7th Conference on Sustainability in Civil Engineering (CSCE'25)
(An International Conference)
Department of Civil Engineering
Capital University of Science and Technology, Islamabad Pakistan

SPATIAL DOWNSCALING OF GRACE DATASET USING RANDOM FOREST APPROACH FOR A BETTER UNDERSTANDING OF HYDROLOGICAL DROUGHTS IN PAKISTAN

[a] Mustajab Ali, [b] Usman Ali*

a: Department of Civil Engineering, Mirpur University of Science and Technology (MUST), Mirpur Azad Kashmir, Pakistan.
mustajab.ce@must.edu.pk
b: Department of Civil Engineering, Mirpur University of Science and Technology (MUST), Mirpur Azad Kashmir, Pakistan.
aliu70536@gmail.com
*Corresponding author

Abstract- Climate change and increasing urbanization along with anthropogenic activities has intensified the risk of hydrological droughts in a water stressed country like Pakistan. The overall goal of this study; i) to create a Random Forest machine learning framework to downscale GRACE based Terrestrial Water Storage Anomalies (TWSA) from 1° down to 0.25° spatial resolution, and ii) to assess drought frequency and intensity in various regions of Pakistan using an enhanced Terrestrial Water Storage dataset and the TWS-Drought Severity Index (TWS-DSI). A high-accuracy Random Forest model (R^2 = 0.9967) was trained on 1° resolution TWSA data and applied to finer resolution inputs to derive 0.25° TWSA estimates. The downscaled dataset revealed a sharp decline in water storage from 2011 onwards, with extreme drought conditions (TWS-DSI ≤ -1.6) observed post-2020 in regions such as Sindh and southern Punjab. The results obtained highlight the possibilities of ML being used for hydrological monitoring as well as the necessity of cohesive water resources planning that promotes Sustainable Development Goals 6 (Clean Water) and 13 (Climate Action).

Keywords- GRACE, Hydrological Droughts, Random Forest, TWSA, Pakistan

1 Introduction

Terrestrial water storage covers all continental water available in tree canopies, snow, rivers, lakes or reservoirs, wetlands, and groundwater and is a key aspect of the water and energy balance of the world [1]. It is instrumental in influencing the availability of water resources [2] and regulating water flux interactions among diverse components of the Earth system [3]. Additionally, fluctuations in TWS are intrinsically connected to occurrences of droughts [3, 1], floods [4], and variations in global sea levels [5]. In spite of its significance, global TWS has been comparatively under-explored in relation to hydrological fluxes (such as river flows, evapotranspiration, and groundwater flow) due to insufficient large-scale observations and the complexities associated with explicitly delineating all TWS components within hydrological modeling [1]. The Gravity Recovery and Climate Experiment (GRACE) satellite mission, which was specifically designed to enable the comprehensive monitoring of Earth's gravitational field variations. This has consequently provided researchers with enhanced opportunities to refine and substantiate the terrestrial water storage (TWS) simulations that are integral components of these advanced hydrological models [1][2][5]. The GRACE satellite's TWS data, in conjunction with various model simulations, has been employed for an extensive array of applications that encompass diverse fields, including the critical evaluation of water resource availability [2], the thorough examination of anthropogenic influences on the hydrological cycle [1], the quantification of aquifer depletion rates, the vigilant monitoring of drought conditions, and the thorough assessment of potential flood risks associated with climatic fluctuations [3].

Drought, characterized as a gradually developing phenomenon, ranks among the most expensive natural disasters, directly impacting water resources, agricultural growth, socioeconomic advancement, and ecosystem vitality, and is frequently associated with armed conflicts [6]. A straightforward definition of drought onset involves a decrease in the atmospheric/climatic water balance (commonly referred to as meteorological drought); however, it can evolve into a more intricate natural hazard. For example, a sustained deficit in the climatic water balance may result in diminished streamflow, leading to a hydrological drought [7]; alternatively, it could cause a significant decrease in soil moisture, thereby precipitating an agricultural drought [8]. While meteorological, hydrological, and agricultural droughts are immediate manifestations of reduced water within the hydrological cycle [9]. These droughts subsequently engender numerous detrimental socio-economic, socio-political, and environmental repercussions, including but not limited to water insecurity, food insecurity, economic detriment, and deterioration of water quality, all of which are anticipated to intensify in a progressively warmer future [10].

