

A SATELLITE-BASED APPROACH FOR QUANTIFYING TERRESTRIAL WATER
CYCLE INTENSITY AND ITS RELATIONSHIP WITH GROUNDWATER
AVAILABILITY IN DATA-SCARCE ARID REGIONS

by

FABIAN JACHI ZOWAM

(Under the Direction of Adam M. Milewski)

ABSTRACT

Climate change is speeding up parts of the global water cycle, and the water cycle intensity (WCI) is a tool to quantify this acceleration. Given that groundwater is a component of the global water cycle, this dissertation aims to understand how WCI changes (attributable to climate change) affect groundwater availability, specifically for arid regions. A previous effort to quantify the WCI over any landscape made use of ground-based datasets and focused on historical trends (1945 – 2014). This dissertation: (1) validated a remote sensing approach for quantifying the WCI over the contiguous United States (CONUS) for a more recent period (2001 – 2019) – to capture current climate change trends. Next, it (2) employed data analytics to predict groundwater level anomalies (GWLAs) across an arid region within the CONUS determined by the results of the WCI analysis, using observations from existing groundwater monitoring wells and remotely sensed predictor variables, such as, precipitation, soil moisture, evapotranspiration, and vegetation cover. Finally, it (3) explored the dynamic relationship between the results of

the first two objectives for the arid region of interest, based on an innovative approach to statistical correlation and causation analysis.

The water cycle is speeding up over about half of the CONUS particularly the west, and the state of Arizona might be experiencing much higher WCI rates on average compared to other arid regions of the CONUS. A multi-model approach to predict monthly GWLAs across multiple aquifers in Arizona between January 2010 and December 2019 demonstrated satisfactory performance, and the predictive accuracy was much higher for the unconsolidated sand and gravel aquifers. Finally, a moderate to strong negative lead-lag relationship between groundwater and WCI anomalies (GWLAs leading WCI anomalies) was revealed at various sites across the study area. Some of these locations were contained within Active Management Areas (AMAs) – areas characterized by high groundwater reliance and the enforcement of the strictest groundwater regulations.

This study underscores the importance of groundwater monitoring and strategic management in vulnerable areas. The exclusive use of remotely sensed variables ensures that data scarce and vulnerable regions are well represented and the study's objectives can be replicability globally.

INDEX WORDS: climate change, groundwater, CONUS, Arizona

A SATELLITE-BASED APPROACH FOR QUANTIFYING TERRESTRIAL WATER
CYCLE INTENSITY AND ITS RELATIONSHIP WITH GROUNDWATER
AVAILABILITY IN DATA-SCARCE ARID REGIONS

by

FABIAN JACHI ZOWAM

BS, Obafemi Awolowo University, Nigeria, 2015

MS, Georgia State University, 2020

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2024

© 2024

Fabian Jachi Zowam

All Rights Reserved

A SATELLITE-BASED APPROACH FOR QUANTIFYING TERRESTRIAL WATER
CYCLE INTENSITY AND ITS RELATIONSHIP WITH GROUNDWATER
AVAILABILITY IN DATA-SCARCE ARID REGIONS

by

FABIAN JACHI ZOWAM

Major Professor:	Adam M. Milewski
Committee:	Marshall J. Shepherd
	Valentine A. Nzengung

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2024

DEDICATION

I dedicate this work to my parents for their immense support and prayers. My father, Gabriel, came to the U.S about forty-eight years ago for undergraduate and graduate studies and I grew up with stories of his experience. It was a privilege to walk the path he once walked. He and my mother, Florence, are my biggest motivation to succeed.

ACKNOWLEDGEMENTS

It was a long, challenging, and rewarding journey and I want to very specially thank my advisor, Dr. Milewski, and committee members (Drs. Shepherd and Nzungu) for making it a memorable experience in many ways. I am also grateful to the water resources and remote sensing (WRRS) lab group for the memories together and collaborations through the years. I wish to specially acknowledge the department of geology and the University of Georgia, for the enabling environment and resources that made this work possible. To my biggest support system, my family, I cannot thank you enough. Most importantly, I want to thank God for making this dream a reality.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
 CHAPTER	
1 INTRODUCTION.....	1
References.....	4
2 A SATELLITE-BASED APPROACH FOR QUANTIFYING TERRESTRIAL WATER CYCLE INTENSITY	6
Abstract.....	7
Introduction	8
Materials and Methods	13
Results	17
Discussion	26
Conclusion	28
References.....	30
3 GROUNDWATER LEVEL PREDICTION USING MACHINE LEARNING AND GEOSPATIAL INTERPOLATION MODELS.....	42
Abstract.....	43
Introduction	44

	Materials and Methods	46
	Results	62
	Discussion	72
	Conclusion	75
	References.....	79
4	CLIMATE VARIABILITY AND GROUNDWATER LEVELS: A CORRELATION AND CAUSATION ANALYSIS	90
	Abstract.....	91
	Introduction	92
	Materials and Methods	95
	Results	100
	Discussion	111
	Conclusion	114
	References.....	114
5	CONCLUSION.....	123
	References.....	124

LIST OF TABLES

	Page
Table 2-1: Results of the t-test for the first run (first 17 points for 17 different climate classes).....	18
Table 3-1: Final input variables used in the study and the processing involved	54
Table 3-2: Performance of Individual SVR Models	63
Table 3-3: RF model performance evaluated at the four test wells (15 percent test split)	64
Table 3-4: RF model performance at validation wells using the EBK predictions as input	67
Table 3-5: RF model performance using all monitoring wells for EBK predictions	70
Table 4-1: Summary of variables used in the study	97
Table 4-2: Summary of correlation analysis at all monitoring well locations	100
Table 4-3: Results of the GC tests conducted at the wells with the largest lagged correlation coefficients	105
Table 4-4: Summary of ccf analysis between GWLA and WCIA at the ungauged sites	109

LIST OF FIGURES

	Page
Figure 2-1: Flow chart showing the conceptual framework, data acquisition, and processing steps towards the main objectives of the study	13
Figure 2-2: Present day (1980–2016) Köppen–Geiger climate classification map showing 17 classes across the CONUS and the 17 randomly selected sample points for the first sample run	15
Figure 2-3: Comparison of GPM and PRISM precipitation datasets based on R^2 , NSE, and KGE.....	19
Figure 2-4: The Dfc climate region where GPM and PRISM showed the weakest agreement.....	20
Figure 2-5: Ranking of the 17 climate classes based on relative agreements between the GPM and PRISM gridded precipitation products, according to R^2 , NSE, and KGE values	21
Figure 2-6: Average annual (a) water cycle intensity, (b) precipitation, and (c) evapotranspiration from 2001 to 2019	22
Figure 2-7: El Niño and La Niña years over the period of study	24
Figure 2-8: The difference in (a) water cycle intensity, (b) precipitation, and (c) evapotranspiration between the annual averages of 2001–2009 and 2010–2019, and their respective differences without ENSO years (d–f).....	25

Figure 3-1: Map of the study area showing the locations of groundwater monitoring wells	48
Figure 3-2: Maps of initial predictor variables resampled to $0.125^{\circ} \times 0.125^{\circ}$ grid resolution	49
Figure 3-3: One-dimensional SVR, where x is the input or independent variable and y is the dependent variable	56
Figure 3-4: A simple RF Model with three DTs.....	58
Figure 3-5: (a) A single semivariogram. Source: Maliva (2016) [71]. (b) The EBK model showing a distribution of semivariograms	59
Figure 3-6: Model design showing train/test split ratios for the spatial (a) and temporal (b) evaluations and the locations of the split wells	60
Figure 3-7: RF schematic with three observations, the seven predictor variables, three randomly selected predictors to build each tree, and three DTs.....	62
Figure 3-8: Map showing the locations of the four test wells for temporal evaluations after the second (85:15) train/test split.....	64
Figure 3-9: Plots comparing the RF-predicted GWLA at each test well with observed values. The numbers 1–4 correspond to the four test wells shown in Figure 8 and Table 3	65
Figure 3-10: Location of the validation wells (1–3). The validation wells were removed from the dataset before performing the EBK to ensure unbiased model evaluation.....	66
Figure 3-11: Partial dependence plots (PDPs) of the validation RF model.....	67

Figure 3-12: EBK model output for January 2010 showing (a) GWL predictions at $0.023^\circ \times 0.023^\circ$, (b) standard errors of those predictions at $0.023^\circ \times 0.023^\circ$, and (c) resampled GWL prediction surface at a $0.125^\circ \times 0.125^\circ$ grid size	68
Figure 3-13: Spatial locations of all the monitoring wells used in this study. Wells 1–7 were used to evaluate the performance of the final RF model.....	69
Figure 3-14: GWLA predictions for January 2010 after model deployment	71
Figure 3-15: Average GWLA for the period of study (January 2010 to December 2019), and the locations of the monitoring wells (black circles) used in this study	72
Figure 4-1: Study Area showing groundwater monitoring wells and surficial geology	95
Figure 4-2: Methodological flow chart	99
Figure 4-3: Location of the GC test wells	103
Figure 4-4: Ccf plot for Well 43.....	104
Figure 4-5: AMAs and correlation wells	106
Figure 4-6: Scatter plots for Well 33 illustrate contemporaneous (lag = 0) and lagged (lag = +1) relationships between WCI and GWL anomalies	107
Figure 4-7: Locations where ccf analysis between GWLA and WCIA were conducted	108

CHAPTER 1

INTRODUCTION

All of the waters within the earth-atmosphere system are continuously cycled through a complex process known as the water cycle. As a cycle, there are no start or end points but rather a series of processes or phases that are a part of a continuous system. These phases include precipitation, evaporation, transpiration, surface runoff, and infiltration. Climate change is known to impact this system (Trenberth, 2014; Malinowski and Skoczko, 2018). But what exactly is ‘climate change’, and why is it important?

According to the Intergovernmental Panel on Climate Change (IPCC), climate change is simply the long-term (decades or longer) variation in average weather conditions of a place (Kumar, 2012), brought about by anthropogenic activities and emissions that increase the concentrations of greenhouses in the atmosphere (Kumar, 2012, Trenberth, 2011; Trenberth, 2018). These gases are methane, carbon dioxide, and nitrous oxide (Trenberth, 2011), and they trap outgoing radiation, resulting in the accumulation of heat and warming of the planet (Trenberth, 2011; Trenberth, 2018). Between 1970 and 2004, the global anthropogenic greenhouse gas emissions increased by 70 percent (IPCC., 2007), and between 1990 and 2019, the warming effects associated with these gasses increased by 45 percent (36 percent for carbon dioxide alone) (Indicators, 2015). This ultimately leads to a change in climate (MacCracken, 2001).

This human-induced change in climate is changing precipitation patterns and the water cycle (Trenberth, 2011). Hydrologic extremes are becoming more extreme

(Kundzewicz, 2008; Malinowski and Skoczko, 2018) and the difference in precipitation patterns between wet and dry areas is intensifying (Dao et al., 2023). In particular, these changes are likely to impact the natural recharge of groundwater in many parts of the world. For instance, for most of the United States, a warming of 1.5 °C could lead to a loss of over 100 billion cubic meters of groundwater storage within four years (Condon et al., 2020). This can be especially troubling for arid and semi-arid regions where surface water availability is already limited.

Therefore, this dissertation addresses the impact of climate change on groundwater availability in arid regions through a systematic investigation comprising three distinct stages or phases:

The first phase (chapter 2) aims to quantify the terrestrial water cycle response to climate change for any point on the globe, using remote sensing. The terrestrial water cycle is the continuous movement and re-distribution of water between the earth's land surface and atmosphere, and groundwater is a component of that cycle. Huntington et al. (2018) proposed a framework for quantifying the terrestrial water cycle intensity (WCI) across the contiguous United States (CONUS), but their approach is primarily beneficial to data-rich regions like the CONUS. As a result, this phase will include the validation of a remote sensing WCI approach (for the CONUS) that will be more beneficial for data-scarce regions. Regions where the water cycle is intensifying will show a positive change in WCI ($\Delta WCI > 0$) and a weakening cycle will show a negative change ($\Delta WCI < 0$). (Huntington et al., 2018).

Predicting groundwater in arid regions can be challenging due to several factors including the limited number of monitoring wells, insufficient and incomplete data, and

inaccurate estimates of aquifer recharge and discharge (Tladi et al., 2023). The second phase of the investigation (chapter 3) explores ways to overcome this challenge, using a combination of machine learning and geospatial interpolation models, and remotely sensed input variables to predict groundwater levels at a regional scale much smaller than the CONUS. Machine learning is a subset of artificial intelligence that learns patterns in datasets in order to make predictions, and its application can be seen in our everyday lives. For example, movie streaming services (such as Netflix) recommend new movies based on the movies we've seen previously, the time of the day we saw those movies, and so on. Online marketplaces (such as Amazon) recommend products to us based on items we've previously purchased or tried to purchase. Similarly, the focus here is to learn relationships between groundwater levels observed at monitoring wells and various remotely sensed explanatory variables such precipitation, soil moisture, land surface temperature, and so on, to predict groundwater levels at locations without monitoring wells. The study area selection for this phase will be based on the findings and interpretations from phase 1.

Changes in terrestrial water cycle intensities from phase 1 serve as indicators of climate change, and groundwater levels from phase 2 are our measure of groundwater availability. Phase 3 (chapter 4) assesses the dynamic relationships between the research products of Phase 1 and 2 to evaluate dependencies and interactions between both variables at local scales within the region of interest.

All of this would be done using remotely sensed datasets from satellites with global coverage, meaning that these can be replicated at any point on the globe. Still, the biggest beneficiaries of this remote sensing approach are the data sparse regions of the world where field based hydrological variables are limited or largely inaccessible. This study promises

to be a valuable resource for monitoring groundwater resources in the face of global warming and climate change threats, and it is my hope that it contributes substantially to the existing body of knowledge towards planning for an unpredictable climate future.

References

- Condon, L. E., Atchley, A. L., & Maxwell, R. M. (2020). Evapotranspiration depletes groundwater under warming over the contiguous United States. *Nature communications*, 11(1), 873.
- Dao, P. U., Heuzard, A. G., Le, T. X. H., Zhao, J., Yin, R., Shang, C., & Fan, C. (2023). The impacts of climate change on groundwater quality: A review. *Science of the Total Environment*, 169241.
- Indicators, C. C. (2015). Climate Change Indicators: Greenhouse Gases. Retrieved May, 30, 2020.
- IPCC (Intergovernmental Panel on Climate Change). (2007). Climate change 2007: synthesis report. Summary for policymakers. *Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Kumar, C. P. (2012). Climate change and its impact on groundwater resources. *International Journal of Engineering and Science*, 1(5), 43-60.
- Kundzewicz, Z. W. (2008). Climate change impacts on the hydrological cycle. *Ecohydrology & Hydrobiology*, 8(2-4), 195-203.
- MacCracken, M. C. (2001). Global warming: A science overview. In *Global Warming and Energy Policy* (pp. 151-159). Boston, MA: Springer US.

Malinowski, Ł., & Skoczko, I. (2018). Impacts of climate change on hydrological regime and water resources management of the Narew River in Poland. *Journal of Ecological Engineering*, 19(4).

Tladi, T. M., Ndambuki, J. M., Olwal, T. O., & Rwanga, S. S. (2023). Groundwater Level Trend Analysis and Prediction in the Upper Crocodile Sub-Basin, South Africa. *Water*, 15(17), 3025.

Trenberth, K. E. (2011). Changes in precipitation with climate change. *Climate research*, 47(1-2), 123-138.

Trenberth, K. E. (2014). Water cycles and climate change. *Global environmental change*, 2014, 31-37.

Trenberth, K. E. (2018). Climate change caused by human activities is happening and it already has major consequences. *Journal of energy & natural resources law*, 36(4), 463-481.

CHAPTER 2

A SATELLITE-BASED APPROACH FOR QUANTIFYING TERRESTRIAL WATER
CYCLE INTENSITY ¹

¹ Zowam, F.J., Milewski, A.M., and Richards IV, D.F. 2023. *Remote Sens.* 15(14), 3632.

Reprinted here with permission of the publisher.

Abstract

The terrestrial water cycle intensity (WCI) is a widely used tool to quantify the impact of climate change on the distribution of global water resources. In this study, a satellite-based WCI was tested by comparing the parameter-elevation regressions on independent slopes model (PRISM) precipitation estimates with those of the Global Precipitation Measurement (GPM) satellite mission across the contiguous United States (CONUS), based on an existing Köppen–Geiger climate classification for the CONUS. Both precipitation products were not statistically different across all climate classes. Consequently, satellite-based WCI changes between two multiannual periods (2001 to 2009 and 2010 to 2019) were calculated at a 0.1-degree spatial resolution using the GPM and a validated global evapotranspiration dataset. This study showed that: (1) The water cycle is speeding up in many parts of the CONUS, particularly the West, driven by recent increases in both precipitation and evapotranspiration through much of the region. (2) The El Niño-Southern Oscillation (ENSO) may be influencing the WCI of the CONUS by driving precipitation in the west, southeast, and parts of the north, and dryness in the northeast regions. The hydrological impacts of these results cannot be generalized. However, flood and drought risks, water availability and quality issues remain key primary concerns.

Introduction

The terrestrial water cycle is the continuous circulation and redistribution of water between the earth's land surface and atmosphere. The intensity of this circulation over any spatial-temporal scale is expressed as the sum of precipitation and evapotranspiration [1]. Globally, precipitation and evapotranspiration rates are increasing as mean air temperatures continue to rise due to global warming [2]. The period from 2012 to 2021 was the warmest decade recorded since 1901, where the warmest years were 2016 and 2020 [3]. In the United States, the last four decades, prior to 2021, have seen faster warming than the global average [3]. It is expected that these increased temperatures will strengthen the atmosphere's evaporative demand, causing an increase in evapotranspiration rates due to water availability. This was evident during two fifteen-year periods spanning from 1984 to 2015, where global evapotranspiration increased by at least 3.57 cubic miles/year as the areas covered by water also increased [4]. Increases in global evapotranspiration rates will increase the likelihood and frequencies of droughts in many parts of the world just as flood risks increase with increasing precipitation intensities. While floods and droughts are both undesirable and extreme hydrological events, droughts can particularly disrupt the natural recharge of both surface and groundwater systems and potentially threaten global water security. Thus, the terrestrial water cycle intensity (WCI) is a valuable tool for monitoring and assessing changes in specific components of the water cycle in response to a changing climate.

By using climate variables from the parameter-elevation regressions on independent slopes model (PRISM) to calculate WCI changes for the contiguous United

States (CONUS) between the averages of 1945 to 1974 and 1985 to 2014 [1], two possible considerations became apparent. (1) The temporal spread of their study did not exclusively address current climate change anomalies. In the CONUS, the last two decades prior to 2021 have shown most of the warmest years in a 120-year record [3]. (2) PRISM uses point measurements of climatic observations from a wide range of monitoring networks to develop spatially continuous climate datasets. Such methodology will be less effective in data-sparse regions of the world. In fact, most of the earth's surface lacks adequate in situ precipitation measuring stations [5–7]. In developing countries, the situation is not improving [8–10]. Cases of inadequacy have been well reported over the continents of Africa [10–15] and South America [12,16–19]. To put this into perspective, the average rain gauge density in the United States is 1.3 gauges per 1000 sq. km [20], but Nigeria, one of the largest developing African countries, has only 87 operational stations and is 970 short of achieving a density of 1 gauge per 874 sq. km [21]. In many regions of Africa and South America, the rain gauge density is as low as 1 gauge per 100,000 sq. km [22]. Of course, understanding WCI changes in these regions where hydrometeorological data may be limited is as useful as it is over data-rich areas. Therefore, addressing both issues is to evaluate the feasibility of using satellite-sourced variables to understand the implications of recent temperature anomalies for the WCI of the CONUS.

