.K., Yilmaz, A.G., Muttill, N., 2015. Optimal design of rain gauge network in the Middle Yarra River catchment, Australia. *Hydrol. Process.* 29, 2582–2599.
 Alexakis, D., Tsanis, I., 2016. Comparison of multiple linear regression and artificial neural network models for downscaling TRMM precipitation products using MODIS data. *Environ. Earth Sci.* 75, 1077.
 Back, L.E., Bretherton, C.S., 2005. The relationship between wind speed and precipitation in the Pacific ITCZ. *J. Clim.* 18, 4317–4328.

- Bookhagen, B., Burbank, D.W., 2006. Topography, relief, and TRMM-derived rainfall variations along the Himalaya. *Geophys. Res. Lett.* 33.
- Chao, L., Zhang, K., Li, Z., Zhu, Y., Wang, J., Yu, Z., 2018. Geographically weighted regression based methods for merging satellite and gauge precipitation. *J. Hydrol.* 558, 275–289.
- Cheema, M.J.M., Bastiaanssen, W.G., 2012. Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *Int. J. Remote Sens.* 33, 2603–2627.
- Chen, F., Liu, Y., Liu, Q., Li, X., 2014. Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity. *Int. J. Remote Sens.* 35, 3074–3093.
- Chen, Y., Huang, J., Sheng, S., Mansaray, L.R., Liu, Z., Wu, H., Wang, X., 2018. A new downscaling-integration framework for high-resolution monthly precipitation estimates: combining rain gauge observations, satellite-derived precipitation data and geographical ancillary data. *Remote Sens. Environ.* 214, 154–172.
- Chen, F., Gao, Y., Wang, Y., Li, X., 2020. A downscaling-merging method for high-resolution daily precipitation estimation. *J. Hydrol.* 581, 124414.
- Dahri, Z.H., Ludwig, F., Moors, E., Ahmad, B., Khan, A., Kabat, P., 2016. An appraisal of precipitation distribution in the high-altitude catchments of the Indus basin. *Sci. Total Environ.* 548, 289–306.
- Darand, M., Amanollahi, J., Zandkarimi, S., 2017. Evaluation of the performance of TRMM multi-satellite precipitation analysis (TMPA) estimation over Iran. *Atmos. Res.* 190, 121–127.
- Davenport, M.L., Nicholson, S.E., 1993. On the relation between rainfall and the normalized difference vegetation index for diverse vegetation types in East Africa. *Int. J. Remote Sens.* 14, 2369–2389.
- Didier, N., 2015. Comparison of spatial and temporal cloud coverage derived from CloudSat, CERES, ISCCP and their relationship with precipitation over Africa. *Am. J. Remote Sens.* 3, 17–28.
- Duan, Z., Bastiaanssen, W., 2013. First results from version 7 TRMM 3B43 precipitation product in combination with a new downscaling-calibration procedure. *Remote Sens. Environ.* 131, 1–13.
- Ezzine, H., Bouziane, A., Ouazar, D., Hasnaoui, M.D., 2017. Downscaling of open coarse precipitation data through spatial and statistical analysis, integrating NDVI, NDWI, elevation, and distance from sea. *Adv. Meteorol.* 2017.
- Fang, J., Du, J., Xu, W., Shi, P., Li, M., Ming, X., 2013. Spatial downscaling of TRMM precipitation data based on the orographical effect and meteorological conditions in a mountainous area. *Adv. Water Resour.* 61, 42–50.
- Foody, G., 2003. Geographical weighting as a further refinement to regression modelling: an example focused on the NDVI–rainfall relationship. *Remote Sens. Environ.* 88, 283–293.
- Fotheringham, A.S., Charlton, M.E., Brunsdon, C., 1998. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environ. Plan A* 30, 1905–1927.
- Fotheringham, A.S., Brunsdon, C., Charlton, M., 2003. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons.
- Gao, J., Li, S., 2011. Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using geographically weighted regression. *Appl. Geogr.* 31, 292–302.
- Gao, Z., Long, D., Tang, G., Zeng, C., Huang, J., Hong, Y., 2017. Assessing the potential of satellite-based precipitation estimates for flood frequency analysis in ungauged or poorly gauged tributaries of China's Yangtze River basin. *J. Hydrol.* 550, 478–496.