Anticipated alterations in climatic conditions are expected to exacerbate both the intensity and frequency of droughts globally, driven by transformations in precipitation patterns and rising temperatures [11]. Similar to various other locales, numerous regions in Pakistan face challenges stemming from hydro-meteorological anomalies [12]. Prolonged dry spells may yield catastrophic consequences, negatively affecting water resources and agriculture [13]. Recognized as a region characterized by arid climates, Pakistan experiences scant rainfall coupled with high temperatures [14]. Climate change has precipitated significant repercussions for both the socioeconomic and environmental landscapes in Pakistan [11], as well as in its neighboring regions within Southwest Asia.

An extensive body of literature has emerged focusing on drought assessment utilizing indices such as the standardized precipitation index (SPI), Palmer's drought severity index (PDSI), a standardized runoff index (SRI) and soil moisture drought index (SMI) [10, 1]. These traditional indices were employed to monitor and forecast meteorological, agricultural, and hydrological droughts. Recently, a novel drought index, the TWS drought severity index (TWS-DSI) [15], has been introduced to analyze droughts concerning the vertically available water storage, contrasting with the individual storages or fluxes utilized in conventional indices.

Our research introduces an innovative methodology for assessing drought severity that was previously absent in earlier investigations. The principal objectives of this investigation include (i) the formulation of a Random Forest-based machine learning framework to refine the spatial granularity of GRACE-derived Terrestrial Water Storage Anomalies (TWSA), enhancing its resolution from 1° to 0.25°, and (ii) the application of the improved TWSA dataset to assess drought frequency and intensity across different regions of Pakistan using the TWS-Drought Severity Index (TWS-DSI), in support of the aims outlined in Sustainable Development Goals 6 and 13.

2 Research Methodology

2.1 Study Area

Pakistan is located at latitude 23.5°–37.5° N and longitude 62°–75° E with an area equals to 803,940 km^2(Figure 1). Pakistan's hydrology is largely shaped by the Indhetics5 tributary rivers of Indus, Jhelum, Chenab answered, Ravi, and Sutlej, which constitute the Indus Basin and host nearly 90% of irrigated agricultural production [16]. The three main climatic regions of the country are arid, semi-arid, and sub-humid. Rainfall is also irregular, with annual precipitation of less than 200 mm in Balochistan, while in the northern part this surpasses 1,500 mm [11, 14]. Previous episodes of drought especially between 1998-2002 have had catastrophic socio-economic impacts displacing more than 3 million people and generating $2.5 billion in losses [6]. Due to inadequate surface water supplies, farmers are fulfilling their escalating demands through the unregulated overexploitation of groundwater in Pakistan [11]. The concurrent utilization of both groundwater and surface water is occurring on nearly 70% of the irrigated agricultural lands.