The tropical rainfall measuring mission (TRMM) was the first satellite devoted to measuring precipitation from space [23]. It was launched in 1997 and estimated tropical rainfall for about 17 years [24–26]. Since the evolution of TRMM and its successor, the GPM mission, several researchers have evaluated the ability of various satellite-derived precipitation products to capture the patterns and intensity of rainfall. In the Middle East,

between 1998 and 2013, three of these products namely: the TRMM Multi-Satellite Precipitation Analysis 3B42 (TRMM-3B42) product, the Climate Prediction Center MORPHing technique (CMORPH), and the precipitation estimation from remotely sensed information using artificial neural networks (PERSIANN), failed to effectively replicate severe daily rainfall events [27]. Similar results were reported for a mountainous region in eastern Italy between 2003 and 2010 [28]. In Africa, the TRMM-3B42 was able to detect spatial and seasonal rainfall patterns, and reasonably estimate high-intensity events over the Blue Nile basin in Sudan between 2001 and 2016 [29]. Over the Three Gorges Reservoir area of China, between 2001 and 2016, all three products showed varying strengths in terms of rainfall amount, extreme precipitation, and rainy-day detection ability [30]. In most of these cases, TRMM proved to be a reliable source for continuous measurements in space and time at the monthly and annual scales. Its successor, the GPM mission, made use of the most advanced instruments in space [31] and improved spatial resolution (0.25° to 0.1°), revisit times (3 h to 30 min), and latitudinal coverage (quasi-global) compared to TRMM [26]. Despite these advancements, the integrated multi-satellite retrievals for the GPM (IMERG) algorithm, which estimates precipitation from GPM constellation retrievals, incorporates early TRMM estimates in its latest version to produce a consistent, long-term precipitation record [31,32]. Otherwise, systematic biases exist between both missions, underscoring the need to make appropriate considerations when utilizing their respective products. In east-central China, the final IMERG product outperformed the early IMERG, late IMERG, near-real-time, and post-processing TRMM products, providing the most accurate estimation of daily and monthly rainfall [33]. Over Singapore, the GPM IMERG performed better than two TRMM products (3B42 and

3B42RT) in detecting precipitation, capturing variabilities, and providing more accurate daily estimates [24]. Over a mountainous region in southwest China, the GPM IMERG and 3B42 products were evaluated against observed rain gauge data from a dense gauge network, where the IMERG product showed improved capabilities to capture rainfall variability and detect medium and high-intensity events but tended to overestimate the high-intensity events [34]. In the CONUS, the GPM IMERG performed better than the 3B42 product in capturing precipitation intensity variations, reducing missed-precipitation bias for winter and summer precipitation, reducing false-precipitation bias for summer precipitation, and showing better consistency in capturing spatial distribution patterns at monthly time scales, based on comparisons with ground-based, gauge-corrected radar systems [35]. Globally, IMERG enhances precipitation detection capacity, outperforms other satellite products in its ability to capture spatiotemporal variability of extreme events, and is one of the best alternatives to ground-based measurements [36].

On the other hand, satellite remote sensing of ET generally has lesser efficiencies than precipitation, with significant disparities among various observatory satellites. However, in a recent study, Elnashar et al. (2021) [37] ranked the performance of 12 global (satellite) ET datasets after validation against flux eddy covariance ET from 645 sites based on six metrics. The authors synthesized the best-performing products into a single, much-improved ensemble global ET dataset across all land cover types, from 1982 to 2019. Their methodological approach provided some insight for this study.

The main objectives of this study are twofold. (1) To develop a framework to test the ability of a remote sensing WCI by first evaluating the reliability of satellite products to estimate precipitation over the CONUS. Monthly PRISM and GPM IMERG

precipitation estimates between January 2001 and December 2019 will be compared for different climates based on three metrics and a statistical test on their difference. (2) Upon successful completion of the first objective, calculate WCI changes between two separate multiannual periods (2001–2009 and 2010–2019), using the GPM precipitation product and the Elnashar et al. (2021) validated ET datasets (Figure 1).

Shifts in the water cycle are among the most consequential effects of climate change [38]. There are also naturally occurring climate variability signals that potentially modify regional climates and alter WCI patterns. The quasi-periodic fluctuation in equatorial Pacific Ocean temperatures, otherwise known as the El Niño—Southern Oscillation (ENSO) is the most prominent year-to-year climate variation affecting underlying weather and climate patterns on Earth [39]. Sea surface temperatures (SSTs) act as natural indicators of these events, where above-average temperatures may indicate a warm (El Niño) phase and below-average temperatures, a cold (La Niña) phase. Various impacts of El Niño and La Niña on the climatology of the CONUS have been well reported—some of which include: the correlation between summer rainfall in the central CONUS from 1950 to 1990 and El Niño events [40], correlations between the 1988/1993 North America summer droughts/floods and La Niña/El Niño signals in those respective years [41], the influence of El Niño on winter precipitation in California from 1901 to 2010 [42], and the attribution of the 2012 drought in the south and south-central CONUS to the recurring 2010-11 and 2011-12 La Niña episodes [43]. Therefore, in addition to calculating satellite-based WCIs, this study also aims to quantify the influence of these events on WCI changes across the CONUS for the selected period of study.

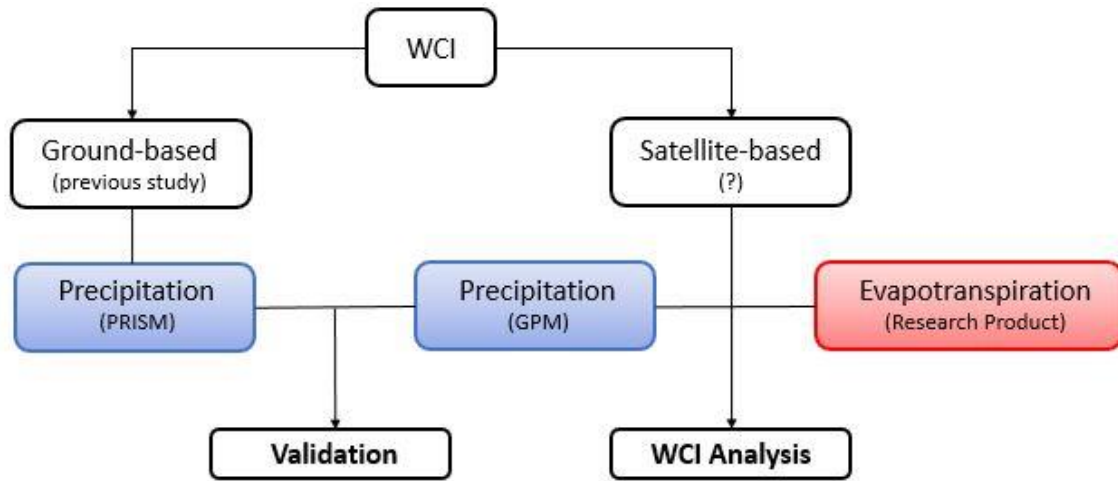


Figure 1. Flow chart showing the conceptual framework, data acquisition, and processing steps towards the main objectives of the study. The blue boxes represent the precipitation datasets, and the red box represents the evapotranspiration dataset used in the study.

Materials and Methods

The final run of the IMERG fusion, the most suitable version for research purposes [44], was used in this study. It was accessed and downloaded from the NASA data portal (<https://gpm.nasa.gov/data/directory>, accessed on 11 May 2021). A gridded global precipitation dataset with 0.1-degree spatial (~11 km) and monthly temporal scale was downloaded in Tiff format for January 2001 to December 2019.

A gridded PRISM precipitation raster dataset for the CONUS with 4 km (~0.04 degree) cell size and monthly temporal scale was also downloaded. This dataset, acquired in Bil formats from the Oregon State PRISM data portal (<https://prism.oregonstate.edu/>, accessed on 11 November 2022) for January 2001 to December 2019, was resampled to the spatial resolution of the GPM rasters.

The [37] validated global ET research product was accessed from the Harvard Dataverse research data repository (<https://doi.org/10.7910/DVN/ZGOUED>, accessed on 3 July 2021). A gridded monthly ET dataset with 0.01-degree (~1 km) spatial resolution was downloaded in Tiff format for the same period covered in both precipitation datasets. The downloaded rasters were upscaled to the spatial resolution of the precipitation datasets.

The gridded GPM and PRISM products were compared for various climates based on the Beck et al. (2018) [45] present-day Köppen–Geiger climate classification (Figure 2). The first letters in the classification scheme (A, B, C, D) represent the main climate types: tropical, arid, temperate, and cold [45]. The second (W, S, f, s, w, m) and third letters (h, k, a, b, c,) indicate moisture and air temperature characteristics, respectively [46]. The Csa climate class, for example, represents a temperate, mild climate with more precipitation in the winter than in summer where the summer months are hot, and is identical to the Csb except that it has hotter summers (Figure 2).

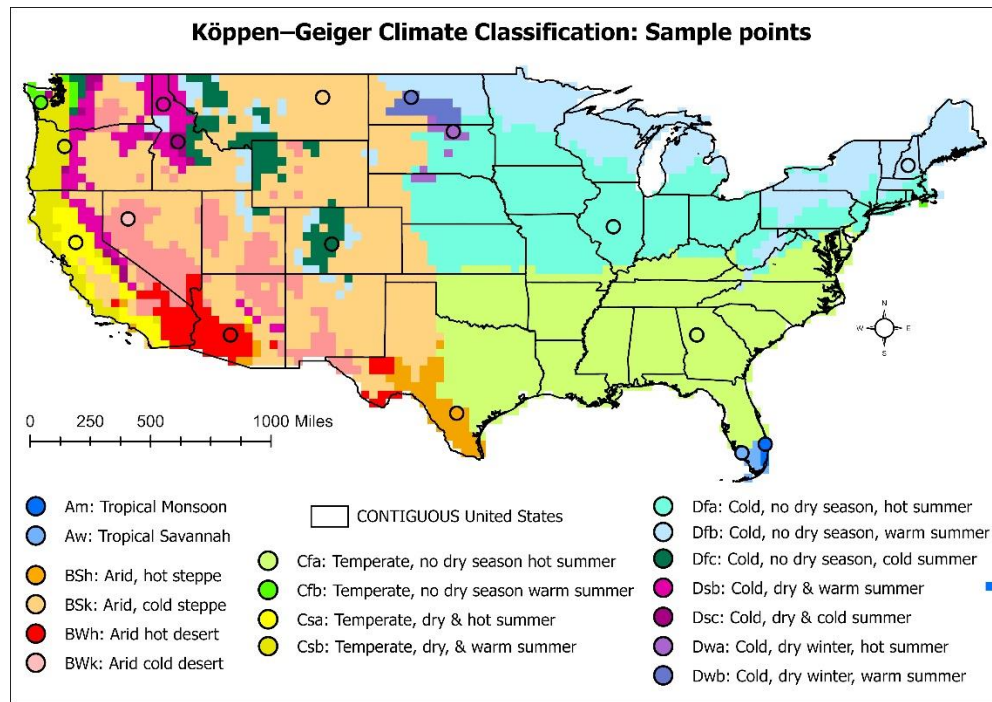


Figure 2. Present day (1980–2016) Köppen–Geiger climate classification map by [45] showing 17 classes across the CONUS and the 17 randomly selected sample points for the first sample run.

For each of the 17 different climate classes over the CONUS, 17 (0.1-degree) cells (each representing a climate class) were randomly selected (Figure 2). Monthly precipitation estimates were then extracted from both precipitation datasets, and the procedure was repeated twice to obtain a total of three sample runs and 51 (0.1-degree) cells—three cells representing each climate class.

A paired t-test of the difference in means between the GPM and PRISM precipitation estimates was performed on each set of 17 randomly selected points, where a null hypothesis of zero was tested at a significance level of 0.05 (5%). First, a differencing technique to eliminate spatial and time correlations was applied to the datasets. Differencing both datasets reduced the number of observations by one.

The datasets were further compared based on three model evaluation metrics: coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and Kling–Gupta efficiency (KGE). The NSE compared each observed (GPM) value with its corresponding simulated (PRISM) value, normalizing their difference by the variance of the observed time series. It ranges from negative infinity to one, with higher values indicating a well-varied observed time series that aligns closely with the simulated values. Like the NSE, the KGE ranges from negative infinity to one but with more straightforward relationships between correlation, variability, and bias [47].

Following successful evaluations at the 51 randomly selected points, monthly WCI was calculated for each 0.1-degree cell across the CONUS by summing up the GPM IMERG precipitation and corresponding evapotranspiration values from Elnashar et al. (2021). The total annual WCI for each year was obtained by adding these monthly WCI values, year by year, and an annual average was calculated by averaging the cell values over the period of study (2001 to 2019). In addition, the WCI difference between the annual averages of 2001 to 2009 and 2010 to 2019 was calculated to show temporal trends and spatial patterns of change between both multiannual periods.

El Niño (or La Niña) events in the Niño 3.4 region of the tropical Pacific Ocean have been defined differently in various studies, such as occurring when the December-January-February (DJF) SST anomaly exceeds $\pm 0.5^\circ\text{C}$ [48,49], or when five consecutive three-month running means of SST anomalies exceed $\pm 0.5^\circ\text{C}$ between July and June [50]. Although definitions may vary, such events generally occur when anomalies exceed $\pm 0.5^\circ\text{C}$ for several months [51]. In this study, a straightforward approach was used to analyze SST anomalies in the tropical Pacific Ocean. We simply calculated annual

averages of the three-month running means of SST anomalies obtained from [52]. The anomalies were originally derived by calculating SST departures from 30-year average baseline temperature conditions [52]. Applying the threshold of ± 0.5 °C, years with significant warm and cold phases were identified. These years were excluded from the dataset and WCI calculations were repeated. However, years that exceeded both El Niño and La Niña thresholds were considered neutral, indicating a lack of clear dominance of either El Niño or La Niña conditions, and therefore were not removed from the datasets.

Results

This section consists of two parts to address the study's main objectives: Validation and WCI analyses.

Validation

The t-tests yielded identical results across all three sample runs, although only the results from the first run are presented (Table 1). P-values were greater than our chosen significance level of 0.05, so we would accept the null hypothesis that both precipitation products are not statistically different. The confidence interval (CI), which represents a range of acceptable null hypotheses, also included a zero at every sample point. Thus, we could still not rule out a zero difference in means between both datasets and would, therefore, accept the null hypothesis that they are not statistically different [53]. Also, (Cohen's d) effect size calculations showed negligible differences between the GPM and PRISM precipitation datasets. Because of how the tests were set up, negative values imply that the mean of the GPM dataset was lesser than that of the PRISM dataset.

Table 1. Results of the t-test for the first run (first 17 points for 17 different climate classes). P-values and CIs were generated from the paired t-tests. The CI is a range of acceptable null hypotheses defined by lower and upper confidence limits [53]. Narrower CIs indicate smaller uncertainties in our estimates, and vice versa. The effect size represents the magnitude of the difference between GPM and PRISM datasets at each sample point.

ID	Class: Description	P-Value	CI	Effect Size
1	Am: Tropical Monsoon	0.98	−7.96–8.13	0.0006
2	Aw: Tropical Savannah	0.95	−6.84–7.24	0.0017
3	BWh: Arid, desert, hot	0.98	−2.01–1.96	−0.0006
4	BWk: Arid, desert, cold	0.90	−0.98–1.12	0.0030
5	BSh: Arid, steppe, hot	0.99	−5.35–5.27	−0.0004
6	BSk: Arid, steppe, cold	0.99	−3.14–3.09	−0.0004
7	Csa: Temperate, dry summer, hot summer	0.99	−8.44–8.58	0.0003
8	Csb: Temperate, dry summer, warm summer	0.93	−8.27–9.08	0.0022
9	Cfa: Temperate, no dry season, hot summer	0.99	−6.05–6.00	−0.0002
10	Cfb: Temperate, no dry season, warm summer	0.82	−7.80–9.86	0.0045
11	Dsb: Cold, dry summer, warm summer	0.98	−3.82–3.93	0.0007
12	Dsc: Cold, dry summer, cold summer	0.98	−3.40–3.30	−0.0010
13	Dwa: Cold, dry winter, hot summer	0.99	−3.50–3.45	−0.0004
14	Dwb: Cold, dry winter, warm summer	0.99	−2.79–2.83	0.0003
15	Dfa: Cold, no dry season, hot summer	1.00	−6.06–6.07	0.0000
16	Dfb: Cold, no dry season, warm summer	0.93	−5.28–5.76	0.0024
17	Dfc: Cold, no dry season, cold summer	0.89	−2.24–1.94	−0.0046

Further evaluations based on the three performance metrics showed that the Dfc (cold, no dry season, cold summer) climate class, having the lowest R^2 on all three sample runs, the lowest NSE for the second and third runs, and the lowest KGE value on the first run, demonstrated relatively weak agreement compared to other classes (Figure 3). Representing the climate of the Rockies (Figure 4), it is still unclear whether the weak agreements were due to terrain complexities, climatic factors, both, or other factors. Therefore, further investigation may be necessary to specifically determine the underlying causes of the relatively poor performance of the GPM satellite product over the Dfc climate.

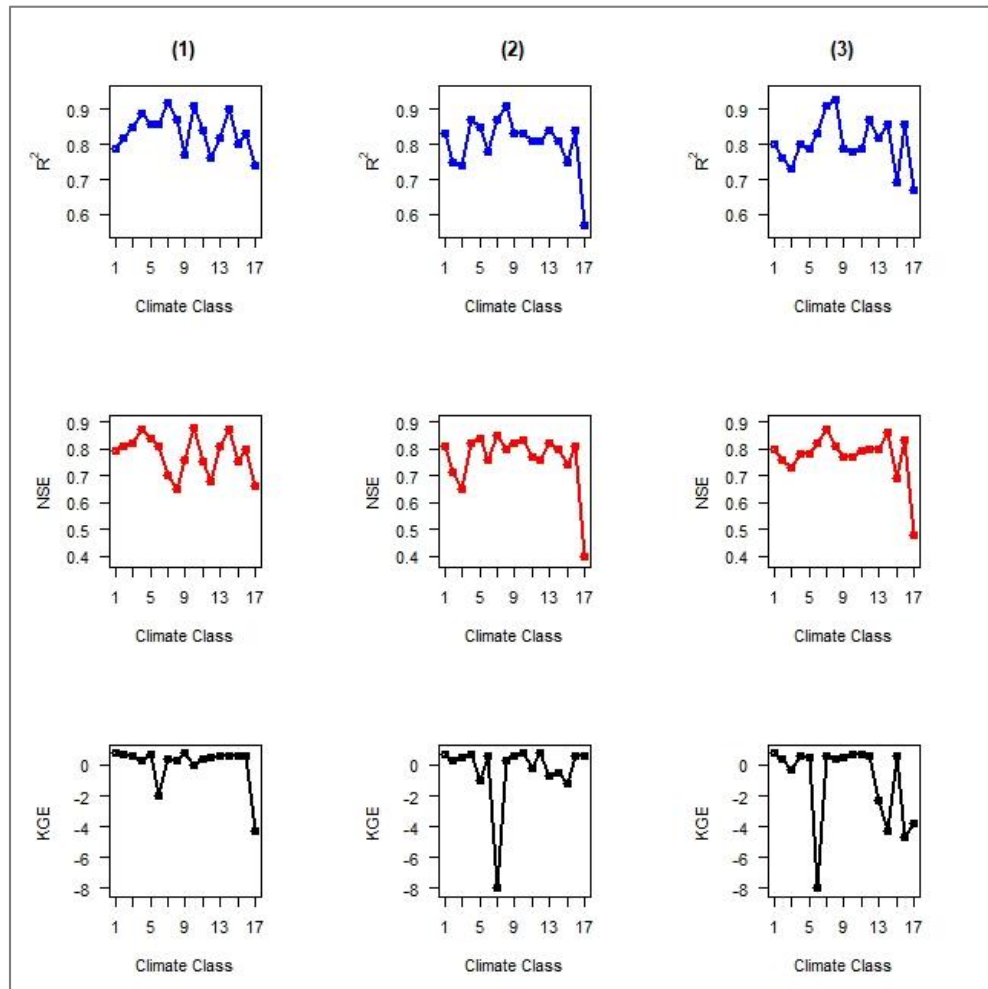


Figure 3. Comparison of GPM and PRISM precipitation datasets based on R^2 , NSE, and KGE. The numbers (1), (2), and (3) represent the first, second, and third sample runs. The Dfc climate is represented as climate class 17.

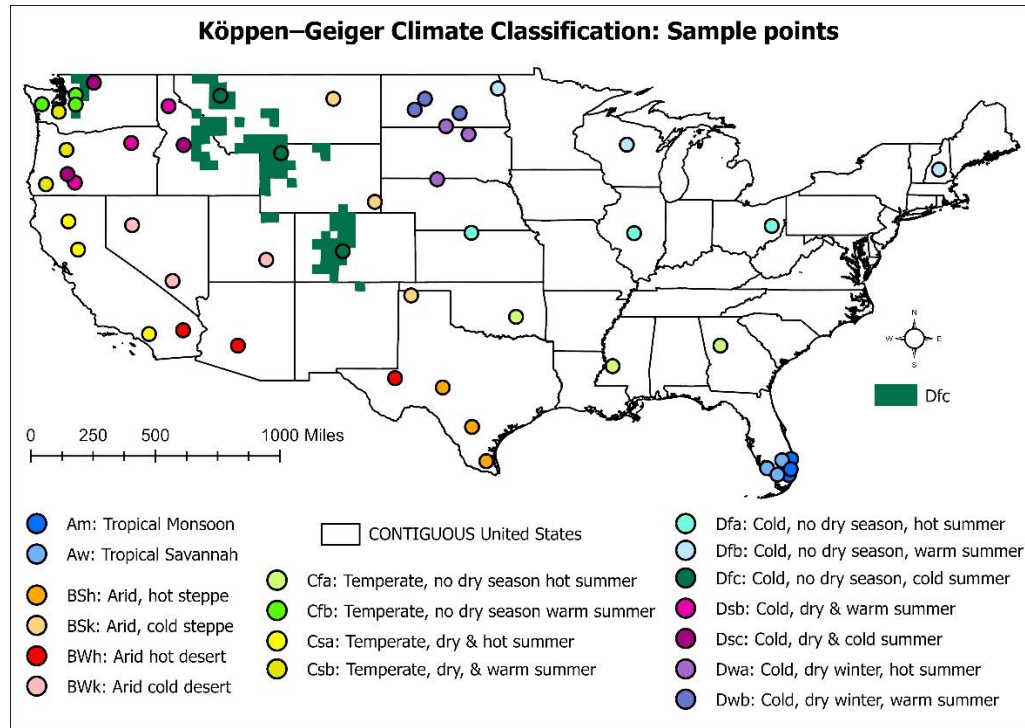


Figure 4. The Dfc climate region where GPM and PRISM showed the weakest agreement.