- Gebremichael, M., Krajewski, W.F., 2004. Assessment of the statistical characterization of small-scale rainfall variability from radar: analysis of TRMM ground validation datasets. *J. Appl. Meteorol.* 43, 1180–1199.
- Gebremichael, M., Anagnostou, E.N., Bitew, M.M., 2010. Critical steps for continuing advancement of satellite rainfall applications for surface hydrology in the Nile River basin 1. *J. Afr. Water Res.* 46, 361–366.
- Guan, H., Wilson, J.L., Xie, H., 2009. A cluster-optimizing regression-based approach for precipitation spatial downscaling in mountainous terrain. *J. Hydrol.* 375, 578–588.
- Hadria, R., Boudhar, A., Ouatiki, H., Lebrini, Y., Elmansiouri, L., Gadaoui, F., Lionboui, H., Benabdellouahab, T., 2019. Combining use of TRMM and ground observations of annual precipitations for meteorological drought trends monitoring in Morocco. *Am. J. Remote Sens.* 7, 25–34.
- Harmsen, E.W., Mesa, S.G., Cabassa, E., Ramírez-Beltran, N.D., Pol, S.C., Kuligowski, R.J., Vasquez, R., 2008. Satellite sub-pixel rainfall variability. *Int. J. Syst. Appl. Eng. Dev.* 2, 91–100.
- Hasson, S., Lucarini, V., Khan, M.R., Petitta, M., Bolch, T., Gioli, G., 2014. Early 21st century snow cover state over the western river basins of the Indus River system. *Hydroclim. Earth Syst. Sci.* 18, 4077–4100.
- Henn, B., Newman, A.J., Livneh, B., Daly, C., Lundquist, J.D., 2018. An assessment of differences in gridded precipitation datasets in complex terrain. *J. Hydrol.* 556, 1205–1219.
- Huffman, G.J., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., Schneider, U., 1997. The global precipitation climatology project (GPCP) combined precipitation dataset. *Bull. Am. Meteorol. Soc.* 78, 5–20.
- Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* 8, 38–55.
- Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., 2009. Improving the global precipitation record: GPCP version 2.1. *Geophys. Res. Lett.* 36.
- Immerzeel, W., Rutten, M., Droogers, P., 2009. Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sens. Environ.* 113, 362–370.
- Jia, S., Zhu, W., Lü, A., Yan, T., 2011. A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sens. Environ.* 115, 3069–3079.
- Jing, W., Yang, Y., Yue, X., Zhao, X., 2016. A comparison of different regression algorithms for downscaling monthly satellite-based precipitation over North China. *Remote Sens.* 8, 835.
- Kidd, C., Bauer, P., Turk, J., Huffman, C., Joyce, R., Hsu, K.-L., Braithwaite, D., 2012. Inter-comparison of high-resolution precipitation products over Northwest Europe. *J. Hydrometeorol.* 13, 67–83.
- Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y.N., Ushio, T., Nakagawa, K., 2007. Global precipitation map using satellite-borne microwave radiometers by the GSMap project: production and validation. *IEEE Trans. Geosci. Remote Sens.* 45, 2259–2275.
- Kumari, M., Singh, C.K., Bakimchandra, O., Basistha, A., 2017. Geographically weighted regression based quantification of rainfall-topography relationship and rainfall gradient in Central Himalayas. *Int. J. Climatol.* 37, 1299–1309.
- Langella, G., Basile, A., Bonfante, A., Terribile, F., 2010. High-resolution space-time rainfall analysis using integrated ANN inference systems. *J. Hydrol.* 387, 328–342.
- Lemone, M.A., Grossman, R.L., Chen, F., Ikeda, K., Yates, D., 2003. Choosing the averaging interval for comparison of observed and modeled fluxes along aircraft transects over a heterogeneous surface. *J. Hydrometeorol.* 4, 179–195.
- Li, M., Shao, Q., 2010. An improved statistical approach to merge satellite rainfall estimates and raingauge data. *J. Hydrol.* 385, 51–64.
- Liu, J., Zhang, W., Nie, N., 2018. Spatial downscaling of TRMM precipitation data using an optimal subset regression model with NDVI and terrain factors in the Yarlung Zangbo river basin, China. *Adv. Meteorol.* 2018.