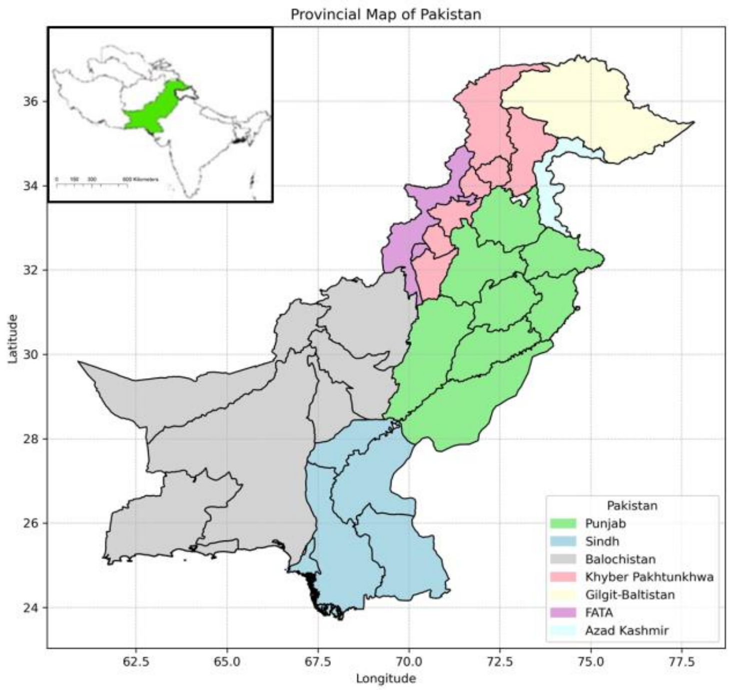

Figure 1 Map of the study area

2.2 Data sources

In the current study, we employed Total Water Storage (TWS) data sourced from the Gravity Recovery and Climate Experiment (GRACE). The GRACE satellite initiative, a collaborative effort between NASA and the German Aerospace Center (DLR), is meticulously structured to evaluate variations in the Earth's gravitational field, thus enabling the monitoring of global changes in the available total water storage. Currently, the gravitational data obtained from GRACE is predominantly disseminated by several processing organizations, including the German Research Center for Geosciences (GFZ), the Jet Propulsion Laboratory (JPL), and the Center for Space Research at the University of Texas (CSR), sustaining a spatial resolution of 1° x 1° [17, 18]. The GRACE data (RL06, level-3) is available at (https://grace.jpl.nasa.gov/data/get-data/) and is presented in terms of equivalent water height, which signifies monthly total water storage [19]. In order to improve the accuracy of the TWSA estimates, we synthesized these three data products. A range of pre-processing techniques were applied to these datasets, including technique called de-stripping filter, Gaussian smoothing, as well as glacier isostatic adjustment (GIA) [20].

2.3 Methodology

2.3.1 GRACE-DSI

For every grid cell, we calculated the GRACE Drought Severity Index (GRACE-DSI) by dividing the GRACE TWS anomaly for month j and year i by the mean of those anomalies represented as $\text{GRACE-DSI}_{i,j} = (\text{TWS}_{i,j} - \langle \text{TWS}_j \rangle)/\sigma_j$, where i spans from 2002 to 2024, and $\langle \text{TWS}_j \rangle$ and $\sigma_j$for the same month and applying the standard deviation [15].

The global GRACE-DSI data follow pseudo-standard normal distribution. Using data from GRACE-DSI, we classified drought into five categories (see Table 1) indicating that certain levels of drought can be recognized by the magnitude of a TWS deficit in both dry and wet regions.

Table 1 Drought classification on the basis of GRACE DSI [15]

Category	Description	GRACE-DSI
W4	Extremely wet	2.0 or greater
W3	Very wet	1.60 to 1.99
W2	Moderately wet	1.30 to 1.59
W1	Slightly wet	0.80 to 1.29
W0	Near normal	0.50 to 0.79
D0	Abnormally dry	-0.50 to -0.79
D1	Moderate drought	-0.80 to -1.29
D2	Severe drought	-1.30 to -1.59
D3	Extreme drought	-1.60 to -1.99
D4	Exceptional drought	-2.0 or less

2.4 Model Design

For machine learning implementation, a Random Forest model was employed to downscale GRACE TWSA. Initially, the model was trained using coarse-resolution (1°) TWSA as the target and multiple resampled predictor variables the same resolution. Once trained, the model was applied to 0.25° resolution inputs to predict downscaled TWSA [20].The dataset was split into 70% for training and 30% for validation. This means that the model learned patterns from 70% of the data, while its performance was tested on the remaining 30% to ensure generalizability. This split helps to identify whether the model is over fitting or can truly predict on unseen data.

Additionally, 5-fold cross-validation was used to enhance the model's reliability. In this technique, the dataset is divided into five equal parts (folds). The model is trained on four folds and validated on the fifth, and this process is repeated five times so that each fold is used once for validation. The results are then averaged across all five iterations. This approach ensures that the model's performance is stable and not overly dependent on how the data is split, making it ideal for small or noisy datasets. Residuals at the coarse scale were calculated by comparing model predictions with actual GRACE data, and the same trained model was then used to generate downscaled TWSA estimates from high-resolution inputs.