Based on recommended R^2 and NSE standards for monthly periods and local scales [54], agreement across all Köppen-Geiger climates was at least satisfactory on average. A ranking of climate classes by their relative agreements between both gridded precipitation products showed that while the Dfc was weakest, the Bwk (arid, desert, cold) class showed the strongest agreement (Figure 5).

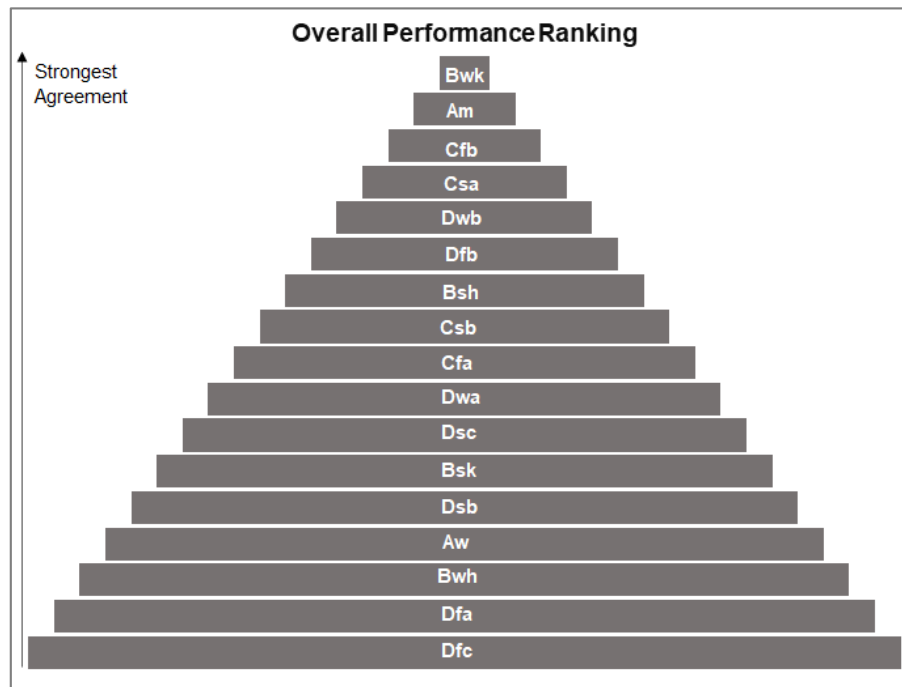


Figure 5. Ranking of the 17 climate classes based on relative agreements between the GPM and PRISM gridded precipitation products, according to R^2 , NSE, and KGE values.

WCI Analyses

The average annual WCI between 2001 and 2019 varied across the CONUS. The west generally showed lower intensities, whereas the highest values were seen in the southeastern CONUS region. However, in the westernmost border around the western Washington region, values of over 4000 mm/yr. were recorded (Figure 6a). Similar patterns were seen for both precipitation and evapotranspiration averages. Two essential ingredients for precipitation are moisture and lift. Thus, with the Gulf of Mexico as a potential supplier of moisture, combined with the existence of several mountain ranges in the region, it is no surprise that there is that much rain in the southeast region. Likewise, the active weather caused by the low-pressure system of the Aleutian Islands explains the excessive amounts of rain along the western boundary (Figure 6b). To illustrate the importance of available

moisture for evapotranspiration, the areas with high precipitation (Figure 6b) appeared to also show relatively high evapotranspiration rates (Figure 6c).

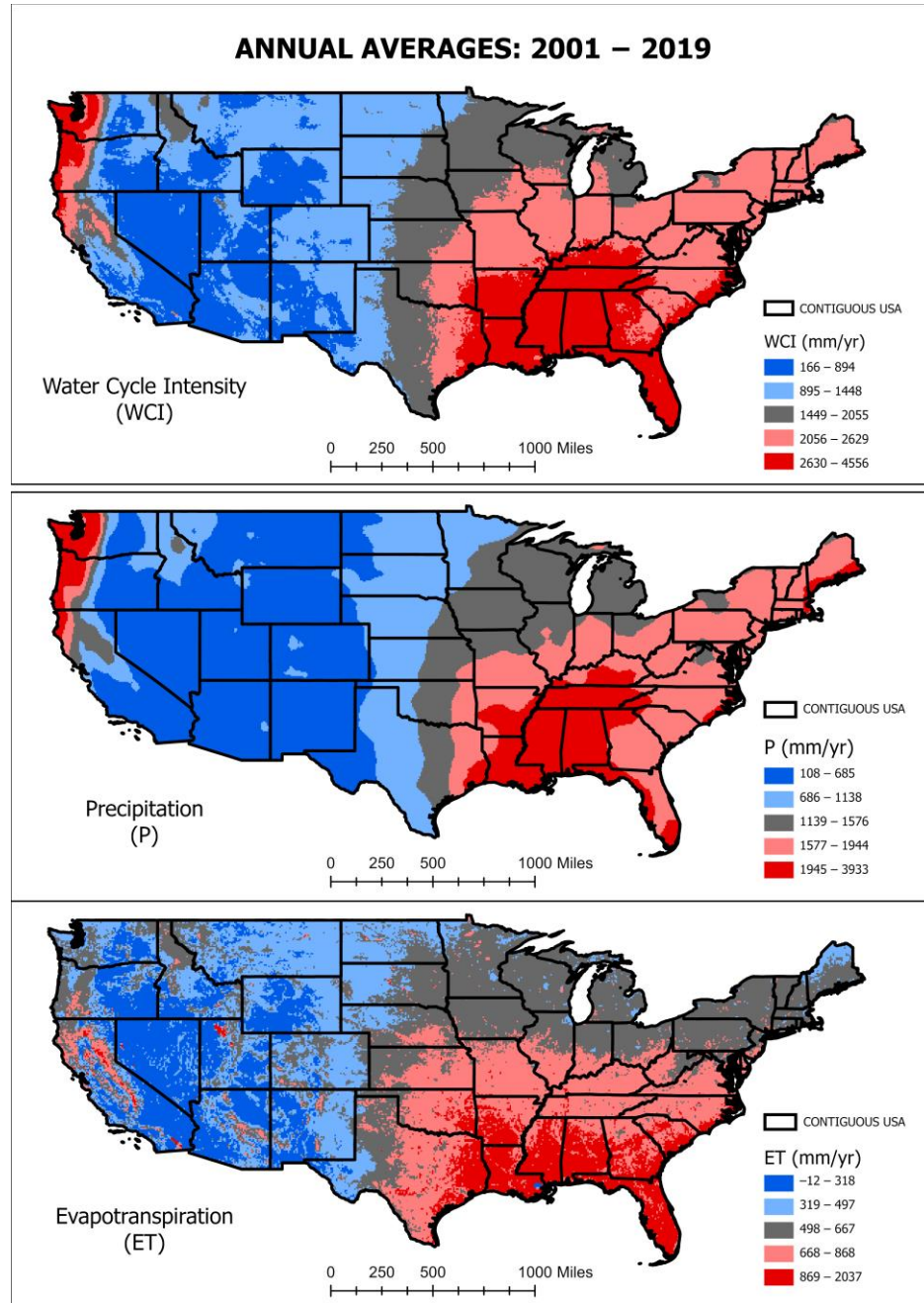


Figure 6. Average annual (a) water cycle intensity, (b) precipitation, and (c) evapotranspiration from 2001 to 2019.

To calculate the differences, the averages of the first period (2001–2009) were subtracted from the second (2010–2019). Following the initial calculations, El Niño, and La Niña years (Figure 7) were then removed from the datasets based on average annual SST anomalies and a threshold of $\pm 0.5^{\circ}\text{C}$. The El Niño years 2002, 2006, and 2015, and La Niña years 2007, 2008, and 2011 (Figure 7) were also identified in a different, independent study that compared average November-December-January (NDJ) and December-January-February (DJF) SST anomalies in the Niño 3.4 region with predefined threshold anomalies [55].

The west CONUS showed notable increases in WCI between both periods (Figure 8a), mostly due to increases in both precipitation and evapotranspiration in the second period (Figure 8b, c). However, our adjusted results (Figure 8d–f) show that during the period of study, ENSO impacted the WCI for the CONUS primarily by bringing more water to the southeast, west, and parts of the north and less water to the northeast regions.

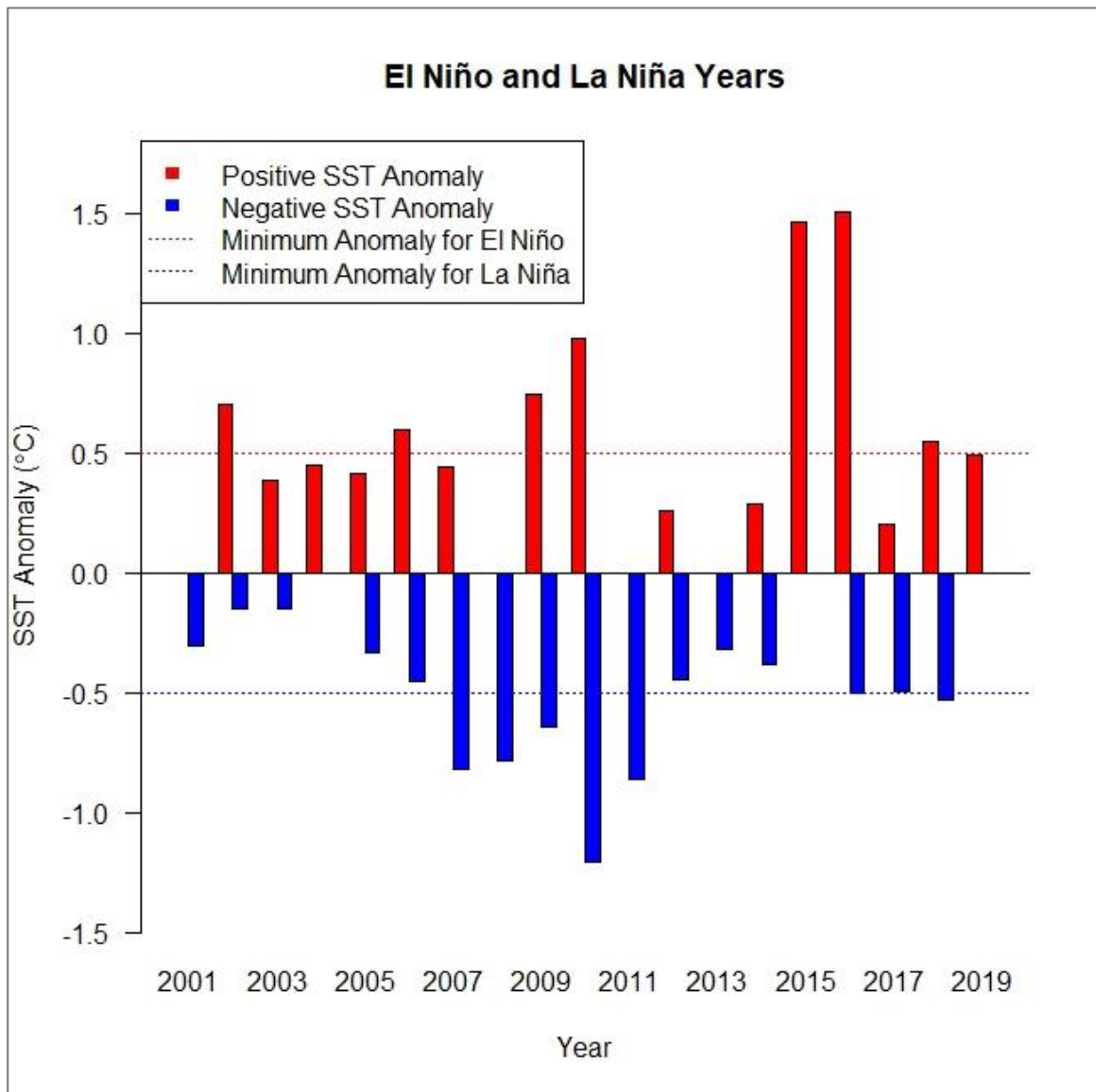


Figure 7. El Niño and La Niña years over the period of study. Average annual SST anomalies were compared with a threshold anomaly of ± 0.5 °C to identify El Niño and La Niña years, respectively. Years such as 2009, 2010, and 2018 where both positive and negative anomalies exceeded their respective El Niño and La Niña threshold anomalies were considered neutral, indicating a lack of clear dominance of either condition.

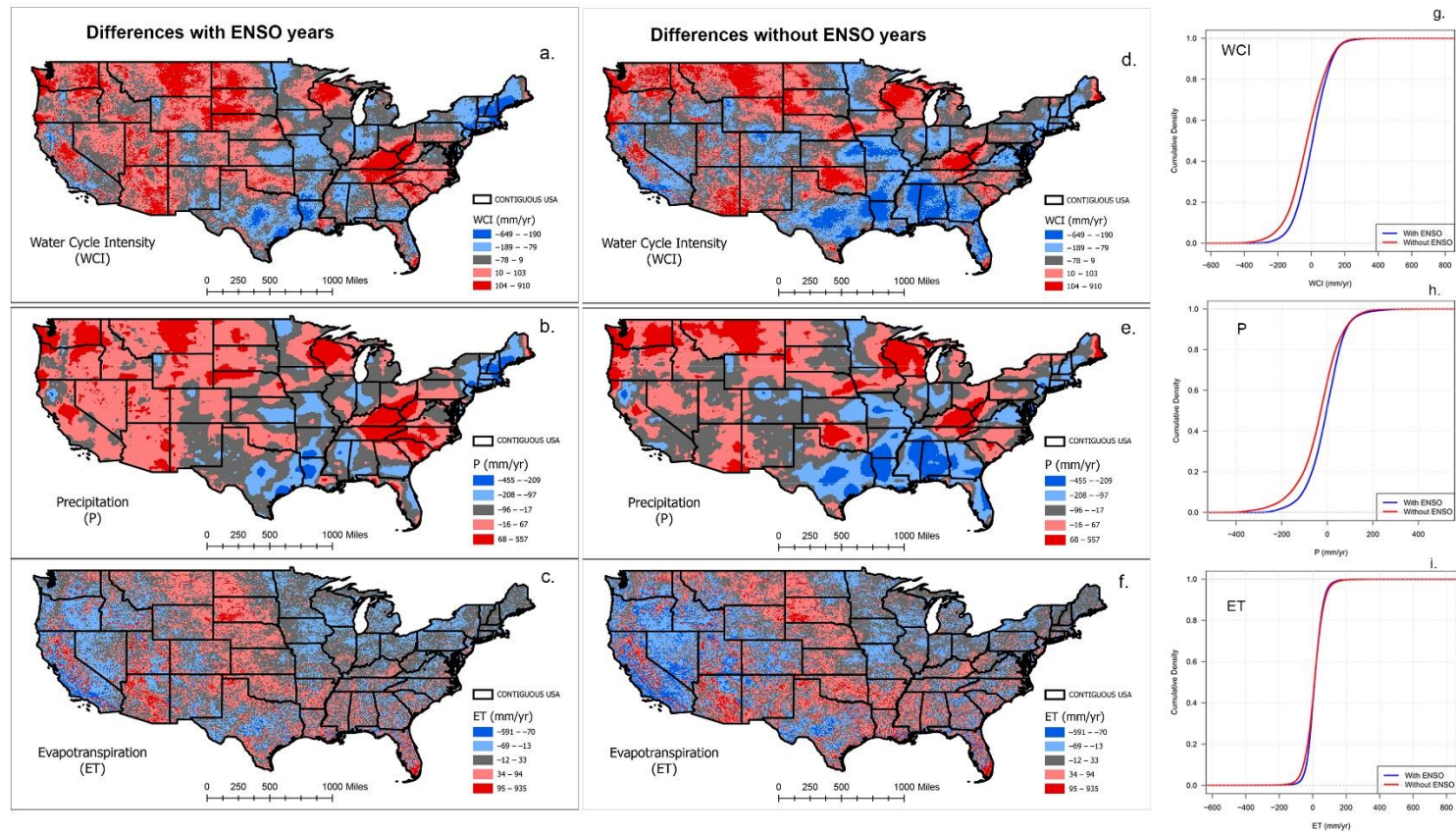


Figure 8. The difference in (a) water cycle intensity, (b) precipitation, and (c) evapotranspiration between the annual averages of 2001–2009 and 2010–2019, and their respective differences without ENSO years (d–f). The cumulative density function (CDF) plots show the cumulative probabilities of change values for water cycle intensity (g), precipitation (h), and evapotranspiration (i) for both scenarios

Discussion

Satellite precipitation retrievals generally experience difficulties due to complex terrains and climate [5,56–61]. The GPM, in particular, is less accurate in mountainous regions than in plains [36,62–66]. Specific to the CONUS, snowfall underestimation of the GPM has been recorded in the western mountainous regions [36]. Likewise, we detected an underestimation in the Dfc climate over the Rockies. For each sample point in the region, the NEXRAD-TDWR radar coverage [67] showed a beam height of less than 3000 feet above ground level, indicating good ground-based radar accuracy for the PRISM input. In addition, a local hill shade DEM for each point showed complex mountainous terrain that can affect GPM accuracy and a high elevation where snowfall and seasonal snow cover is expected. The research suggests that the complex terrain and high amounts of snow in the winter played a part in the underestimation of precipitation by the GPM for the Dfc climate region of the CONUS. However, micro-scale studies at finer spatial resolutions that account for more variables, such as local radar propagation characteristics and GPM flyovers to PRISM updates, are necessary to state this conclusively.

The WCI results presented in this study reinforce the intuition that with precipitation and evapotranspiration as indicators, the WCI over any area might increase through one of three scenarios: (1) Increases in both precipitation and evapotranspiration, as seen in parts of the west CONUS. A few studies validate this [68,69], where the warming-induced intensification of the water cycle over the Midwest resulted in increased precipitation and evaporation rates. (2) A dominating influence of precipitation, notably in the Appalachian region of the east CONUS. This scenario may carry flood risks, as flood occurrences have consistently been reported in the south and central Appalachian regions

[70–72]. (3) A dominating influence of evapotranspiration, showing examples in parts of the west and southeast regions. Increases in drought risks in the West have been attributed to a combination of atmospheric, environmental, and consumptive water demands exceeding supply [73,74]. In general, scenarios 1 and 3 may directly impact regional water availability. Groundwater levels, for example, declined in the Midwest where the intensification of the water cycle increased precipitation and evaporation rates [68].

About 53% of the CONUS showed a positive WCI change, which implies that the water cycle is currently speeding up in and around more than half of the country. Of this, 58.2% showed increases in both average precipitation and evapotranspiration (scenario 1). An increase in precipitation but decrease in evapotranspiration (scenario 2) constituted 21.2%, while the reverse (scenario 3) made up 19.1%. Without ENSO years, 40.4% of the CONUS showed an intensified water cycle where 55.6%, 18.5%, and 24.5% represented the three scenarios, respectively, thereby providing supporting evidence that ENSO may be influencing the WCI of the CONUS. In both cases, however, the dominant representation of a faster water cycle for the periods of investigation is characterized by a simultaneous increase in both precipitation and evapotranspiration, where the increase in evapotranspiration is presumably the result of higher air temperatures and an increase in moisture availability. In a general sense, the impacts of ENSO on the WCI of the CONUS were primarily driven by changes in precipitation (Figure 8g–i).

A faster water cycle presents a variety of hydrological consequences. Extreme precipitation and flood events become more likely, drought risks intensify, and the availability and quality of surface and groundwater resources can be impacted. For example, in coastal areas where aquifers are susceptible to saltwater intrusion, decreasing

groundwater recharge rates may lower the water table, allow the infiltration of saltwater, and consequently increase the severity of water quality issues. Unfortunately, these effects cannot be generalized for the CONUS. Therefore, further work at local, watershed, and regional scales is required to categorically express representative hydrological impacts for the various spatial (and temporal) scales of interest.

Conclusion

This study addressed two potential considerations within the framework of the Huntington et al. (2018) [1] study. For the same region, and with similar hydroclimatic variables, we evaluated the feasibility of a remote sensing WCI for a different, more recent period, allowing us to relate WCI changes to current climate change using satellite-based datasets. The period of investigation (2001–2019) was also determined by the availability of complete-year data for the GPM IMERG precipitation (2001–present) and evapotranspiration (1982–2019) datasets. To address the limitations of satellite-based precipitation estimates, particularly in complex terrains, we incorporated comparisons with ground-based PRISM estimates to provide validation for the satellite-based dataset and additional context on its limitations. We showed that the GPM IMERG precipitation product was not statistically different from ground-based gridded PRISM estimates across all Köppen–Geiger climates within the CONUS despite inherent challenges over complex mountainous terrains.