- López, P.L., Immerzeel, W.W., Sandoval, E.A.R., Sterk, G., Schellekens, J., 2018. Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the Magdalena River basin in Colombia. *Front. Earth Sci.* 6, 68.
- Ma, Z., Zhou, Y., Hu, B., Liang, Z., Shi, Z., 2017. Downscaling annual precipitation with TMPA and land surface characteristics in China. *Int. J. Climatol.* 37, 5107–5119.
- Ma, Z., Tan, X., Yang, Y., Chen, X., Kan, G., Ji, X., Lu, H., Long, J., Cui, Y., Hong, Y., 2018. The first comparisons of IMERG and the downscaled results based on IMERG in hydrological utility over the Ganjiang River Basin. *Water* 10, 1392.
- Meersmans, J., Van Weverberg, K., De Baets, S., Ridder, De, Palmer, F., Van Wesemael, B., Quine, T., 2016. Mapping mean total annual precipitation in Belgium, by investigating the scale of topographic control at the regional scale. *J. Hydrol.* 540, 96–105.
- Mishra, A.K., 2019. Investigating changes in cloud cover using the long-term record of precipitation extremes. *Meteorol. Appl.* 26, 108–116.
- Mukherjee, S., Joshi, R., Prasad, R.C., Vishvakarma, S.C., Kumar, K., 2015. Summer monsoon rainfall trends in the Indian Himalayan region. *Theor. Appl. Climatol.* 121, 789–802.
- Nakaya, T., 2015. Geographically Weighted Regression (GWR) Software. GWR 4.0. ASU GeoDa Center website.
- Rizwan, M., Li, X., Jamal, K., Chen, Y., Chaudhary, J.N., Zheng, D., Anjum, L., Ran, Y., Pan, X., 2019. Precipitation variations under a changing climate from 1961–2015 in the source region of the Indus River. *Water* 11, 1366.
- Seyoum, W.M., 2018. Characterizing water storage trends and regional climate influence using GRACE observation and satellite altimetry data in the Upper Blue Nile River Basin. *J. Hydrol.* 566, 274–284.
- Shafeeqe, M., Yi, L., 2020. A tri-approach for diagnosing gridded precipitation datasets for watershed glacio-hydrological simulation in mountain regions. *Hydrology and Earth System Sciences Discussions*, pp. 1–49.
- Shafeeqe, M., Luo, Y., Wang, X., Sun, L., 2019. Revealing vertical distribution of precipitation in the glacierized upper indus basin based on multiple datasets. *J. Hydrometeorol.* 20, 2291–2314.
- Shafeeqe, M., Luo, Y., Wang, X., Sun, L., 2020. Altitudinal distribution of meltwater and its effects on glacio-hydrology in glacierized catchments, Central Asia. *J. Afr. Water Res.* 56, 30–52.
- Shi, Y., Song, L., 2015. Spatial downscaling of monthly TRMM precipitation based on EVI and other geospatial variables over the Tibetan Plateau from 2001 to 2012. *Mt. Res. Dev.* 35, 180–194.
- Shi, Y., Song, L., Xia, Z., Lin, Y., Myneni, R.B., Choi, S., Wang, L., Ni, X., Lao, C., Yang, F., 2015. Mapping annual precipitation across mainland China in the period 2001–2010 from TRMM3B43 product using spatial downscaling approach. *Remote Sens.* 7, 5849–5878.
- Sokol, Z., Blížnák, V., 2009. Areal distribution and precipitation–altitude relationship of heavy short-term precipitation in the Czech Republic in the warm part of the year. *Atmos. Res.* 94, 652–662.
- Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* 81, 2035–2046.
- Tahir, A.A., Chevallier, P., Arnaud, Y., Neppel, L., Ahmad, B., 2011. Modeling snowmelt-runoff under climate scenarios in the Hunza River basin, Karakoram Range, Northern Pakistan. *J. Hydrol.* 409, 104–117.
- Tang, G., Long, D., Hong, Y., 2016. Systematic anomalies over inland water bodies of High Mountain Asia in TRMM precipitation estimates: no longer a problem for the GPM era? *IEEE Geosci. Remote Sens. Lett.* 13, 1762–1766.
- Trenberth, K.E., Shea, D.J., 2005. Relationships between precipitation and surface temperature. *Geophys. Res. Lett.* 32.