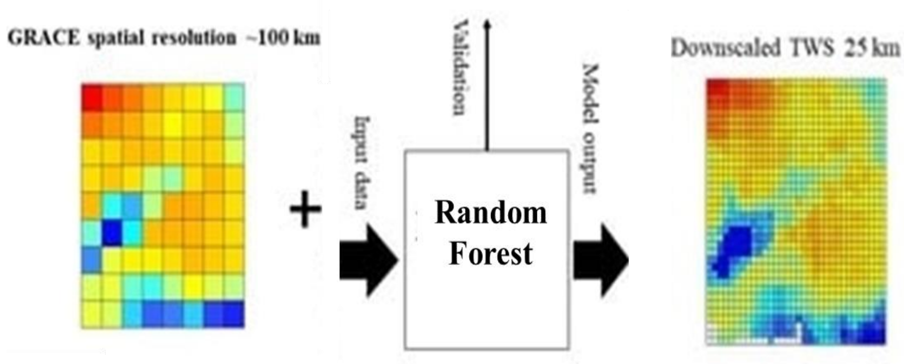

Figure 2 Illustration of the spatial downscaling of GRACE dataset from 1° to 0.25 °

3 Results

3.1 Downscaled GRACE time series

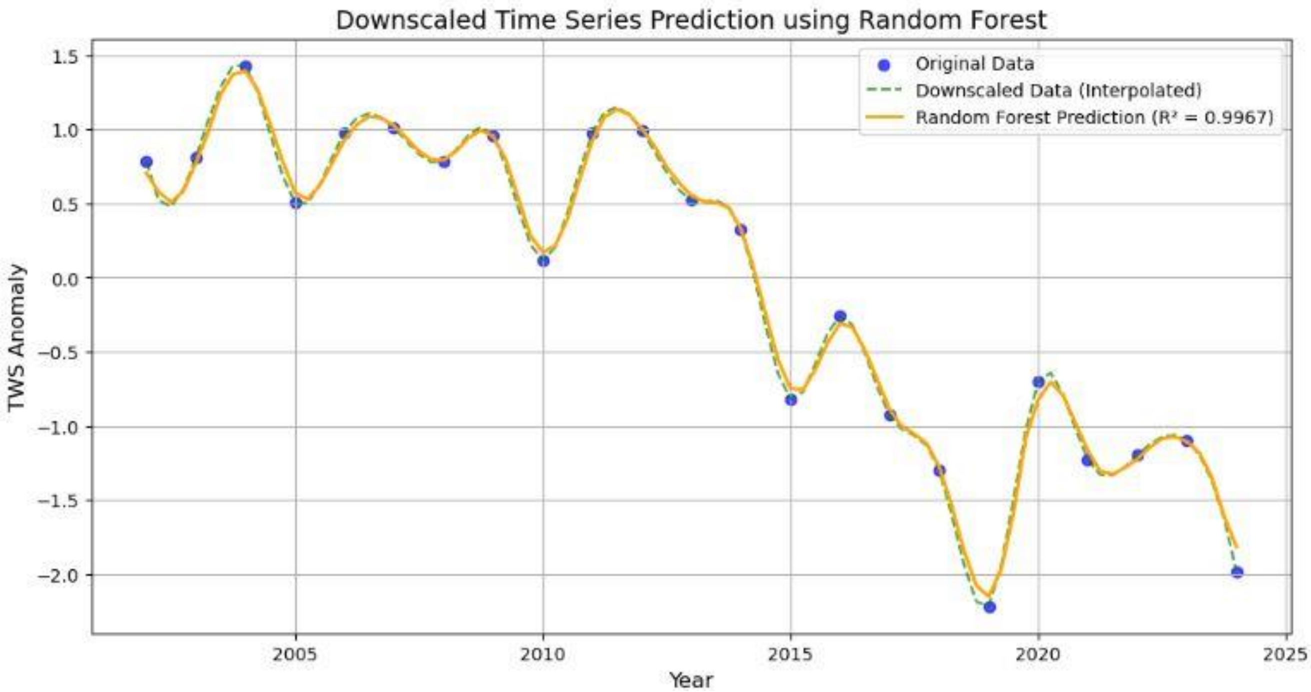

Figure 3Hydrological drought assessment in Pakistan using downscaled GRACE Dataset