Using the framework introduced by Huntington et al. (2018), we present the first satellite-based quantification of the terrestrial water cycle intensity. Comparing the results of both studies, the spatial patterns of average WCI changes presented in this study contrast

those of the framework study. As we've shown, between the averages of 2001 to 2009 and 2010 to 2019, faster WCIs were seen around the west CONUS, contrary to the Huntington et al. (2018) study that showed a vulnerability of the east CONUS between the averages of 1945 to 1974 and 1985 to 2014. This not only substantiates the influence of current temperature anomalies on the spatial patterns of regional water cycle intensities for the CONUS but the susceptibility of the global water cycle to climate variabilities. In addition, we attempted to more accurately associate climate change with shifts in WCIs by eliminating years that showed strong occurrences of El Niño and La Niña events from our calculations. Our recommendation for future studies is to address and quantify the impacts of other relevant cyclic patterns, particularly those not directly influenced by ENSO (El Niño and La Niña) events.

Satellite missions for hydroclimatic variables different from those utilized in this study have also enhanced various other studies relating to the water cycle. Examples include GRACE (Gravity Recovery and Climate Experiment) [75] and its follow-on mission GRACE-FO [38] for water storage anomalies, SMOS (soil moisture and ocean salinity) [76,77] and SMAP (soil moisture active passive) [78] for soil moisture, and the Water Cycle Observation Mission (WCOM) for various components of the water cycle [79]. While WCIs generally inform the availability and distribution of water resources, remote sensing provides additional benefits including the availability and consistency of measurements to calculate the WCIs for wider spatial coverages and hydrologically inaccessible and data-sparse areas. A satellite-based WCI approach can be applied to any geographical region, covering spatial and temporal scales for which reliable satellite-based estimates are available.

In particular, the ability to calculate the WCI for any point on the globe and be able to determine if WCI changes are primarily driven by changes in precipitation, evaporation, or both, can allow regional water management agencies to make better-informed decisions on the storage and efficient distribution of water supplies. For example, negative WCI changes where precipitation is the decreasing variable and evapotranspiration is constant suggests that more water may need to be imported to the region or stored than previously in order to support the same amount of evapotranspiration. Conversely, positive WCI changes driven by evapotranspiration indicate the vital role of limited soil moisture in reducing the vapor pressure deficit in the atmosphere. Possible remedies may include supplementary water supply systems such as irrigation schemes to sustain agricultural practices, or importation to secure water for households and industries. Leveraging insights from WCI changes and harnessing the benefits of remote sensing for such calculations will help mitigate and overcome various water resources and associated socio-economic issues. It is our hope that this knowledge contributes to further work in refining water budget calculations to plan for an uncertain climate future.

References

1. Huntington, T.G.; Weiskel, P.K.; Wolock, D.M.; McCabe, G.J. A new indicator framework for quantifying the intensity of the terrestrial water cycle. *J. Hydrol.* **2018**, *559*, 361–372.
2. Douville, H.; Raghavan, K.; Renwick, J.; Allan, R.P.; Arias, P.A.; Barlow, M.; Cerezo-Mota, R.; Cherchi, A.; Gan, T.Y.; Gergis, J.; et al. Water Cycle Changes. In *Climate Change 2021: The Physical Science Basis. Contribution of Working*

- Group I to 45 the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., et al., Eds.; Cambridge University Press: Cambridge, UK, 2021; pp. 1055–1210. (*in press*).
3. United States Environmental Protection Agency. Climate Change Indicators: Weather and Climate. Available online: <https://www.epa.gov/climate-indicators/climate-change-indicators-us-and-global-temperature> (accessed on 18 November 2022).
 4. Zhan, S.; Song, C.; Wang, J.; Sheng, Y.; Quan, J. A Global Assessment of Terrestrial Evapotranspiration Increase Due to Surface Water Area Change. *Earth's Futur.* **2019**, *7*, 266–282.
 5. Milewski, A.; Seyoum, W.M.; Elkadiri, R.; Durham, M. Multi-Scale Hydrologic Sensitivity to Climatic and Anthropogenic Changes in Northern Morocco. *Geosciences* **2019**, *10*, 13.
 6. Milewski, A.; Elkadiri, R.; Durham, M. Assessment and Comparison of TMPA Satellite Precipitation Products in Varying Climatic and Topographic Regimes in Morocco. *Remote Sens.* **2015**, *7*, 5697–5717.
 7. Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353.
 8. Nkiaka, E.; Nawaz, N.R.; Lovett, J.C. Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: Case study of the Logone catchment, Lake Chad Basin. *Meteorol. Appl.* **2017**, *24*, 9–18.

9. Buytaert, W.; Friesen, J.; Liebe, J.; Ludwig, R. Assessment and Management of Water Resources in Developing, Semi-arid and Arid Regions. *Water Resour. Manag.* **2012**, *26*, 841–844.
10. Hughes, D.A. Comparison of satellite rainfall data with observations from gauging station networks. *J. Hydrol.* **2006**, *327*, 399–410.
11. Abdelmoneim, H.; Soliman, M.R.; Moghazy, H.M. Evaluation of TRMM 3B42V7 and CHIRPS Satellite Precipitation Products as an Input for Hydrological Model over Eastern Nile Basin. *Earth Syst. Environ.* **2020**, *4*, 685–698.
12. Massari, C.; Brocca, L.; Pellarin, T.; Abramowitz, G.; Filippucci, P.; Ciabatta, L.; Maggioni, V.; Kerr, Y.; Prieto, D.F. A daily 25 km short-latency rainfall product for data-scarce regions based on the integration of the Global Precipitation Measurement mission rainfall and multiple-satellite soil moisture products. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 2687–2710.
13. Satgé, F.; Defrance, D.; Sultan, B.; Bonnet, M.-P.; Seyler, F.; Rouché, N.; Pierron, F.; Paturel, J.-E. Evaluation of 23 gridded precipitation datasets across West Africa. *J. Hydrol.* **2020**, *581*, 124412.
14. Dieulin, C.; Mahé, G.; Paturel, J.E.; Ejjiyar, S.; Trambly, Y.; Rouché, N.; El Mansouri, B. A new 60-year 1940/1999 monthly-gridded rainfall data set for Africa. *Water* **2019**, *11*, 387.
15. Crétat, J.; Vizzy, E.K.; Cook, K.H. How well are daily intense rainfall events captured by current climate models over Africa? *Clim. Dyn.* **2014**, *42*, 2691–2711.

16. Trejo, F.J.P.; Barbosa, H.Á.; Peñaloza-Murillo, M.A.; Moreno, M.A.; Farias, A. Intercomparison of improved satellite rainfall estimation with CHIRPS gridded product and rain gauge data over Venezuela. *Atmósfera* **2016**, *29*, 323–342.
17. Salio, P.; Hobouchian, M.P.; Skabar, Y.G.; Vila, D. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmos. Res.* **2015**, *163*, 146–161.
18. Rozante, J.R.; Moreira, D.S.; De Goncalves, L.G.; Vila, D.A. Combining TRMM and surface observations of precipitation: Technique and validation over South America. *Am. Meteorol. Soc.* **2010**, *25*, 885–894.
19. de Angelis, C.F.; McGregor, G.R.; Kidd, C. A 3-year climatology of rainfall characteristics over tropical and subtropical South America based on tropical rainfall measuring mission precipitation radar data. *Int. J. Clim.* **2004**, *24*, 385–399.
20. Dzotsi, K.A.; Matyas, C.J.; Jones, J.W.; Baigorria, G.; Hoogenboom, G. Understanding high resolution space-time variability of rainfall in southwest Georgia, United States. *Int. J. Climatol.* **2014**, *34*, 3188–3203.
21. Ngene, B.; Agunwamba, J.; Nwachukwu, B.; Okoro, B. The Challenges to Nigerian Raingauge Network Improvement. *Res. J. Environ. Earth Sci.* **2015**, *7*, 68–74.
22. Brocca, L.; Massari, C.; Pellarin, T.; Filippucci, P.; Ciabatta, L.; Camici, S.; Kerr, Y.H.; Fernández-Prieto, D. River flow prediction in data scarce regions: Soil moisture integrated satellite rainfall products outperform rain gauge observations in West Africa. *Sci. Rep.* **2020**, *10*, 12517.

23. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116.
24. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM Precipitation Products over Singapore. *Remote Sens.* **2017**, *9*, 720.
25. Beria, H.; Nanda, T.; Bisht, D.S.; Chatterjee, C. Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6117–6134.
26. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China. *Remote Sens.* **2016**, *8*, 472.
27. El Kenawy, A.M.; McCabe, M.F.; Lopez-Moreno, J.I.; Hathal, Y.; Robaa, S.M.; Al Budeiri, A.L.; Jadoon, K.Z.; Abouelmagd, A.; Eddenjal, A.; Domínguez-Castro, F.; et al. Spatial assessment of the performance of multiple high-resolution satellite-based precipitation data sets over the Middle East. *Int. J. Clim.* **2019**, *39*, 2522–2543.
28. Mei, Y.; Anagnostou, E.N.; Nikolopoulos, E.I.; Borga, M. Error Analysis of Satellite Precipitation Products in Mountainous Basins. *J. Hydrometeorol.* **2014**, *15*, 1778–1793.
29. Elhamid, A.M.; Eltahan, A.M.; Mohamed, L.M.; Hamouda, I.A. Assessment of the two satellite-based precipitation products TRMM and RFE rainfall records using ground based measurements. *Alex. Eng. J.* **2020**, *59*, 1049–1058.
30. Zhang, T.; Yang, Y.; Dong, Z.; Gui, S. A Multiscale Assessment of Three Satellite Precipitation Products (TRMM, CMORPH, and PERSIANN) in the Three Gorges Reservoir Area in China. *Adv. Meteorol.* **2021**, *2021*, 1–27.

31. Skofronick-Jackson, G.; Petersen, W.A.; Berg, W.; Kidd, C.; Stocker, E.F.; Kirschbaum, D.B.; Kakar, R.; Braun, S.A.; Huffman, G.J.; Iguchi, T.; et al. The Global Precipitation Measurement (GPM) Mission for Science and Society. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1679–1695.
32. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Tan, J. Integrated Multi-Satellite Retrievals for GPM (IMERG) technical documentation. *Nasa/Gsfc Code* **2015**, 612, 2019.
33. Yang, X.; Lu, Y.; Tan, M.L.; Li, X.; Wang, G.; He, R. Nine-Year Systematic Evaluation of the GPM and TRMM Precipitation Products in the Shuaishui River Basin in East-Central China. *Remote Sens.* **2020**, *12*, 1042.
34. He, Z.; Yang, L.; Tian, F.; Ni, G.; Hou, A.; Lu, H. Intercomparisons of Rainfall Estimates from TRMM and GPM Multisatellite Products over the Upper Mekong River Basin. *J. Hydrometeorol.* **2017**, *18*, 413–430.
35. Wang, J.; Petersen, W.A.; Wolff, D.B. Validation of Satellite-Based Precipitation Products from TRMM to GPM. *Remote Sens.* **2021**, *13*, 1745.
36. Pradhan, R.K.; Markonis, Y.; Godoy, M.R.V.; Villalba-Pradas, A.; Andreadis, K.M.; Nikolopoulos, E.I.; Papalexiou, S.M.; Rahim, A.; Tapiador, F.J.; Hanel, M. Review of GPM IMERG performance: A global perspective. *Remote Sens. Environ.* **2022**, *268*, 112754.
37. Elnashar, A.; Wang, L.; Wu, B.; Zhu, W.; Zeng, H. Synthesis of global actual evapotranspiration from 1982 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 447–480.
38. Rodell, M.; Reager, J.T. Water cycle science enabled by the GRACE and GRACE-FO satellite missions. *Nat. Water* **2023**, *1*, 47–59.

39. McPhaden, M.J.; Zebiak, S.E.; Glantz, M.H. ENSO as an Integrating Concept in Earth Science. *Science* **2006**, *314*, 1740–1745.
40. Wang, Z.; Chang, C.P.; Wang, B. Impacts of El Niño and La Niña on the US climate during northern summer. *J. Clim.* **2007**, *20*, 2165–2177.
41. Trenberth, K.E.; Guillemot, C.J. Physical Processes Involved in the 1988 Drought and 1993 Floods in North America. *J. Clim.* **1996**, *9*, 1288–1298.
42. Jong, B.T.; Ting, M.; Seager, R. El Niño's impact on California precipitation: Seasonality, regionality, and El Niño intensity. *Environ. Res. Lett.* **2016**, *11*, 054021.
43. Rippey, B.R. The US drought of 2012. *Weather. Clim. Extrem.* **2015**, *10*, 57–64.
44. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Xie, P.; Yoo, S.H. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theor. Basis Doc. (ATBD)* **2015**, *4*, 26.
45. Beck, H.E.; Zimmermann, N.E.; McVicar, T.R.; Vergopolan, N.; Berg, A.; Wood, E.F. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* **2018**, *5*, 1–12.
46. Kotteck, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* **2006**, *15*, 259–263.
47. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91.

48. Wang, Y.; Hu, K.; Huang, G.; Tao, W.; Asymmetric impacts of El Niño and La Niña on the Pacific–North American teleconnection pattern: The role of subtropical jet stream. *Environ. Res. Lett.* **2021**, *16*, 114040.
49. Jia, X.; Ge, J.; Wang, S. Diverse impacts of ENSO on wintertime rainfall over the Maritime Continent. *Int. J. Clim.* **2016**, *36*, 3384–3397.
50. Lam, H.C.Y.; Haines, A.; McGregor, G.; Chan, E.Y.Y.; Hajat, S. Time-Series Study of Associations between Rates of People Affected by Disasters and the El Niño Southern Oscillation (ENSO) Cycle. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3146.
51. Met Office El Niño Southern Oscillation (ENSO) Region Sea Surface Temperature Forecasts. Available online: <https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/el-nino-la-nina> (accessed on 30 June 2023).
52. Trenberth, K. The Climate Data Guide: Nino SST Indices (Nino 1 + 2, 3, 3.4, 4; ONI and TNI). Available online: <https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni> (accessed on 12 July 2022).
53. Holland, S. Data Analysis in the Geosciences. Available online: <http://strata.uga.edu/8370/lecturenotes/pvaluesConfidenceIntervals.html> (accessed on 7 September 2021).
54. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and water quality models: Performance measures and evaluation criteria. *Trans. ASABE* **2015**, *58*, 1763–1785.

55. Santoso, A.; Mcphaden, M.J.; Cai, W. The defining characteristics of ENSO extremes and the strong 2015/2016 El Niño. *Rev. Geophys.* **2017**, *55*, 1079–1129.
56. Gonzalez, R.L. Consistency in the AMSR-E Snow Products: Groundwork for a Coupled Snowfall and SWE Algorithm. Doctoral dissertation, Colorado State University, Fort Collins, CO, USA, 2019.
57. Behrangi, A.; Bormann, K.J.; Painter, T.H. Using the Airborne Snow Observatory to Assess Remotely Sensed Snowfall Products in the California Sierra Nevada. *Water Resour. Res.* **2018**, *54*, 7331–7346.
58. Tang, G.; Ma, Y.; Long, D.; Zhong, L.; Hong, Y. Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *J. Hydrol.* **2016**, *533*, 152–167.
59. Seyyedi, H.; Anagnostou, E.N.; Beighley, E.; McCollum, J. Satellite-driven downscaling of global reanalysis precipitation products for hydrological applications. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 5077–5091.
60. AghaKouchak, A.; Behrangi, A.; Sorooshian, S.; Hsu, K.; Amitai, E. Evaluation of satellite-retrieved extreme precipitation rates across the central United States. *J. Geophys. Res. Atmos.* **2011**, *116*.
61. Turk, F.; Miller, S. Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1059–1069.
62. Moazami, S.; Najafi, M. A comprehensive evaluation of GPM-IMERG V06 and MRMS with hourly ground-based precipitation observations across Canada. *J. Hydrol.* **2021**, *594*, 125929.

63. Yang, M.; Liu, G.; Chen, T.; Chen, Y.; Xia, C. Evaluation of GPM IMERG precipitation products with the point rain gauge records over Sichuan, China. *Atmos. Res.* **2020**, *246*, 105101.
64. Lu, X.; Tang, G.; Wang, X.; Liu, Y.; Jia, L.; Xie, G.; Li, S.; Zhang, Y. Correcting GPM IMERG precipitation data over the Tianshan Mountains in China. *J. Hydrol.* **2019**, *575*, 1239–1252.
65. Wu, Y.; Zhang, Z.; Huang, Y.; Jin, Q.; Chen, X.; Chang, J. Evaluation of the GPM IMERG v5 and TRMM 3B42 v7 Precipitation Products in the Yangtze River Basin, China. *Water* **2019**, *11*, 1459.
66. Sungmin, O.; Kirstetter, P. Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 270–281.
67. Center, R.O. NEXRAD and TDWR Radar Locations. Available online: <https://www.roc.noaa.gov/WSR88D/Maps.aspx> (accessed on 28 November 2022).
68. Yeh, P.J.F.; Wu, C.H. Recent Acceleration of the Terrestrial Hydrologic Cycle in the U.S. Midwest. *J. Geophys. Res. Atmos.* **2018**, *123*, 2993–3008.
69. Baker, J.M.; Griffis, T.J.; Ochsner, T.E. Coupling landscape water storage and supplemental irrigation to increase productivity and improve environmental stewardship in the U.S. Midwest. *Water Resour. Res.* **2012**, *48*.
70. Miller, D.; Forsythe, J.; Kusselson, S.; Straka III, W.; Yin, J.; Zhan, X.; Ferraro, R. A Study of Two Impactful Heavy Rainfall Events in the Southern Appalachian

- Mountains during Early 2020, Part I; Societal Impacts, Synoptic Overview, and Historical Context. *Remote Sens.* **2021**, *13*, 2452.
71. Wix, J.M. Constructing a Flash Flood Climatology of the Appalachian Mountains: Comparison of Eastern Kentucky and West Virginia Summertime Rain Events. Undergraduate Thesis, Western Kentucky University, Bowling Green, KY, USA, 2010.
 72. Gaffin, D.M.; Hotz, D.G. A precipitation and flood climatology with synoptic features of heavy rainfall across the southern Appalachian Mountains. *Natl. Weather. Dig.* **2000**, *24*, 3–16.
 73. Martin, J.T.; Pederson, G.T.; Woodhouse, C.A.; Cook, E.R.; McCabe, G.J.; Anchukaitis, K.J.; Wise, E.K.; Erger, P.J.; Dolan, L.; McGuire, M.; et al. Increased drought severity tracks warming in the United States' largest river basin. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 11328–11336.
 74. US Bureau of Reclamation. Colorado River Basin Water Supply and Demand Study: Executive Summary. Available online: https://www.usbr.gov/watersmart/bsp/docs/finalreport/ColoradoRiver/CRBS_Executive_Summary_FINAL.pdf (accessed on 18 November 2022).
 75. Eicker, A.; Forootan, E.; Springer, A.; Longuevergne, L.; Kusche, J. Does GRACE see the terrestrial water cycle “intensifying”? *J. Geophys. Res. Atmos.* **2016**, *121*, 733–745.
 76. Durack, P. Ocean Salinity and the Global Water Cycle. *Oceanography* **2015**, *28*, 20–31.

77. Reul, N.; Fournier, S.; Boutin, J.; Hernandez, O.; Maes, C.; Chapron, B.; Alory, G.; Quilfen, Y.; Tenerelli, J.; Morisset, S.; et al. Sea Surface Salinity Observations from Space with the SMOS Satellite: A New Means to Monitor the Marine Branch of the Water Cycle. *Surv. Geophys.* **2014**, *35*, 681–722.
78. Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The Soil Moisture Active Passive (SMAP) Mission. *Proc. IEEE* **2010**, *98*, 704–716.
79. Shi, J.; Dong, X.; Zhao, T.; Du, Y.; Liu, H.; Wang, Z.; Zhu, D.; Ji, D.; Xiong, C.; Jiang, L. The water cycle observation mission (WCOM): Overview. *IEEE Int. Geosci. Remote Sens. Symp.* **2016**, 3430–3433.

CHAPTER 3

GROUNDWATER LEVEL PREDICTION USING MACHINE LEARNING AND
GEOSPATIAL INTERPOLATION MODELS ¹

¹ Zowam, F.J. and Milewski, A.M. Submitted to *Water* 16(19), 2771.

Reprinted here with permission of the publisher.

Abstract

Given the vulnerability of surface water to the direct impacts of climate change, the accurate prediction of groundwater levels has become increasingly important, particularly for dry regions, offering significant resource management benefits. This study presents the first statewide groundwater level anomaly (GWLA) prediction for Arizona across its two distinct aquifer types—unconsolidated sand and gravel aquifers and rock aquifers. Machine learning (ML) models were combined with empirical Bayesian kriging (EBK) geostatistical interpolation models to predict monthly GWLAs between January 2010 and December 2019. Model evaluations were based on the Nash–Sutcliffe efficiency (NSE) and coefficient of determination (R^2) metrics. With average NSE/ R^2 values of 0.62/0.63 and 0.72/0.76 during the validation and test phases, respectively, our multi-model approach demonstrated satisfactory performance, and the predictive accuracy was much higher for the unconsolidated sand and gravel aquifers. By employing a remote sensing-based approach, our proposed model design can be replicated for similar climates globally, and hydrologically data-sparse and remote areas of the world are not left out.