- Vicente-Serrano, S.M., Gouveia, C., Camarero, J.J., Beguería, S., Trigo, R., López-Moreno, J.I., Azorín-Molina, C., Pasho, E., Lorenzo-Lacruz, J., Revuelto, J., 2013. Response of vegetation to drought time-scales across global land biomes. *Proc. Natl. Acad. Sci.* 110, 52–57.
- Wang, W., Lu, H., Yang, D., Sothea, K., Jiao, Y., Gao, B., Peng, X., Pang, Z., 2016. Modelling hydrologic processes in the Mekong River Basin using a distributed model driven by satellite precipitation and rain gauge observations. *PLoS One* 11, e0152229.
- Wang, L., Chen, R., Han, C., Yang, Y., Liu, J., Liu, Z., Wang, X., Liu, G., Guo, S., 2019. An improved spatial-temporal downscaling method for TRMM precipitation datasets in alpine regions: a case study in northwestern China's Qilian mountains. *Remote Sens.* 11, 870.

- Xu, S., Wu, C., Wang, L., Gonsamo, A., Shen, Y., Niu, Z., 2015. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. *Remote Sens. Environ.* 162, 119–140.
- Yaduvanshi, A., Srivastava, P.K., Pandey, A., 2015. Integrating TRMM and MODIS satellite with socio-economic vulnerability for monitoring drought risk over a tropical region of India. *R&J.C. 14–27*.
- Yang, Z., Zhang, Q., & Hao, X. 2016. Evapotranspiration trend and its relationship with precipitation over the loess plateau during the last three decades. *Adv. Meteorol.*, 2016.
- Yang, N., Zhang, K., Hong, Y., Zhao, Q., Huang, Q., Xu, Y., Xue, X., Chen, S., 2017. Evaluation of the TRMM multisatellite precipitation analysis and its applicability in supporting reservoir operation and water resources management in Hanjiang basin, China. *J. Hydrol.* 549, 313–325.
- Yin, Z.-Y., Zhang, X., Liu, X., Colella, M., Chen, X., 2008. An assessment of the biases of satellite rainfall estimates over the Tibetan Plateau and correction methods based on topographic analysis. *J. Hydrometeorol.* 9, 301–326.
- Zeng, C., Yang, L., Zhu, A.-X., Rossiter, D.G., Liu, J., Liu, J., Qin, C., Wang, D., 2016. Mapping soil organic matter concentration at different scales using a mixed geographically weighted regression method. *Geoderma* 281, 69–82.
- Zhan, C., Han, J., Hu, S., Liu, L. & Dong, Y. 2018. Spatial downscaling of GPM annual and monthly precipitation using regression-based algorithms in a mountainous area. *Adv. Meteorol.*, 2018.
- Zhang, Q., Shi, P., Singh, V.P., Fan, K., Huang, J., 2017. Spatial downscaling of TRMM-based precipitation data using vegetative response in Xinjiang, China. *Int. J. Climatol.* 37, 3895–3909.
- Zhang, T., Li, B., Yuan, Y., Gao, X., Sun, Q., Xu, L., Jiang, Y., 2018a. Spatial downscaling of TRMM precipitation data considering the impacts of macro-geographical factors and local elevation in the Three-River headwaters region. *Remote Sens. Environ.* 215, 109–127.
- Zhang, Y., Li, Y., Ji, X., Luo, X., Li, X., 2018b. Fine-resolution precipitation mapping in a mountainous watershed: geostatistical downscaling of TRMM products based on environmental variables. *Remote Sens.* 10, 119.
- Zhang, H., Loáiciga, H.A., Ha, D., Du, Q., 2020. Spatial and temporal downscaling of TRMM precipitation with novel algorithms. *J. Hydrometeorol.* 21, 1259–1278.
- Zheng, X., Zhu, J., 2015. A methodological approach for spatial downscaling of TRMM precipitation data in North China. *Int. J. Remote Sens.* 36, 144–169.
- Zhou, X., Ni, G.-H., Shen, C., Sun, T., 2017. Remapping annual precipitation in mountainous areas based on vegetation patterns: a case study in the Nu River basin. *Hydrol. Earth Syst. Sci.* 21, 999–1015.