The downscaled GRACE Terrestrial Water Storage (TWS) anomaly time series(Figure 3) show interannual to multi-decadal variability from 2002 to 2024 and a clear long-term downward trend as also stated by [1]. The TWS anomaly was characterized by a strong positive signal as seen from the TWS-DSI values (Table 1) during the first period, from 2002 to approximately 2010, with a recurring seasonality in both peaks and lows that indicates stable conditions that are relatively wet. But, as of 2011 there is a gradual downward trend that indicates a shift to dryer conditions beginning at that date, which could be attributed to lower levels of precipitation as well as more water being drained [19] . The most significant in this regard is the increase in the period 2015-2020 in which TWS anomalies have sharply decreased below -1.0(Table 1), indicating extreme and constant wetland hydrological drought in this period. This is consistent with GRACE-based studies [15]& [1] on Drought Severity Index, DSI, at the global scale, which also mark relevant drought events in the period. There seems to be a slight and transitory recovery starting around 2020, potentially due to higher precipitations [21] or managed recharge. The anomaly sharply declines again after 2022 and falls under -2.0 (Table 1) in 2024 indicating that drought conditions are re-establishing and perhaps intensifying [19].

The overall trend is a negative net change of water storage on land in the last two decades, in agreement with recent literature on the subject that highlights the global increase of hydrological droughts. The random forest model reported very high R^2 = 0.9967 which confirms the potential of machine learning to improve temporal resolution and predictions of GRACE TWS data for monitoring water resources.