Introduction

Global warming is threatening surface water supply in many parts of the world, particularly in arid regions. In these regions, groundwater, a viable alternative and an important source of freshwater, is often limited [1,2]. Considering that groundwater level (GWL) is an indicator of groundwater availability at any given time, monitoring GWLs provides significant insights into the dynamics of recharge and withdrawals and how they influence the long-term availability of groundwater. In arid regions, this can be challenging due to the inadequate distribution of groundwater wells and the presence of spatial and temporal data gaps in monitoring records [3]. Therefore, accurate and reliable predictive tools are essential for supporting the sustainable management of groundwater in these areas [3].

The relationships between GWL fluctuations and explanatory variables are generally complex and nonlinear [4–7]. However, machine learning (ML) algorithms can effectively learn these relationships. One of such algorithms is the support vector machine (SVM) for regression purposes (SVR), especially when embedded with the radial basis function (RBF) kernel [5,8]. Another algorithm capable of learning these relationships is random forest (RF) [9], which is the most employed ML technique for GWL predictions [10]. Both SVMs and RF are known to give accurate results [10,11].

Several studies have shown the RBF–SVR model to outperform various other techniques, such as artificial neural network (ANN) [12–14], radial basis function neural network (RBF–NN) [15], the autoregressive integrated moving average (ARIMA) model [16], RF [8], and the gradient boosting mechanism (GBM) [8]. These studies also attribute

the success of SVMs to their strong prediction capability and the ability to generalize well to unseen data.

Likewise, various studies have reported success using RF for GWL prediction. For example, it outperformed K-nearest neighbor (KNN), ANN, and SVR based on root mean square error (RMSE) values during testing [17]; ANN and SVR based on R-squared (R^2), mean absolute error (MAE), and RMSE values in training and R^2 and MAE values in testing [11]; multilinear regression (MLR) based on R^2 , MAE, and RMSE values in both training and testing [18]; decision trees (DTs) and SVRs based on its R^2 and RMSE values in testing [19]; and the XGBoost regressor based on its MAE and RMSE values in testing [20].

In addition to traditional ML methods, geostatistical interpolation (GI) techniques are also commonly used for GWL prediction. Kriging is the most utilized GI technique [21,22], where the measured values of a variable (GWL, in this case) at specific locations are used to make predictions at unmeasured areas. It relies on the correlation between the measured values as a function of distance, depicted by a semivariogram, to assign weights that describe the contribution of each measured point to the prediction at unmeasured locations [23]. Kriging presents an important advantage over other interpolation methods in being able to quantify and minimize prediction uncertainties [23].

Classical kriging (CK), the traditional form of kriging, relies on a single semivariogram assumed to be the true representation of the measured data. In contrast, empirical Bayesian kriging (EBK), an advanced kriging approach, incorporates multiple semivariograms to account for the uncertainties associated with estimating a single semivariogram [23]. Thus, EBK is a more robust kriging algorithm [23], and studies such

as Bouhout et al. (2022) [24] and Hussain et al. (2016) [25] have demonstrated its superior performance in GWL prediction applications.

Our review of the existing literature suggests that RBF–SVR, RF, and EBK models are some of the most effective GWL prediction tools. To enhance the prediction accuracy, we propose an approach that integrates all three techniques to predict the monthly GWLAs across the state of Arizona (representative of arid/semi-arid systems) between January 2010 and December 2019, using remotely sensed predictor variables.

Materials and Methods

Natural groundwater recharge in arid regions is often limited, raising concerns about the sustained availability of freshwater in these regions. Climate change exacerbates these issues by intensifying the hydrologic cycle, resulting in increased evapotranspiration rates and a reduction in the soil moisture available to recharge groundwater systems [26,27]. In a recent study, Zowam et al. (2023) [27] quantified terrestrial water cycle intensity (WCI) changes across the contiguous United States (CONUS) attributable to climate change and showed that the state of Arizona might be experiencing much higher relative WCI rates on average than other arid regions in the CONUS. These factors underscore the need for accurate prediction of the GWL in such regions to continue to effectively manage the potentially limited freshwater resources therein.

Study Area and Target Variable

The study area, covering about 114,000 mi², is located in the southwest U.S. (Figure 1). Its surficial geology is characterized primarily by unconsolidated deposits in the south/southwest–northwest corner of the state, with various rocks dominating the other

regions [28]. The consolidated rocks are mainly sedimentary and extrusive igneous (volcanic) rocks and constitute the mountain ranges that border the basins filled with unconsolidated materials [29]. These rocks are the main sources of sedimentary materials that fill the basin and have very low permeability and groundwater flow rates [29]. Many communities depend solely on groundwater to meet their water needs, which has led to a long history of over-extraction in many parts of the state [30]. In the late 1980s, the Colorado river arrived in Arizona and eased some of the pressure on groundwater to meet these needs, but the prolonged drought in the Colorado river basin, coupled with projected warming temperatures, is expected to not only reduce the surface water availability in the state but also further stress the aquifers in the region [30].

Daily GWL data from 59 monitoring wells were downloaded from the National Groundwater Monitoring Network (NGWMN) portal (<https://cida.usgs.gov/ngwmn/index.jsp>, accessed on 29 January 2023). Among these wells, 38 were drilled into unconsolidated sand and gravel aquifers and are managed by the Arizona Department of Water Resources (ADWR), while the remaining 21 were drilled into consolidated rock aquifers (Figure 1). Three of the rock aquifer wells used in this study are maintained by the U.S. Geological Survey (USGS), and the rest are managed by the ADWR. The depths of these wells varied and ranged from 89 to 1600 feet below ground surface (bgs) for the sand and gravel aquifers, and about 25 to 851 feet bgs for the rock aquifers, and the aquifers were predominantly unconfined. Daily GWL measurements were aggregated into monthly averages from January 2010 to December 2019.

Missing data is the most common challenge in real-world ML applications [31], and various methods exist to address this issue. The simplest of these methods is mean

imputation (MI) [31]. In MI, the mean values of available observations are used to fill in missing observations, which has proven to work well with small variance distributions [32,33], i.e., distributions with a coefficient of variation (CV) less than 10% [34]. Therefore, with CV values ranging from 0.01 to 0.91% for the unconsolidated aquifers and 0.001 to 0.21% for the rock aquifers, the MI method was ideal for the target variable (monthly GWL). Missing monthly values were replaced with the annual average GWL for the given year.

To compute monthly GWLA, we calculated the mean GWL at each well and subtracted the monthly measurements from this mean value.

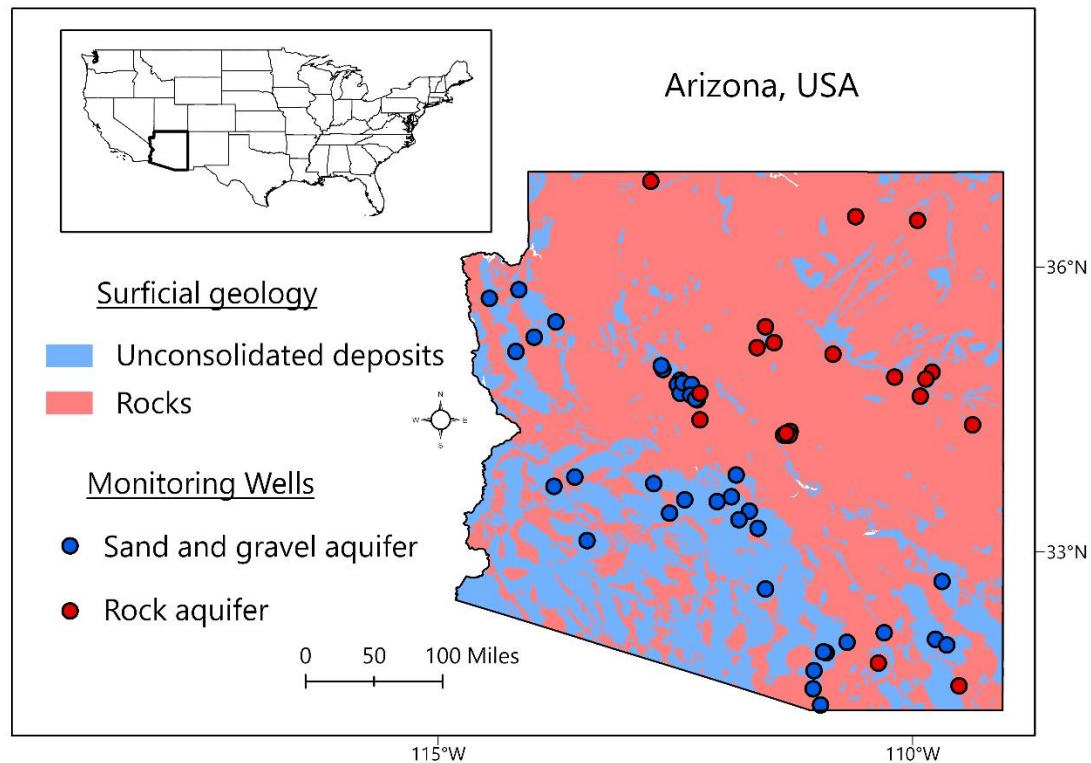


Figure 1. Map of the study area showing the locations of groundwater monitoring wells.

Predictor Variables

The selection of the input variables was informed by the Seyoum et al. (2019) [35] study as well as established hydrogeological principles. The initial variables included precipitation, soil moisture, evapotranspiration, land surface temperature, vegetation index, curve number, saturated hydraulic conductivity, and groundwater storage anomalies (Figure 2).

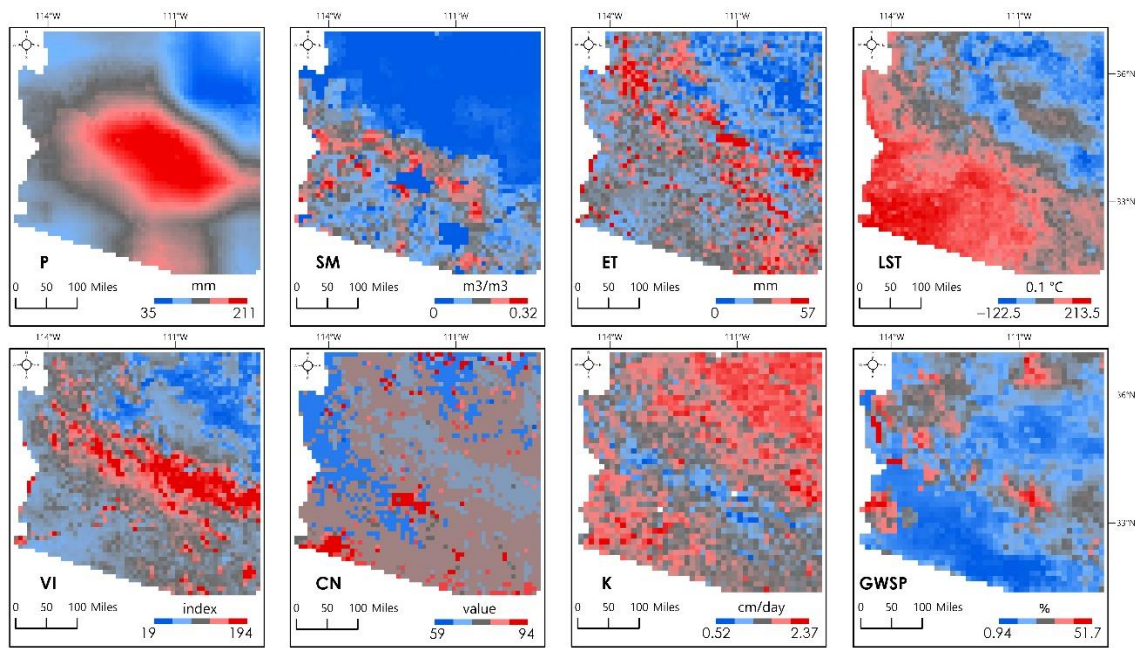


Figure 2. Maps of initial predictor variables resampled to $0.125^\circ \times 0.125^\circ$ grid resolution. P = precipitation, SM = soil moisture, ET = evapotranspiration, LST = land surface temperature, VI = vegetation index, CN = curve number, K = saturated hydraulic conductivity, GWSP = groundwater storage percentile. CN and K are representative values for the period of study, and all other variables are for January 2010.

Precipitation (P)

Precipitation is the most important hydrological variable for predicting groundwater recharge [36], thus playing a crucial role in determining GWL. In the last two decades, satellite-based precipitation measurement techniques have seen significant

advancements, and the Global Precipitation Measurement Mission (GPM) using the Integrated Multi-Satellite Retrievals for the GPM (IMERG) algorithm stands out as one of the best alternatives to ground-based measurements [27,37]. In particular, the GPM mission demonstrates significant potential to mitigate the challenges associated with estimating precipitation in arid regions [38]. In addition, under light rainfall conditions (typical of these regions), IMERG tends to produce lower detection errors and generally more accurate estimates [39]. The final run of the IMERG system provides the most accurate precipitation measurements, making it ideal for research purposes [40]. A monthly IMERG (final run) dataset with a $0.1^\circ \times 0.1^\circ$ grid resolution was downloaded from the National Aeronautics and Space Administration (NASA) data portal for January 2010 to December 2019 (<https://gpm.nasa.gov/data/directory>, accessed on 11 May 2021).

Soil Moisture (SM)

SM and GWL tend to demonstrate a negative relationship [41–43], which can be much stronger for shallow groundwater [41]. For this study, we utilized a research product that integrated measurement efforts from both the European Space Agency (ESA) and NASA (<https://doi.org/10.1594/PANGAEA.940409>, accessed on 12 February 2023). The dataset was generated by downscaling ESA's Climate Change Initiative (CCI) data using NASA's Soil Moisture Active and Passive (SMAP) data [44]. The global dataset had a grid size of 9 km ($\sim 0.08^\circ \times 0.08^\circ$) and a daily temporal resolution covering a 43-year period from 1978 to 2020 [44]. We converted the daily rasters into monthly averages spanning our study period and extracted the monthly data for the study area.

Evapotranspiration (ET)

The combined effects of liquid water losses from soil surfaces (evaporation) and its uptake by plants (transpiration) limit the amount that infiltrates the ground, thereby affecting GWLs. We utilized a global dataset with a fine grid size of 1 km ($\sim 0.01^\circ \times 0.01^\circ$) and a monthly temporal resolution (<https://doi.org/10.7910/DVN/ZGOUED>, accessed on 3 July 2021). The dataset was obtained by synthesizing the best-performing satellite ET products following validation against flux eddy covariance ET and performed better than local products across the United States, China, and the continent of Africa [45].

Land Surface Temperature (LST)

LST tends to exhibit a positive relationship with GWL, which is more pronounced for shallow groundwater [41]. This study utilized a gap-filled, continuous LST dataset generated by filling in missing pixels in the Moderate-Resolution Imaging Spectroradiometer (MODIS) 1 km resolution LST daily product [46]. Daytime (1:30 PM) and nighttime (1:30 AM) global datasets were downloaded from the Iowa State University research repository (<https://doi.org/10.25380/iastate.c.5078492.v3>, accessed on 12 February 2023). The downloaded rasters were converted into monthly averages, and the final LST dataset was obtained by averaging the daytime and nighttime monthly estimates.

Vegetation Index (VI)

The presence of vegetation may affect groundwater recharge (and GWLs) in various ways, including slowing down runoff and enhancing ET through transpiration. VI values are unitless and help to visualize the locations and relative abundance of green vegetation. A monthly dataset with a $0.1^\circ \times 0.1^\circ$ grid resolution was accessed and downloaded from the NASA Earth Observations (NEO) data portal

(https://neo.gsfc.nasa.gov/view.php?datasetId=MOD_NDVI_M&date=2010-12-01, accessed on 12 February 2023).

Curve Number (CN) and Runoff Depth (R)

CN is a dimensionless parameter that characterizes the runoff potential of a surface. It is influenced by land use, soil characteristics, and the antecedent moisture conditions of soils. Lower numbers typically correspond to permeable soils with high infiltration rates, while higher numbers are associated with impervious surfaces and limited infiltration capacities. The CNs utilized in this study were generated using a 250 m hydrological soil group dataset (HYSOG250m) and the 2015 ESA–CCI 300 m land cover dataset, and available at a 250 m grid resolution (<https://doi.org/10.6084/m9.figshare.7756202.v1>, accessed on 12 February 2023).

With the CN data, we generated a monthly dataset of R for the period of study using Grove et al. (1998)'s [47] equations:

$$R = \frac{(P - 0.2S)^2}{(P + 0.8S)} \quad (1)$$

Given that:

$$S = \left(\frac{1000}{CN} \right) - 10 \quad (2)$$

Where R = runoff depth, P = precipitation, S = potential maximum retention, and CN is the Curve Number [42].

Soil Saturated Hydraulic Conductivity (K_s and PK_s)

Soil saturated hydraulic conductivity describes the ability of soils to transmit water under saturated conditions [48]. Using remotely sensed environmental variables and the RF ML technique, Gupta et al. (2021) [49] generated four global K_s maps representing four

different soil depths at a 1 km ($\sim 0.01^\circ \times 0.01^\circ$) resolution. All four maps were utilized in this study (<https://doi.org/10.5281/zenodo.3935359>, accessed on 11 October 2023).

To obtain a continuous and dynamic K_s dataset, we multiplied each K_s value by the monthly precipitation estimates. We selected precipitation as it is a primary driver of water input into the soil and significantly influences K_s . By integrating the precipitation data with K_s , we added temporal variability to the K_s variable, making it suitable for our ML models.

$$P \times K_{si} = (PK_s)_i \quad (3)$$

Here, P is the precipitation dataset, K_s is the soil-saturated hydraulic conductivity dataset at a depth of i , and i ranges from 1 to 4. The four resulting outputs from equation 3 were averaged into a single, comprehensive PK_s dataset for the study area.

Groundwater Storage Percentile (GWSP)

Using terrestrial water storage (TWS) observations from the Gravity Recovery and Climate Experiment (GRACE) satellite mission and a numerical model representing the interactions between water and energy across the Earth's surface, scientists at NASA are able to determine weekly groundwater conditions, expressed as percentiles, based on comparison with historical data [50]. These percentiles indicate the probability of occurrence within the 1948 to 2014 period of record and are generated at a spatial resolution of $0.125^\circ \times 0.125^\circ$ over North America from April 2002 to the present [50]. We downloaded the monthly averages of the GWSP from the Giovanni data portal (<https://giovanni.gsfc.nasa.gov/giovanni/>, accessed on 10 October 2023) for our period of study. Considering all the input datasets, the GWSP had the largest spatial resolution ($0.125^\circ \times 0.125^\circ$), requiring all the other datasets to be resampled (upscaled) to the spatial resolution of the GWSP rasters.

Groundwater recharge is typically computed using a simple water balance approach assuming negligible changes in soil water storage in the unsaturated zone [51]. Based on this idea, we introduced a secondary input variable, called the recharge index (RI), to represent the balance between the water inflows and outflows within each $0.125^\circ \times 0.125^\circ$ grid and the amount potentially available to recharge groundwater.

$$RI = P - ET - R \quad (4)$$

where P is precipitation; ET is evapotranspiration; R is runoff depth; and RI is the recharge index for a given grid.

GWLs from monitoring wells located within a grid were assumed to be representative of the entire grid, as were the predictor variables. Ultimately, this study utilized six predictor variables, namely soil moisture (SM), land surface temperature (LST), the vegetation index (VI), saturated hydraulic conductivity (K_s), the groundwater storage percentile (GWSP), and the recharge index (RI) (Table 1).

Table 1. Final input variables used in the study and the processing involved.

ID	Variable	Processing	Unit
1	Soil moisture (SM)	$\times \circ$	m^3/m^3
2	Land surface temperature (LST)	$\times \Delta \circ$	$^\circ C$
3	Vegetation index (VI)	$\circ$	—
4	Saturated hydraulic conductivity (PK_s) *	$\circ \Delta \square$	mm^2/day
5	Groundwater storage percentile (GWSP)	—	%
6	Recharge index (RI) *	$\circ \circ \square$	mm

Note(s): $\times$ = convert daily data into monthly averages; Δ = raster averaging; $\circ$ = resampling to $0.125^\circ \times 0.125^\circ$; $\square$ = raster arithmetic operations. NB: Secondary variables are marked with an asterisk (*), and the pipe symbol (|) separates the processing applied to each individual variable.

Model Algorithms

Developing an acceptable ML model with a monthly temporal resolution requires approximately 10–12 years of data [52]. Here, we utilize 10 years of monthly data (2010–2019), meeting the acceptable sample size threshold. Selecting algorithms or models that are most appropriate for the data is another important aspect of ML. Getting this step wrong could result in unreliable predictions, leading to a disappointing predictive performance and misleading conclusions [53]. In this study, we selected two ML algorithms (and a GI technique) from a pool of candidates. The selected techniques are discussed in detail below:

SVM and SVR

SVM was introduced by Vladimir Vapnik based on the idea of nonlinear mapping of input vectors to a high-dimensional feature space and constructing an optimal hyperplane to effectively separate the different groups or classes within that space [54]. The generalization capabilities of SVMs led to the development of the less popular SVR for real-value (regression) problems [55].

First proposed in 1996 by Harris Drucker and his colleagues [56], the SVR has become an effective tool for prediction problems, demonstrating excellent generalization abilities and a high prediction accuracy [55]. It works by incorporating a loss function (known as the epsilon insensitive margin of error, ϵ) in the form of a flexible tube formed symmetrically above and below the estimated function, where the prediction errors (ζ^*) within the tube are accepted and those that fall outside are penalized (Figure 3). The objective of the SVR is to find the narrowest possible tube around the estimated function in a way that minimizes the prediction errors [55,57].

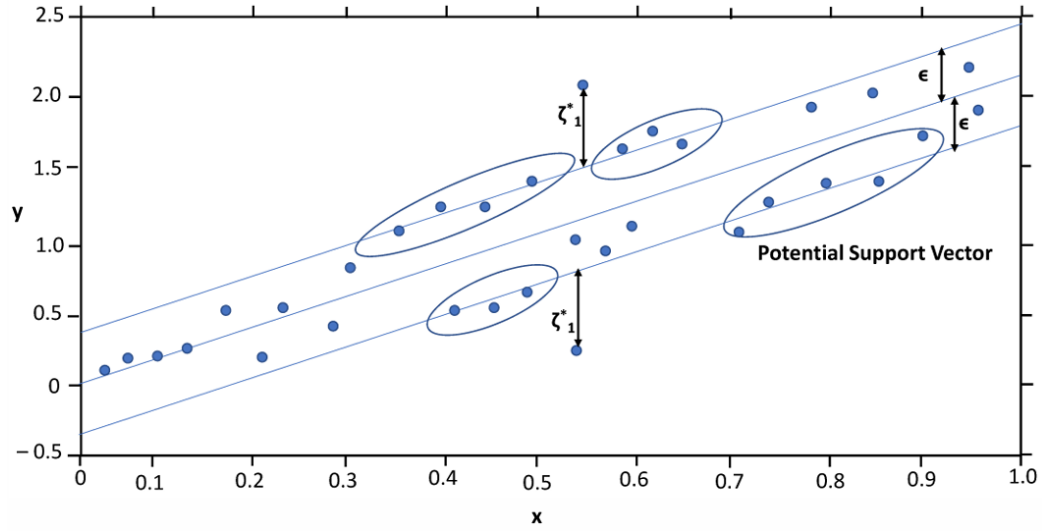


Figure 3. One-dimensional SVR, where x is the input or independent variable and y is the dependent variable. Source: Awad and Khanna (2015) [55].

In general, the performance of the SVR model depends on the tube size (epsilon, ϵ), the regularization constant (C), and the choice of kernel function [58–60]. The C hyperparameter controls the complexity of the model, where large values may lead to overfitting [58,61]. By overfitting, the model learns the training data well but generalizes poorly, i.e., predicts poorly on new unseen data. Kernel functions are used to transform the data into the higher-dimensional feature space, enabling linear machine learning to improve the representation of the nonlinear relationships that exist in the original input space [62]. While there is no guide to the appropriate kernel functions for specific datasets, the most commonly used are the RBF and polynomial functions [63]. RBFs are versatile kernels used when there is a lack of prior knowledge about the data [64]. Such models (RBF-SVR) require an additional hyperparameter, gamma (γ), in addition to C and ϵ , which controls the width of the RBF [60,65,66]. The C hyperparameter must be a positive number, while ϵ and γ can be positive or zero [66].

RF

The RF ML algorithm is an ensemble algorithm of multiple trees that improves the prediction accuracies of the single DT algorithm [19,67]. The different DTs are trained with subsets of the input variables and bootstrapped samples of the original training data such that each DT is unique, resulting in reduced variance [11]. By bootstrapping, samples are randomly drawn (with replacement) from the original training data, maintaining the sample size of the training data. Because the sampling is carried out with replacement, a particular observation may appear multiple times in a bootstrapped sample.

Decision points in the DT structure are called nodes. At the nodes, tree branches are created based on the splitting criteria (Figure 4). The first node (without prior branching) is the root [68]. From the root, each node is split using the best variable among the subset of input variables chosen at that node [69]. The leaf is the final node (with no further branching) associated with an output value [68].

RF model hyperparameters are the number of trees (DTs) and the number of input variables in the random subset at each node [11,69]. The final predictions are either determined by majority votes from individual DTs (classification) or by averaging the predictions from all the trees (regression) [69,70].

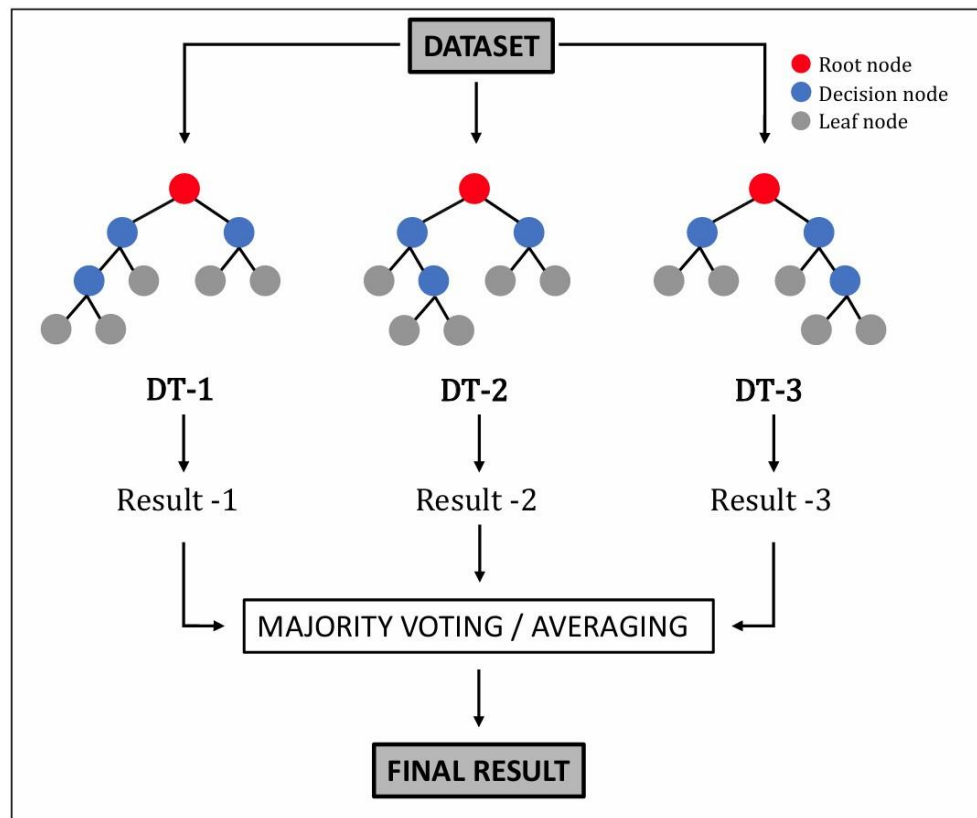


Figure 4. A simple RF Model with three DTs.

EBK

EBK differs from CK in the way it optimizes the parameter uncertainty associated with creating a single semivariogram and the way it automates the optimization process. In a single semivariogram, the semivariance (the y-axis) measures the spatial dependency between pairs of observations or samples, and the lag (the x-axis) is their separation distance (Figure 5a). Depending on the characteristics of the data, a semivariogram may display three important components: a sill, a range, and a nugget (Figure 5a). The range is the distance (or lag) beyond which samples are not spatially autocorrelated, and the sill is the semivariance at that distance. The nugget is the y-intercept of the semivariogram and

represents variability at distances much smaller than the minimum spacing between pairs of sample points.

With EBK, the input data are divided into subsets, specifying parameters such as the size of the subsets (subset size) and the degree of overlap between them (overlap factor). Within each subset, a semivariogram distribution is produced (Figure 5b), and predictions are made for each location using the distribution from one or more subsets [23].

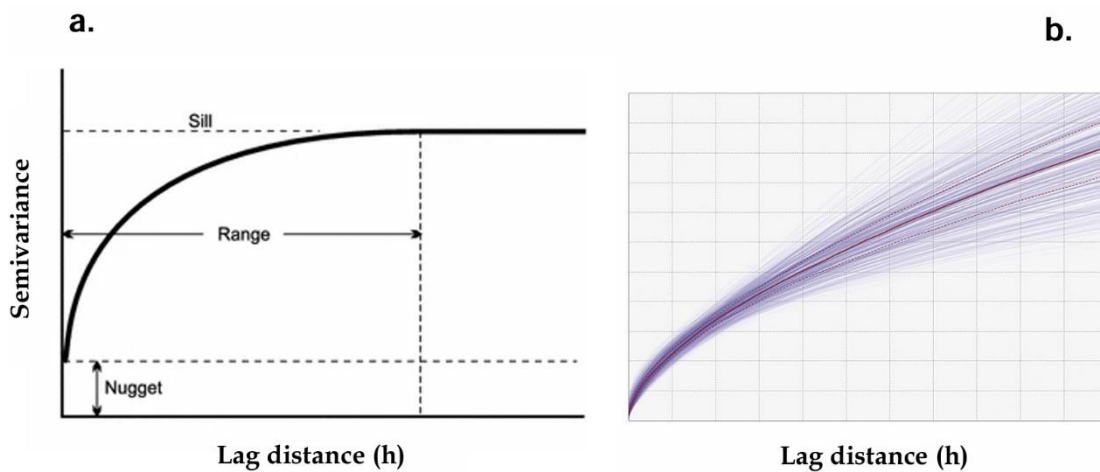


Figure 5. (a) A single semivariogram. Source: Maliva (2016) [71]. (b) The EBK model showing a distribution of semivariograms. Source: Krivoruchko (2012) [23]. The red dotted lines represent the lower and upper quartiles, and the solid red line represents the median of the semivariogram distribution [23,72].

Model Design

The initial phase of the analysis aimed to assess the feasibility of ML to capture spatiotemporal patterns of GWLAs across the study area. For spatial patterns, we trained 12 different SVR models (each corresponding to a specific month) on sixty percent of the dataset using manually tuned values of γ , C , and ϵ and tested each of the trained models with the remaining forty percent (Figure 6a). Each model was trained using three predictor

variables—LST, RI, and the previous month’s GWLA (except for the first month). Incorporating GWL from a prior time step is common practice. It is the most employed type of input data in GWL prediction [52]. To evaluate ML feasibility for temporal patterns, we utilized the same dataset as the test set but used all six predictor variables and the predicted output—replacing the “previous month GWLA” variable to ensure consistency in the number of predictors for each observation (Figure 6b). Following another random train/test split procedure, an RF model was trained on eighty-five percent of the data and tested on the remaining fifteen percent. The models were evaluated for their performance and predictive accuracy using the Nash–Sutcliffe efficiency (NSE) and the coefficient of determination (R^2), respectively. The NSE ranges from $-\infty$ to 1, while R^2 ranges from 0 to 1. For both metrics, a perfect prediction would yield a value of 1.



Figure 6. Model design showing train/test split ratios for the spatial (a) and temporal (b) evaluations and the locations of the split wells.

Based on the results of the preliminary assessments, it was evident that incorporating some approximation of the GWL as a predictor variable would significantly

enhance the model performance and predictive accuracy. Therefore, we sought to use EBK for those approximations. To ensure that the EBK predictions covered the study area, we obtained additional GWL data from 18 nearby monitoring wells outside the study area. A total of 120 EBK models were constructed using the Geostatistical Wizard interpolation tool in ArcGIS PRO, and monthly predictions from those models were converted into anomalies and incorporated into the dataset. A single RF ML model (Figure 7) was then developed to learn the patterns in substantial portions of the augmented dataset.

This approach was first validated at three monitoring well locations. The measurements corresponding to those locations were removed from the dataset before performing the EBK. Because the accuracy of kriging is significantly influenced by the number and density of kriging points, we could not afford to eliminate additional wells (and their corresponding data) from the dataset. Subsequently, the EBK process was repeated using the entire dataset, and a final RF model was trained on a significant portion of this dataset, tested on a smaller subset, and eventually deployed at locations without monitoring wells.

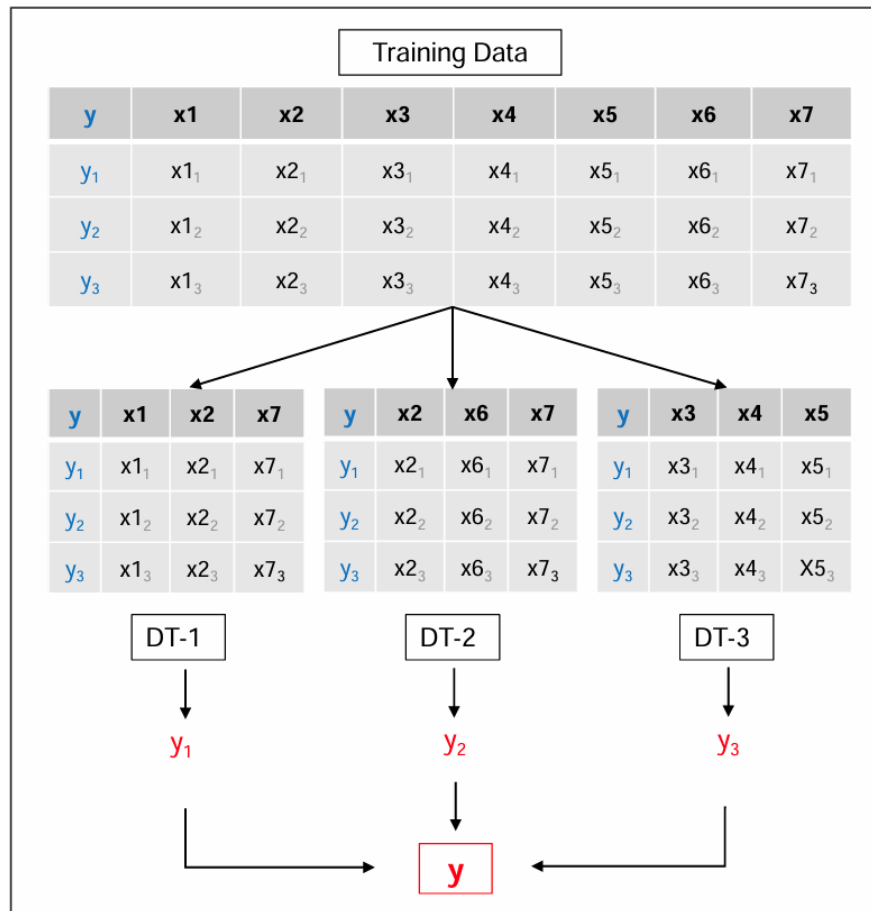


Figure 7. RF schematic with three observations, the seven predictor variables, three randomly selected predictors to build each tree, and three DTs. The target variable is shown in blue, and the predicted outputs are shown in red.

Results

Analyses were done in RStudio (version 2023.06.0+421) and ArcGIS Pro (version 2.9.3). The results are presented below.

Initial Assessment of ML Capabilities

The SVR models effectively captured the spatial patterns of GWLA variation across the study area, based on the NSE and R^2 values (Table 2). Each monthly model

(except for the January model) was trained using the previous month's GWLA data. The unavailability of previous month measurements as input for the first (January) model resulted in its suboptimal performance. However, the predictions from subsequent models were averaged with the January prediction to obtain the final model predictions. Incorporating the suboptimal performance of the first model with the strengths of the subsequent models reduced model uncertainties and produced more accurate predictions (for the subsequent models).

We observed notably high prediction accuracies for the summer months. In particular, the July model, with an NSE of 0.96 and an R^2 of 0.97, demonstrated exceptional performance (Table 2).

Table 2. Performance of Individual SVR Models.

	Model	NSE	R^2
1	January	—	—
2	February	0.88	0.88
3	March	0.71	0.71
4	April	0.51	0.51
5	May	0.87	0.89
6	June	0.90	0.93
7	July	0.96	0.97
8	August	0.87	0.87
9	September	0.80	0.81
10	October	0.77	0.78
11	November	0.83	0.83
12	December	0.91	0.92

The test set (Figure 6b) was augmented with the predicted output of the SVR models, and a single RF model was trained on eighty-five percent of the augmented dataset.

Evaluation at the fifteen percent test wells (Figure 8) revealed much higher predictive accuracies for the sand and gravel aquifers (Table 3, Figure 9).

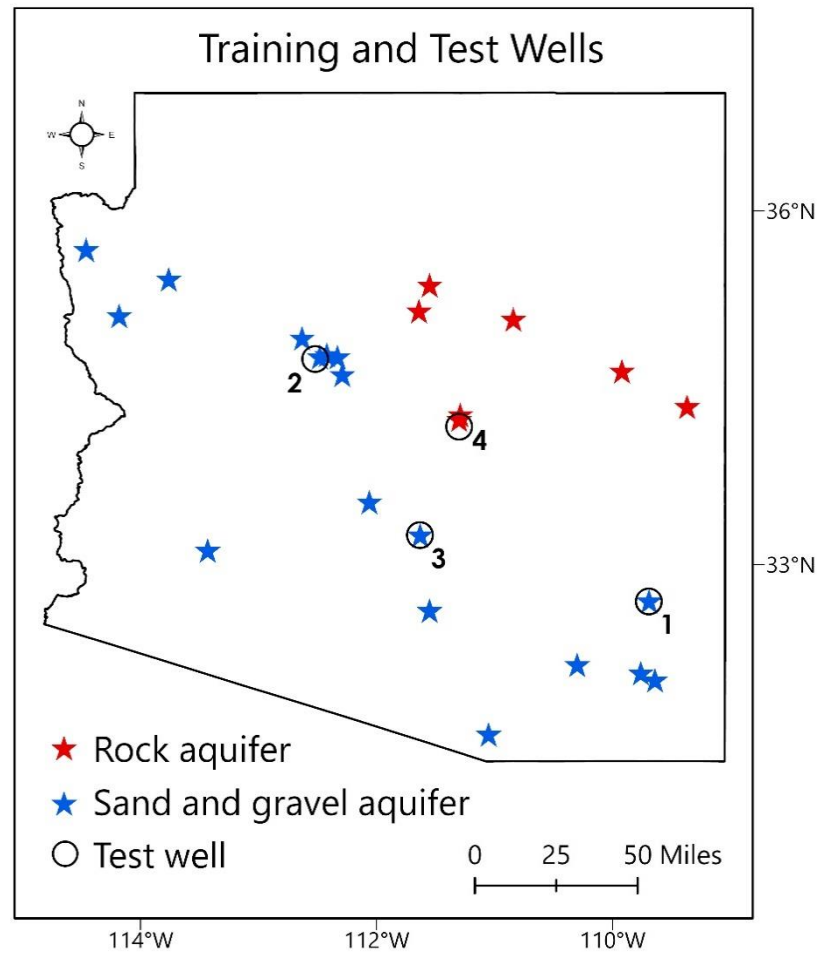


Figure 8. Map showing the locations of the four test wells for temporal evaluations after the second (85:15) train/test split.

Table 3. RF model performance evaluated at the four test wells (15 percent test split).