3.1 Spatial plot

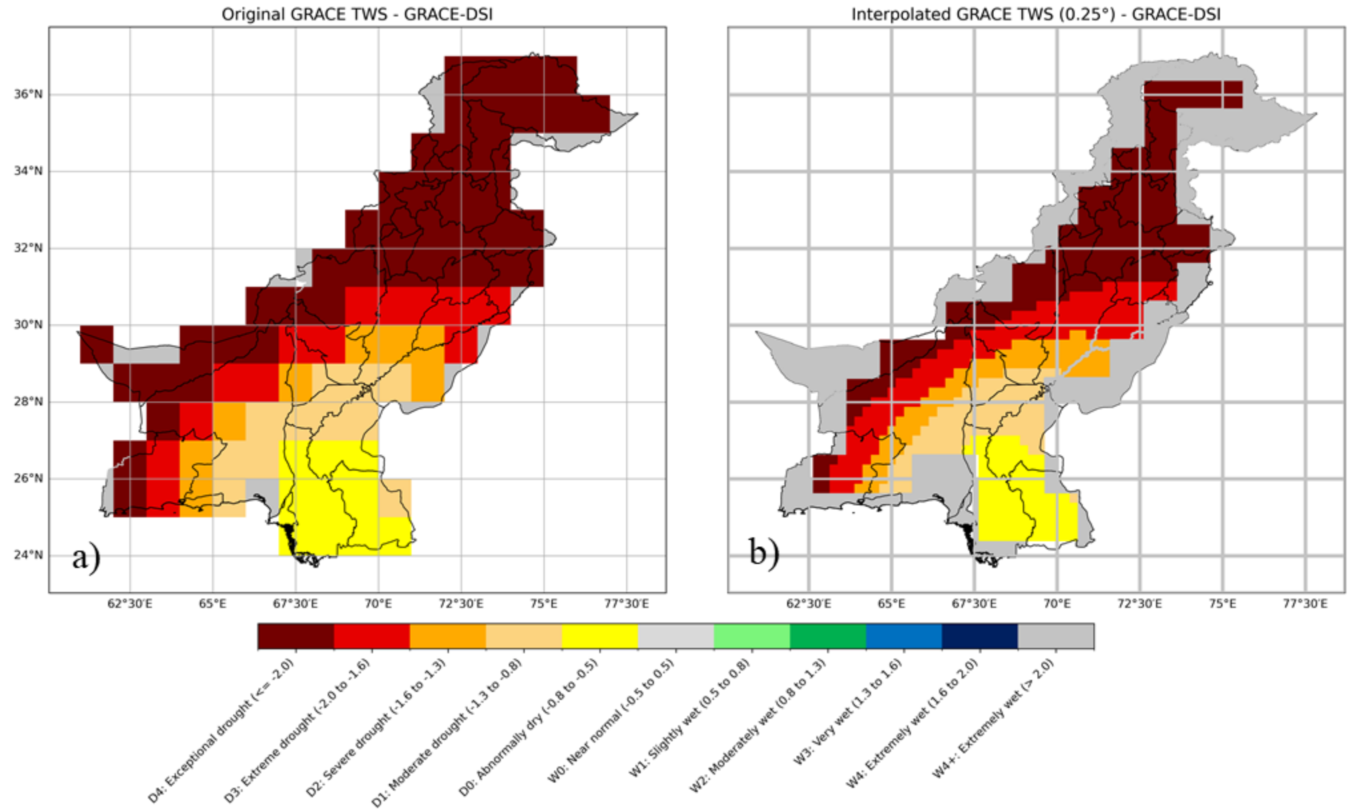

Figure 4 GRACE TWS DSI mapping of Pakistan a) 1° GRACE TWS DSI mapping of Pakistan b) 0.25° downscaled GRACE TWS DSI mapping of Pakistan

GRACE TWS DSI provides a clear perspective on the regional distribution across the country, with areas of significant drought conditions in the northern and western Pakistan contrasting with the south-eastern part of the Indus Basin experiencing significant moisture anomalies [19] as seen in Figure 4. Parts of the north, including Khyber Pakhtunkhwa and Gilgit-Baltistan, have been experiencing severe to exceptional hydrological drought (D3–D4) in reference to Table 1& Figure 4 largely caused by below normal winter precipitation which is typically associated with enhanced western disturbances that are expected to become weaker during years of strong El Niño conditions [22]. Indeed, a strong El Niño pattern prevailed during the 2023-24 winter, coinciding with lower-than-average snowfall and subsequent water storage in the cryosphere across the Hindu Kush-Himalaya, otherwise Pakistan's main water towers [23].

Balochistan in the west experienced similar trends with (D2–D3) in reference to Table 1& Figure 4 droughts, continuing the pattern of chronic susceptibility to winter rainfall failures, and confirming its risky dependence on erratic precipitation [24]. Punjab's role as a transition zone presented as moderate anomalies (D0-D1) in reference to Table 1& Figure 4 as a result of decreased inputs upstream and continued depletion of groundwater reserves due to high abstraction for agriculture, a pattern already well-established in the literature on based groundwater studies in South Asia [11]. Contrarily, residual wetness (W1-W2) in reference to Table 1& Figure 4 in the southeast province of Sindh corresponded to the hydrological signature of the 2022 super floods, that saw the recharging of subsurface aquifers and soils with water, as previously detected by GRACE sensors after flooding thresholds . This pattern of TWS anomalies, which is not homogeneous across space, reinforces the relevance of both short-term atmospheric fluctuations and long-term hydrological memory, and highlights once more the necessity of regionally tailored water management policies.

4 Practical Implementation

By observing the trends, the increased intensity of hydrological droughts can be assessed, and mitigation strategies can be adopted. The results of the current study provide useful information that complements national adaptation efforts in line with the United Nations Sustainable Development Goals (SDGs) of 2030, particularly to SDG 6, Clean Water and Sanitation and SDG 13, Climate Action.