ID	Test Well	Aquifer Type	NSE	R ²
1	Artesia School [D-08-26 33CDC1]	Sand and gravel	0.87	0.87
2	Geiler [B-16-02 21BAA2]	Sand and gravel	0.80	0.80

3	Queen Creek [D-02-07 22BBC]	Sand and gravel	0.84	0.87
4	PE-11 [A-10-10 11ACB]	Rock	-0.13	0.28

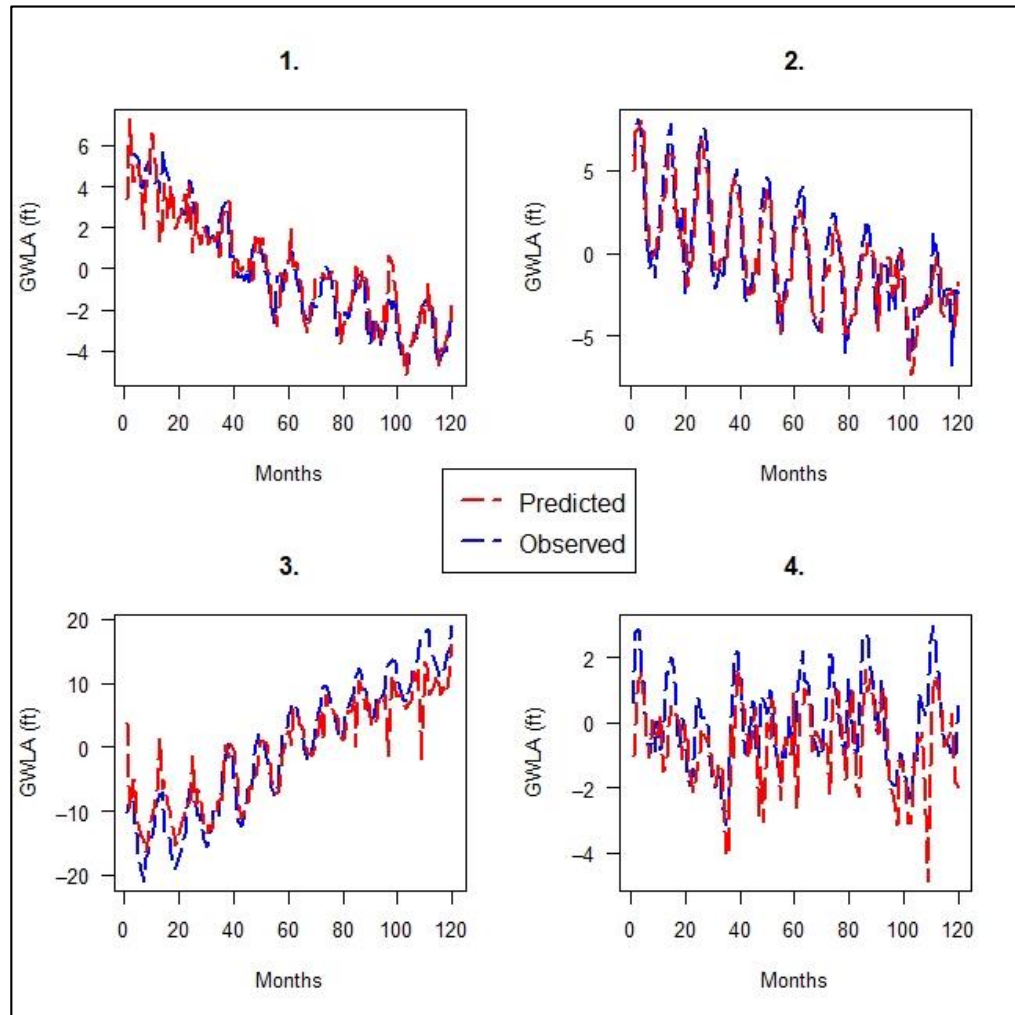


Figure 9. Plots comparing the RF-predicted GWLA at each test well with observed values.

The numbers 1–4 correspond to the four test wells shown in Figure 8 and Table 3.

Integrating EBK GWL Predictions

The prediction output of each EBK model was integrated into the dataset as an input variable, and another RF model was trained on the updated dataset without the validation

wells (Figure 10), which were excluded from the EBK process. Excluding these wells allowed us to simulate real-world scenarios better and ensure unbiased model evaluations. Subsequently, the trained RF model was evaluated at all three validation locations, and the performance was, again, relatively higher for the unconsolidated sand and gravel aquifers (Table 4).

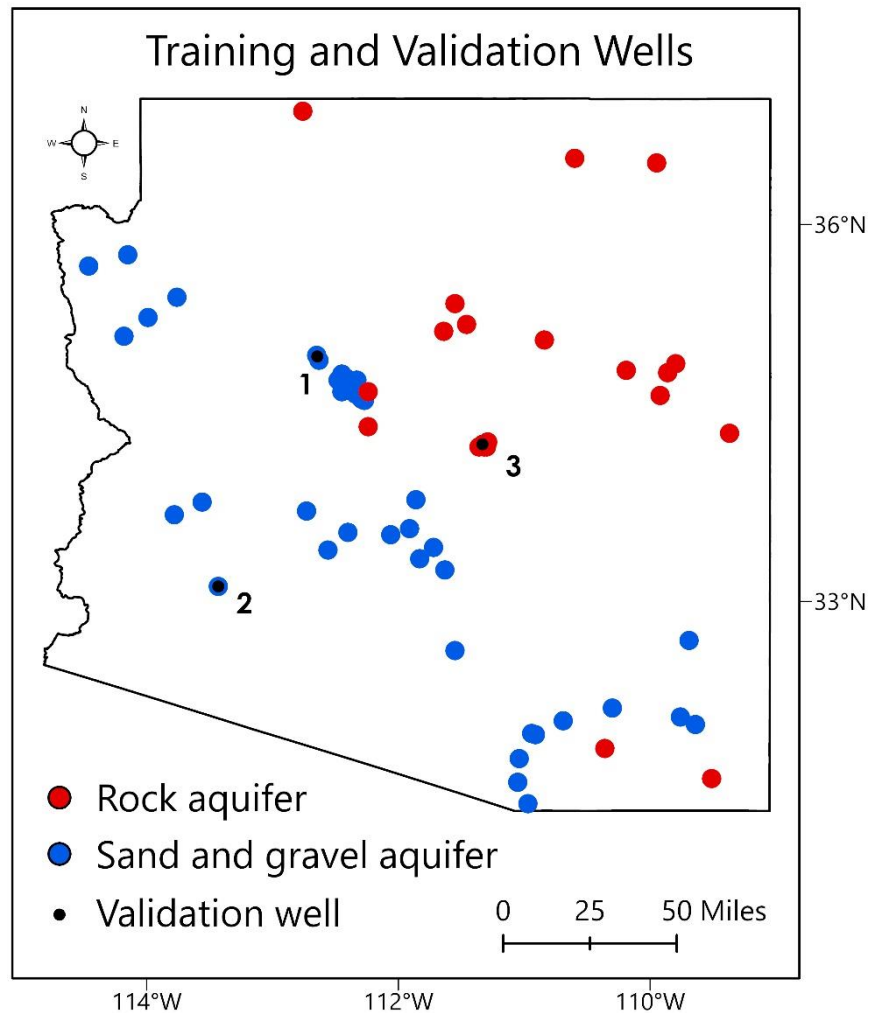


Figure 10. Location of the validation wells (1–3). The validation wells were removed from the dataset before performing the EBK to ensure unbiased model evaluation.

Table 4. RF model performance at validation wells using the EBK predictions as input.

ID	Validation Well	Aquifer Type	NSE	R ²
1	Antelope Wash [B-18-04 25AAA2]	Sand and gravel	0.81	0.83
2	Turtleback [C-03-11 31DBB]	Sand and gravel	0.63	0.65
3	Rumsey Park [A-10-10 04ABB]	Rock	0.41	0.41

The partial dependence plots (Figure 11) describe the relationships between each of the seven predictors (while keeping the others constant) and the RF-model-predicted output. Recall that the RI variable was calculated based on the precipitation, evapotranspiration, and runoff depth values (Equation (4)) and that the runoff depths were derived from precipitation (Equations (1) and (2)), so it primarily reflects changes in evapotranspiration (Figure 11). Additionally, ET is influenced by LST, which explains why the RI and LST variables exhibited similar relationships with GWL (Figure 11).

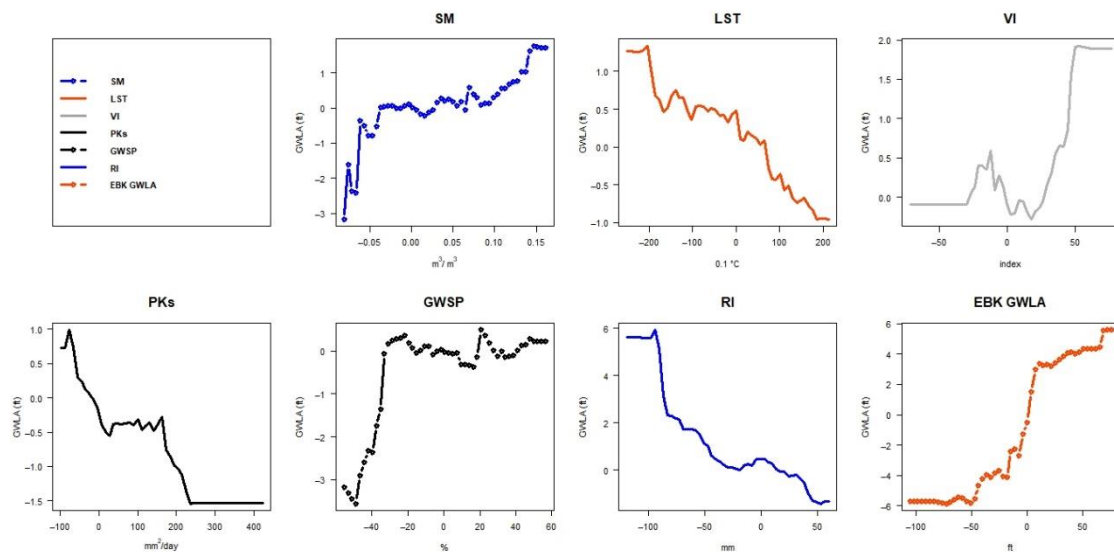


Figure 11. Partial dependence plots (PDPs) of the validation RF model. Each plot illustrates the relationship between the given predictor variable and the predicted output, showing how changes in the former influence the latter.

Following satisfactory validation performance, we repeated the EBK process using all 59 monitoring wells. As before, the interpolated GWL surfaces had a default grid size of $0.023^\circ \times 0.023^\circ$, likely determined by the geographical extent of the study area (Figure 12a). The standard errors of the prediction were also computed, with higher values representing larger prediction uncertainties (Figure 12b). The interpolated GWL surfaces were resampled to $0.125^\circ \times 0.125^\circ$, matching the grid size for this study (Figure 12c). Monthly anomalies were calculated (for each monitoring well) from the resampled surfaces and added to the dataset as a new predictor variable. A final RF model was developed using the complete augmented dataset, excluding seven wells reserved for testing (Figure 13).

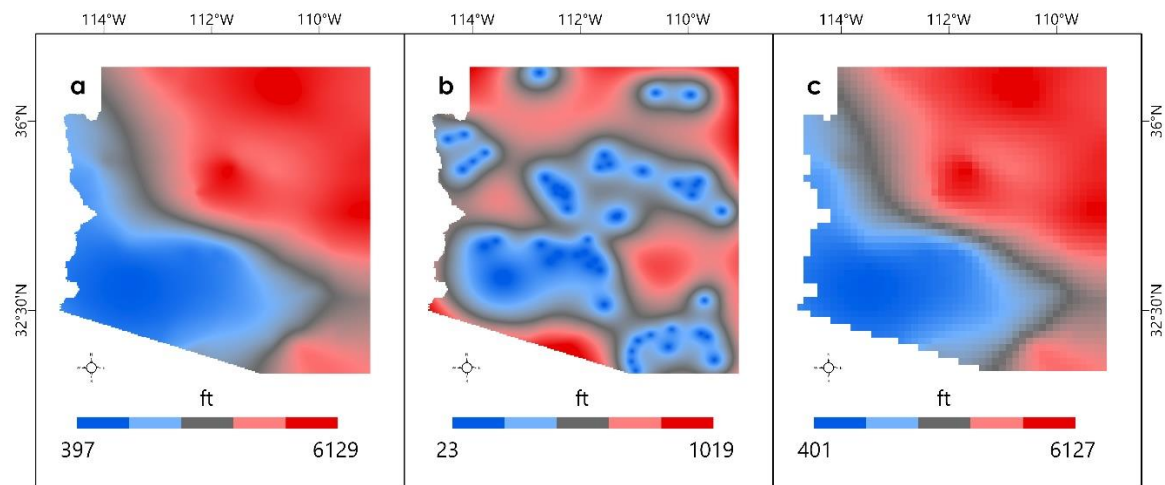


Figure 12. EBK model output for January 2010 showing (a) GWL predictions at $0.023^\circ \times 0.023^\circ$, (b) standard errors of those predictions at $0.023^\circ \times 0.023^\circ$, and (c) resampled GWL prediction surface at a $0.125^\circ \times 0.125^\circ$ grid size.

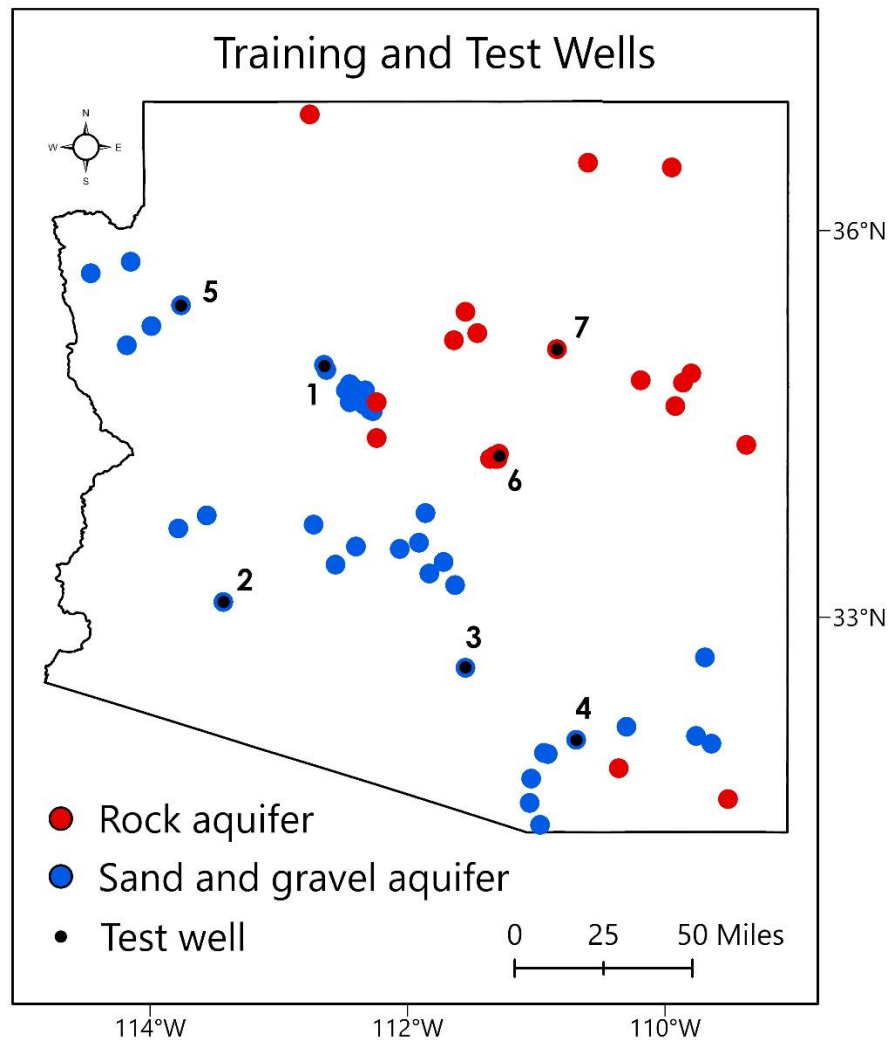


Figure 13. Spatial locations of all the monitoring wells used in this study. Wells 1–7 were used to evaluate the performance of the final RF model.

For the third time in this study, the RF model showed much better predictions at locations with monitoring wells drilled into unconsolidated sand and gravel aquifers compared to those in rock aquifers (Table 5). The average NSE and R^2 values are 0.88 and 0.92 for the former and 0.32 and 0.37 for the latter, respectively (Table 5).

Table 5. RF model performance using all monitoring wells for EBK predictions. Test wells marked with an asterisk (*) indicate that they were also used for validation.

ID	Test Well	Aquifer Type	NSE	R²
1	Antelope Wash [B-18-04 25AAA2] *	Sand and gravel	0.98	0.99
2	Turtleback [C-03-11 31DBB] *	Sand and gravel	0.86	0.88
3	Friendly Corners [D-09-08 29BCC]	Sand and gravel	0.82	0.87
4	Pantano Wash North [D-16-16 15ABD]	Sand and gravel	0.88	0.94
5	Truxton South [B-24-14 33ADA]	Sand and gravel	0.84	0.90
6	GC-3 [A-11-10 26DAB]	Rock	0.44	0.46
7	[A-19-14 03AAC1]	Rock	0.20	0.28

Following a satisfactory validation and test performance, we deployed the final RF model to make predictions at ungauged locations across the study area (Figure 14) and calculated the averages of those predictions for each $0.125^\circ \times 0.125^\circ$ grid (Figure 15).

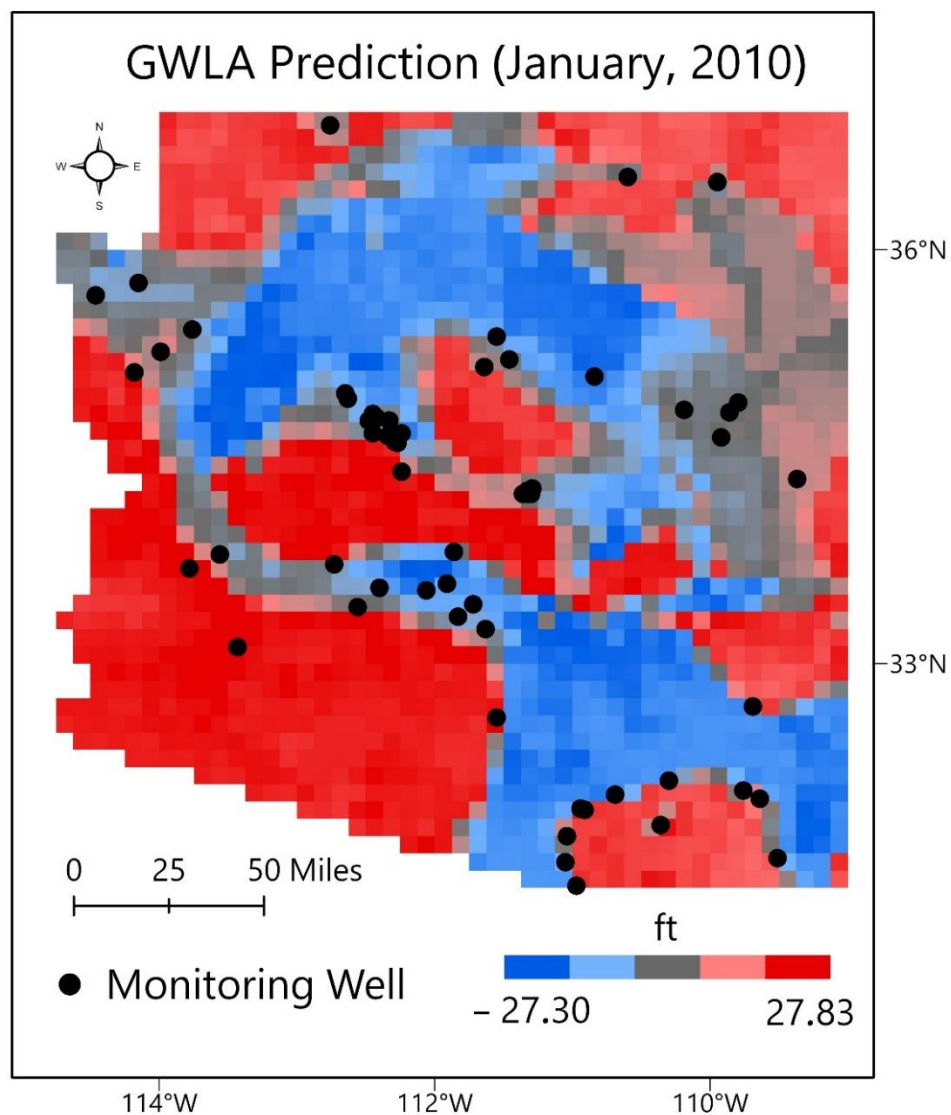


Figure 14. GWLA predictions for January 2010 after model deployment. The monitoring wells (black circles) represent both training and test wells for the final RF model.

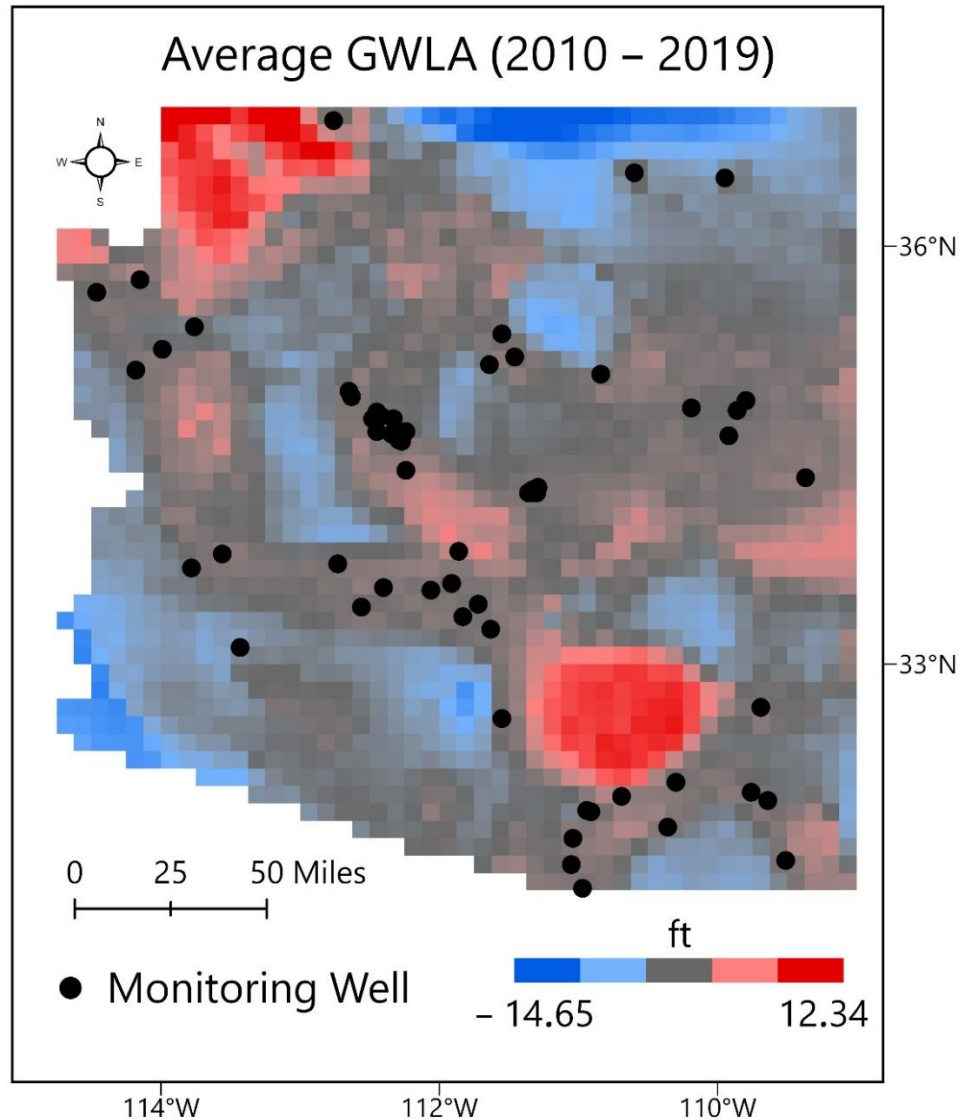


Figure 15. Average GWLA for the period of study (January 2010 to December 2019), and the locations of the monitoring wells (black circles) used in this study.