5 Conclusion

The following conclusions can be drawn from the conducted study:

1. Our analysis reveals a significant decreasing trend of TWS after 2011; with the most extreme drought conditions (TWS-DSI $\leq$ -1.6) observed post-2020. This marks a critical hydrological shift, especially for southern regions such as Sindh and southern Punjab, underscoring a growing risk of prolonged water scarcity in already vulnerable areas.
2. The TWS-DSI spatial distribution highlights a stark regional contrast: severe to exceptional drought conditions dominate northern and western Pakistan due to weakened winter precipitation influenced by El Niño patterns, while southeastern Sindh shows residual wetness linked to the long-lasting effects of the 2022 super floods.
3. These trends are closely associated with the multiple drivers of climate extremes and increase in human water extractions, particularly groundwater extraction and river regulation via reservoirs.
4. With an R^2 of 0.9967, the Random Forest model demonstrated a high predictive accuracy and proved the potential of machine learning to improve the resolution of satellite-based tools for the monitoring of hydrology, such as GRACE.
5. Our results represent a transition to a more anthropogenically mediated hydrological system, with a particular focus on the sensitivity of the Indus River basin to climatic stress and provide actionable insights for climate-resilient water governance, directly supporting Sustainable Development Goals 6 (Clean Water and Sanitation) and 13 (Climate Action).

Reference

[1] Pokhrel et al., "Pokhrel, Y., Felfelani, F., Satoh, Y., Boulange, J., Burek, P., Gädeke, A., ... & Wada, Y. (2021). Global terrestrial water storage and drought severity under climate change. Nature Climate Change, 11(3), 226-233.," 2021.

[2] Rodell et al., "Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M. H. (2018). Emerging trends in global freshwater availability. Nature, 557(7707), 651-659.," 2018.

[3] Tapley et al., "Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., ... & Velicogna, I. (2019). Contributions of GRACE to understanding climate change. Nature climate change, 9(5), 358-369.," 2019.

[4] Pokhrel et al., "Pokhrel, Y. N., Hanasaki, N., Yeh, P. J., Yamada, T. J., Kanae, S., & Oki, T. (2012). Model estimates of sea-level change due to anthropogenic impacts on terrestrial water storage. Nature Geoscience, 5(6), 389-392.," 2012.

[5] Rodell et al., "Rodell, M., Barnoud, A., Robertson, F. R., Allan, R. P., Bellas-Manley, A., Bosilovich, M. G., ... & Seneviratne, S. I. (2024). An abrupt decline in global terrestrial water storage and its relationship with sea level change. Surveys in Geophysics, 45(6),," 2024.

[6] Ali et al., "Ali, M., Gillani, S. H., & Ali, U. (2024). Flash Drought Monitoring in Pakistan Using Machine Learning Techniques and Multivariate Drought Indices. Technical Journal, 3(ICACEE), 717-729.," 2024.

[7] Wu et al., "Wu, G., Chen, J., Shi, X., Kim, J. S., Xia, J., & Zhang, L. (2022). Impacts of global climate warming on meteorological and hydrological droughts and their propagations. Earth's Future, 10(3), e2021EF002542.," 2022.

[8] Zhang et al., "Zhang, Y., Hao, Z., Feng, S., Zhang, X., Xu, Y., & Hao, F. (2021). Agricultural drought prediction in China based on drought propagation and large-scale drivers. Agricultural Water Management, 255, 107028.," 2021.

[9] Yang et al., "Yang, D., Yang, Y., & Xia, J. (2021). Hydrological cycle and water resources in a changing world: A review. Geography and Sustainability, 2(2), 115-122.," 2021.

[10] Akhtar et al., "Akhtar, T., Mushtaq, H., & Hashmi, M. Z. U. R. (2020). Drought monitoring and prediction in climate vulnerable Pakistan: Integrating hydrologic and meteorologic perspectives. Hydrology and Earth System Sciences Discussions, 2020, 1-29.," 2020.

[11] Ashraf et al., "Ashraf, S., Ali, M., Shrestha, S., Hafeez, M.A., Moiz, A. and Sheikh, Z.A., (2022). Impacts of climate and land-use change on groundwater recharge in the semi-arid lower Ravi River basin, Pakistan. Groundwater for Sustainable Development, 17, p.100743," 2022.

[12] Amjad., "Amjad, M. (2023). Hydrometeorological Extremes and Climate Profiling of Pakistan. In Water and Environment for Sustainability: Case Studies from Developing Countries (pp. 1-16). Cham: Springer International Publishing.," 2023.

[13] Rehman et al., "Rahman, M. A., Ali, M., Mojid, M. A., Anjum, N., Haq, M. E., Kainose, A., & Dissanayaka, K. D. C. R. (2023). Crop coefficient, reference crop evapotranspiration and water demand of dry-season Boro rice as affected by climate variability: A case study from," 2023.

[14] Otto et al., "Otto, F. E., Zachariah, M., Saeed, F., Siddiqi, A., Kamil, S., Mushtaq, H., ... & Clarke, B. (2023). Climate change increased extreme monsoon rainfall, flooding highly vulnerable communities in Pakistan. Environmental Research: Climate, 2(2), 025001.," 2023.

[15] Zhao et al., "Zhao, M., Velicogna, I., & Kimball, J. S. (2017). Satellite observations of regional drought severity in the continental United States using GRACE-based terrestrial water storage changes. Journal of Climate, 30(16), 6297-6308.," 2017.

[16] Janjua et al., "Janjua, S., Hassan, I., Muhammad, S., Ahmed, S., & Ahmed, A. (2021). Water management in Pakistan's Indus Basin: challenges and opportunities. Water Policy, 23(6), 1329-1343.," 2021.

[17] Landerer et al., "Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. Water resources research, 48(4).," 2012.

[18] Ali et al., "Ali, S., Liu, D., Fu, Q., Cheema, M. J. M., Pham, Q. B., Rahaman, M. M., ... & Anh, D. T. (2021). Improving the resolution of GRACE data for spatio-temporal groundwater storage assessment. Remote Sensing, 13(17), 3513.," 2021.

[19] Ali et al., "Ali, S., Khorrami, B., Jehanzaib, M., Tariq, A., Ajmal, M., Arshad, A., ... & Khan, S. N. (2023). Spatial downscaling of GRACE data based on XGBoost model for improved understanding of hydrological droughts in the Indus Basin Irrigation System (IBIS). Rem," 2023.

[20] Chen et al., "Chen, L., He, Q., Liu, K., Li, J., & Jing, C. (2019). Downscaling of GRACE-derived groundwater storage based on the random forest model. Remote Sensing, 11(24), 2979.," 2019.

[21] You et al., "You, Y., Ting, M., & Biasutti, M. (2024). Climate warming contributes to the record-shattering 2022 Pakistan rainfall. npj Climate and Atmospheric Science, 7(1), 89.," 2022.

[22] Latif et al., "Latif, M., Shireen, H., Adnan, S., Ahmed, R., & Hannachi, A. (2024). Drought variability in Pakistan: Navigating historical patterns in a changing climate with global teleconnections. Theoretical and Applied Climatology, 155(8), 8379-8400.," 2024.

[23] Masood et al., "Masood, M. U., Haider, S., Rashid, M., Lodhi, M. U. K., Pande, C. B., Alshehri, F., ... & Sammen, S. S. (2023). The Effect of the El Nino Southern Oscillation on Precipitation Extremes in the Hindu Kush Mountains Range. Water, 15(24), 4311.," 2023.

[24] Ali et al., "Ali, M., Ali, U., & Gillani, S. H. FLASH DROUGHT DETECTION AND RISK EXPOSURE IN PAKISTAN: A MACHINE LEARNING APPROACH WITH MULTIVARIATE INDICES.," 2024.

[25] Modi et al., "29. Modi, P., Hanasaki, N., Yamazaki, D. et al. 2022 Sensitivity of subregional distribution of socioeconomic conditions to the global assessment of water scarcity. Commun Earth Environ 3, 144. https://doi.org/10.1038/s43247-022-00475-w," 2022.

[26] Munir et al., "Munir, M. U., Ahmad, A., Hopmans, J. W., Belgacem, A. O., & Baig, M. B. (2021). Water scarcity threats to National Food Security of Pakistan—Issues, implications, and way forward. Emerging challenges to food production and security in asia, Middle East, a," 2021.