Discussion

Although ML is able to understand complex relationships between GWLs and contributing factors, this study revealed much better predictive performance for the unconsolidated material aquifers, where the relationships are generally more

straightforward. Validation wells 1 and 3 were in close proximity to other wells (Figure 10) and could have benefited from the spatial autocorrelation of GWLs between them, but well 3 did not. In fact, validation well 2 performed better than well 3, despite not having that advantage (Figure 10, Table 4). This discrepancy may have been due to several factors, including the intricate heterogeneities and geologic structures in rock aquifers. The PDP for the PK_s variable (Figure 11) also showed a distinction between the two aquifer types. This suggests that even with increased precipitation, changes in aquifer properties (such as reduced permeability in rock aquifers) can restrict groundwater flow and contribute to a declining GWL, considering all other factors at play. Both aquifers showed a negative relationship with PK_s , but GWLs were relatively higher for the unconsolidated material aquifers (Figure 11). This shows that under similar hydrological conditions and assuming all other factors are kept constant, the properties of unconsolidated aquifer materials may allow them to maintain higher GWLs compared to rock aquifers.

The discrepancy in model performance could also have been due to both model and data limitations. RF, the most employed ML algorithm for GWL prediction [10], failed to effectively capture GWL trends in a dolomite rock aquifer in a semi-arid region [73] and was also outperformed in an aquifer with fractured hydrogeology in a similar climate [74], as shown by two separate studies. In both cases, deep learning (DL) models demonstrated the best performance. DL is a branch of ML that is based on the concept of deep neural networks and is especially known to outperform traditional (shallow) ML techniques in applications involving large amounts of data [75,76] and high dimensionality [75]. But drilling a large number of wells into hard rock formations in dry regions might not be practical for various reasons, including cost and limited water availability [77], which can

pose significant challenges for ML-based GWL predictions in rock aquifers in arid regions. Specifically, fewer wells were drilled into the consolidated rocks in the study area compared to the basin fill unconsolidated materials [29], which potentially limited the adequate representation of the geologic complexities of these rock aquifers. However, the relatively weak performance of the RF models for the rock aquifers presents opportunities for future research to investigate and enhance the model performance in such complex geologic and climate settings. These efforts should begin with acquiring the maximum amount of good-quality data for a comprehensive analysis.

The prediction errors from the EBK models (Figure 12b) underscore the importance of the spatial density and distribution of the kriging data in ensuring the reliability of predictions. The largest uncertainties were seen around the boundaries of the study area, and the predictions in the vicinity of the monitoring wells were relatively more accurate. Two of the wells excluded in the validation phase (Figure 10) were included in the final model as test wells (Figure 13) and showed improved predictions (Table 4, Table 5), further underscoring the importance of data density and quality in kriging.

Based on the percent increases in the mean squared error (MSE) when important predictor variables are left out, the RI and EBK variables were the most important in the validation RF model, where both variables showed similar levels of importance. However, in the final deployment model, EBK was the most important variable (by a significant margin). This suggests that incorporating spatial interpolation techniques such as EBK can substantially enhance the performance of ML models. Although the kriging process can be tedious and challenging, the model improvements they offer make these efforts worth it.

As shown, the average GWLA for the period of study was predominantly negative. In fact, only about twenty-eight percent of the study area showed a positive average anomaly during this period. This trend reflects the challenges in a dry/arid region with high groundwater demand and withdrawal rates possibly exceeding natural recharge. Historically, groundwater in Arizona has been pumped out faster than it has been replenished by natural means [78,79], resulting in overdraft in many agricultural and urban areas [79]. But as the quest to exploit deeper aquifers continues, the costs of drilling to these depths are much higher, as are the energy costs of pumping water from them [79]. This study therefore could be useful for optimizing the drilling process by identifying locations for new wells and increasing the likelihood of accessing groundwater at optimal depths. Efforts to manage the groundwater overdraft issue began with identifying regions with a high reliance on groundwater (known as Active Management Areas (AMAs)) and subsequently empowering the ADWR to monitor compliance within the AMAs with the regulatory frameworks in place [79,80]. Within these AMAs and beyond, this study can also aid in the monitoring and allocation of groundwater resources by identifying groundwater-deficient areas based on the average predicted GWLA values, and offer data-driven support and recommendations towards the effective management of groundwater resources in vulnerable areas.

Conclusion

Groundwater is the largest reservoir of available freshwater in the world and a critically important resource. Its global relevance is amplified by the direct impacts of climate change on surface water sources, particularly in arid regions. In this study, we

demonstrated the effectiveness of ML in predicting monthly GWLAs when combined with reliable spatial interpolation models and developed the first statewide GWLA prediction model for the state of Arizona. Following satisfactory performance based on average NSE/R² values of 0.62/0.63 and 0.72/0.76 during the validation and testing phases, respectively, monthly GWLA rasters were produced for January 2010 to December 2019. Moving forward, future studies may focus on addressing some of the challenges of applying traditional ML techniques to rock aquifers in dry regions discussed in this study, in terms of leveraging the available data and reducing prediction uncertainties in such complex settings.

With well depths ranging from 25 to 1600 feet, this study demonstrated effectiveness for both shallow and deep aquifers. The model design utilized remotely sensed datasets from satellites with global coverage, enabling replicability for similar climates across the globe. Our remote sensing approach ensures that data-sparse regions of the world, where field-based hydrological variables are limited or largely inaccessible, are not left out. It is our hope that this work contributes substantially to the science of monitoring groundwater resources in the face of global warming and climate change threats, ensuring the availability of groundwater to meet domestic, agricultural, and industrial water needs.

Abbreviations

The list of abbreviations and acronyms used in this paper are tabulated below:

ADWR	Arizona Department of Water Resources
AMA	Active Management Area

ANN	Artificial neural network
ARIMA	Autoregressive integrated moving average
CONUS	Contiguous United States
CK	Classical kriging
CN	Curve number
CV	Coefficient of variation
DL	Deep learning
DT	Decision tree
EBK	Empirical Bayesian kriging
ET	Evapotranspiration
GBM	Gradient boosting mechanism
GI	Geostatistical interpolation
GRACE	Gravity Recovery and Climate Experiment
GWL	Groundwater level
GWLA	Groundwater level anomaly
GWSP	Groundwater storage percentile
KNN	K-nearest neighbors
K _s	Soil saturated hydraulic conductivity
LST	Land surface temperature
MAE	Mean absolute error
MI	Mean imputation
ML	Machine learning

MLR	Multilinear regression
MSE	Mean squared error
NGWMN	National Groundwater Monitoring Network
P	Precipitation
PDP	Partial dependence plot
PKs	Precipitation x soil saturated hydraulic conductivity
R	Runoff depth
RBF	Radial basis function
RBF-NN	Radial basis function neural network
RBF-SVR	Radial basis function support vector regression
RF	Random forest
RI	Recharge index
RMSE	Root mean square error
SM	Soil moisture
SVM	Support vector machine
SVR	Support vector machine for regression
USGS	U.S. Geological Survey
VI	Vegetation index
WCI	Water cycle intensity

References

1. Scanlon, B.R.; Keese, K.E.; Flint, A.L.; Flint, L.E.; Gaye, C.B.; Edmunds, W.M.; Simmers, I. Global synthesis of groundwater recharge in semiarid and arid regions. *Hydrol. Process.* **2006**, *20*, 3335–3370.
2. Dawoud, M.A. Groundwater economics in arid regions: Abu Dhabi Emirate case study. In *Recent Advances in Environmental Science from the Euro-Mediterranean and Surrounding Regions*, Proceedings of Euro-Mediterranean Conference for Environmental Integration (EMCEI-1), Sousse, Tunisia, 20–25 November, 2017; Springer: Cham; pp. 611–613.
3. Liu, Q.; Gui, D.; Zhang, L.; Niu, J.; Dai, H.; Wei, G.; Hu, B.X. Simulation of regional groundwater levels in arid regions using interpretable machine learning models. *Sci. Total Environ.* **2022**, *831*, 154902.
4. Haider, A.; Lee, G.; Jafri, T.H.; Yoon, P.; Piao, J.; Jhang, K. Enhancing Accuracy of Groundwater Level Forecasting with Minimal Computational Complexity Using Temporal Convolutional Network. *Water* **2023**, *15*, 4041.
5. Tao, H.; Hameed, M.M.; Marhoon, H.A.; Zounemat-Kermani, M.; Heddami, S.; Kim, S.; Sulaiman, S.O.; Tan, M.L.; Sa'adi, Z.; Mehr, A.D.; et al. Groundwater level prediction using machine learning models: A comprehensive review. *Neurocomputing* **2022**, *489*, 271–308.
6. Ardana, P.D.H.; Redana, I.W.; Yekti, M.I.; Simpen, I.N. Groundwater Level Forecasting Using Multiple Linear Regression and Artificial Neural Network Approaches. *Civ. Eng. Archit.* **2022**, *10*, 784–799.
<https://doi.org/10.13189/cea.2022.100304>.

7. Najafabadipour, A.; Kamali, G.; Nezamabadi-Pour, H. Application of Artificial Intelligence Techniques for the Determination of Groundwater Level Using Spatio-Temporal Parameters. *ACS Omega* **2022**, *7*, 10751–10764.
8. Sahoo, M.; Kasot, A.; Dhar, A.; Kar, A. On predictability of groundwater level in shallow wells using satellite observations. *Water Resour. Manag.* **2018**, *32*, 1225–1244.
9. Li, B.; Yang, G.; Wan, R.; Dai, X.; Zhang, Y. Comparison of random forests and other statistical methods for the prediction of lake water level: A case study of the Poyang Lake in China. *Hydrol. Res.* **2016**, *47* (Suppl. S1), 69–83.
10. Afrifa, S.; Zhang, T.; Appiahene, P.; Varadarajan, V. Mathematical and machine learning models for groundwater level changes: A systematic review and bibliographic analysis. *Future Internet* **2022**, *14*, 259.
11. Gonzalez, R.Q.; Arsanjani, J.J. Prediction of groundwater level variations in a changing climate: A Danish case study. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 792.
12. Guzman, S.M.; Paz, J.O.; Tagert, M.L.M.; Mercer, A.E. Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines. *Environ. Model. Assess.* **2019**, *24*, 223–234.
13. Yoon, H.; Hyun, Y.; Ha, K.; Lee, K.K.; Kim, G.B. A method to improve the stability and accuracy of ANN-and SVM-based time series models for long-term groundwater level predictions. *Comput. Geosci.* **2016**, *90*, 144–155.
14. Behzad, M.; Asghari, K.; Coppola, E.A., Jr. Comparative study of SVMs and ANNs in aquifer water level prediction. *J. Comput. Civ. Eng.* **2010**, *24*, 408–413.

15. Nie, S.; Bian, J.; Wan, H.; Sun, X.; Zhang, B. Simulation and uncertainty analysis for groundwater levels using radial basis function neural network and support vector machine models. *Res. Technol.—AQUA* **2017**, *66*, 15–24.
16. Tapak, L.; Rahmani, A.R.; Moghimbeigi, A. Prediction the groundwater level of Hamadan-Bahar plain, west of Iran using support vector machines. *J. Res. Health Sci.* **2013**, *14*, 82–87.
17. Tiwari, V.; Verma, M. Prediction of Groundwater Level Using Advance Machine Learning Techniques. In Proceedings of the 3rd IEEE International Conference on Intelligent Technologies (CONIT), Hubli, India, 23–25 June 2023; pp. 1–6.
18. Hikouei, I.S.; Eshleman, K.N.; Saharjo, B.H.; Graham, L.L.; Applegate, G.; Cochrane, M.A. Using machine learning algorithms to predict groundwater levels in Indonesian tropical peatlands. *Sci. Total Environ.* **2023**, *857*, 159701.
19. Kanyama, Y.; Ajoodha, R.; Seyler, H.; Makondo, N.; Tutu, H. Application of machine learning techniques in forecasting groundwater levels in the grootfontein aquifer. In Proceedings of the 2nd IEEE International Multidisciplinary Information Technology and Engineering Conference (IMITEC), Kimberley, South Africa, 25–27 November 2020; pp. 1–8.
20. Alam, M.J.; Kar, S.; Zaman, S.; Ahamed, S.; Samiya, K. Forecasting Underground Water Levels: LSTM Based Model Outperforms GRU and Decision Tree Based Models. In Proceedings of the IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Naya Raipur, India, 30–31 December 2022; pp. 280–283.

21. Biernacik, P.; Kazimierski, W.; Włodarczyk-Sielicka, M. Comparative Analysis of Selected Geostatistical Methods for Bottom Surface Modeling. *Sensors* **2023**, *23*, 3941.
22. Manda, S.; Patil, A. Analysis of groundwater level differences in Ganges basin using geostatistical modeling. *Int. J. Agric. Eng.* **2018**, *11*, 392–396.
23. Krivoruchko, K. Empirical bayesian kriging. *ArcUser Fall* **2012**, *6*, 1145.
24. Bouhout, S.; Haboubi, K.; Zian, A.; Elyoubi, M.S.; Elabdouni, A. Evaluation of two linear kriging methods for piezometric levels interpolation and a framework for upgrading groundwater level monitoring network in Ghiss-Nekor plain, north-eastern Morocco. *Arab. J. Geosci.* **2022**, *15*, 1016.
25. Hussain, M.M.; Bari, S.H.; Tarif, M.E.; Rahman, M.T.U.; Hoque, M.A. Temporal and spatial variation of groundwater level in Mymensingh district, Bangladesh. *Int. J. Hydrol. Sci. Technol.* **2016**, *6*, 188–197.
26. Deshmukh, M.M.; Elbeltagi, A.; Kouadri, S. Climate Change Impact on Groundwater Resources in Semi-arid Regions. In *Climate Change Impact on Groundwater Resources: Human Health Risk Assessment in Arid and Semi-arid Regions*; Panneerselvam, B., Pande, C.B., Muniraj, K., Balasubramanian, A., Ravichandran, N., Eds.; Springer: Cham, Switzerland, 2022; pp. 9–23.
27. Zowam, F.J.; Milewski, A.M.; Richards IV, D.F. A Satellite-Based Approach for Quantifying Terrestrial Water Cycle Intensity. *Remote Sens.* **2023**, *15*, 3632.
28. McCafferty, A.E.; San Juan, C.A.; Lawley, C.J.M.; Graham, G.E.; Gadd, M.G.; Huston, D.L.; Kelley, K.D.; Paradis, S.; Peter, J.M.; Czarnota, K. National-Scale Geophysical, Geologic, and Mineral Resource Data and Grids for the United States,

- Canada, and Australia: Data in Support of the Tri-National Critical Minerals Mapping Initiative: US Geological Survey Data Release. Available online: <https://www.sciencebase.gov/catalog/item/623a013ed34e915b67cddcfa> (accessed on 22 February 2024).
29. Bedinger, M.S.; Anderson, T.W.; Langer, W.H. Groundwater Units and Withdrawal, Basin and Range Province, Arizona. Water-Resources Investigations Report (No. 83-4114-A). Available online: <https://doi.org/10.3133/wri834114A> (accessed on 19 September 2024).
 30. Tillman, F.D.; Flynn, M.E. Arizona Groundwater Explorer: Interactive maps for evaluating the historical and current groundwater conditions in wells in Arizona, USA. *Hydrogeol. J.* **2024**, *32*, 645–661.
 31. Bertsimas, D.; Pawlowski, C.; Zhuo, Y.D. From predictive methods to missing data imputation: An optimization approach. *J. Mach. Learn. Res.* **2018**, *18*, 1–39.
 32. Petrazzini, B.O.; Naya, H.; Lopez-Bello, F.; Vazquez, G.; Spangenberg, L. Evaluation of different approaches for missing data imputation on features associated to genomic data. *BioData Min.* **2021**, *14*, 1–13.
 33. Muharemi, F.; Logofătu, D.; Leon, F. Review on general techniques and packages for data imputation in R on a real world dataset. In *Computational Collective Intelligence, Proceedings of the 10th International Conference, Bristol, UK, 5–7 September 2018*; Springer International Publishing: Cham, Switzerland, 2018; Part II 10, pp. 386–395.
 34. Lande, R. On comparing coefficients of variation. *Syst. Zool.* **1977**, *26*, 214–217.

35. Seyoum, W.M.; Kwon, D.; Milewski, A.M. Downscaling GRACE TWSA data into high-resolution groundwater level anomaly using machine learning-based models in a glacial aquifer system. *Remote Sens.* **2019**, *11*, 824.
36. Moeck, C.; Grech-Cumbo, N.; Podgorski, J.; Bretzler, A.; Gurdak, J.J.; Berg, M.; Schirmer, M. A global-scale dataset of direct natural groundwater recharge rates: A review of variables, processes and relationships. *Sci. Total Environ.* **2020**, 717, 137042.
37. Pradhan, R.K.; Markonis, Y.; Godoy, M.R.V.; Villalba-Pradas, A.; Andreadis, K.M.; Nikolopoulos, E.I.; Papalexiou, S.M.; Rahim, A.; Tapiador, F.J.; Hanel, M. Review of GPM IMERG performance: A global perspective. *Remote Sens. Environ.* **2022**, 268, 112754.
38. Milewski, A.; Elkadiri, R.; Durham, M. Assessment and comparison of TMPA satellite precipitation products in varying climatic and topographic regimes in Morocco. *Remote Sens.* **2015**, *7*, 5697–5717.
39. Mohammed, S.A.; Hamouda, M.A.; Mahmoud, M.T.; Mohamed, M.M. Performance of GPM-IMERG precipitation products under diverse topographical features and multiple-intensity rainfall in an arid region. *Hydrol. Earth Syst. Sci.* **2020**, 2020, 1–27.
40. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Xie, P.; Yoo, S.H. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). *Algorithm Theor. Basis Doc. (ATBD)* **2015**, *4*, 30.
41. Malik, M.S.; Shukla, J.P.; Mishra, S. Effect of groundwater level on soil moisture, soil temperature and surface temperature. *J. Indian Soc. Remote Sens.* **2021**, *49*, 2143–2161.

42. Maihemuti, B.; Simayi, Z.; Alifujiang, Y.; Aishan, T.; Abliz, A.; Aierken, G. Development and evaluation of the soil water balance model in an inland arid delta oasis: Implications for sustainable groundwater resource management. *Glob. Ecol. Conserv.* **2021**, *25*, e01408.
43. Otoko, G.R. Mathematical Relationship between Soil Moisture and Groundwater Level in A Loamy Sand Soil in The Niger Delta Region of Nigeria. *Int. J. Adv. Res. Sci. Eng. Technol.* **2014**, *5*, 1–8.
44. Hongtao, J.; Huanfeng, S.; Xinghua, L.; Lili, L. The 43-Year (1978–2020) Global 9 km Remotely Sensed Soil Moisture Product: PANGAEA. Available online: <https://doi.org/10.1594/PANGAEA.940409> (accessed on 12 February 2023).
45. Elnashar, A.; Wang, L.; Wu, B.; Zhu, W.; Zeng, H. Synthesis of global actual evapotranspiration from 1982 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 447–480.
46. Zhang, T.; Zhou, Y.; Zhu, Z.; Li, X.; Asrar, G.R. A global seamless 1 km resolution daily land surface temperature dataset (2003–2020). *Earth Syst. Sci. Data* **2021**, *2021*, 1–16.
47. Grove, M.; Harbor, J.; Engel, B. composite vs. distributed curve numbers: Effects on estimates of storm runoff depths. *J. Am. Water Resour. Assoc.* **1998**, *34*, 1015–1023.
48. Arisanty, D.; Rahmawati, N.; Rosadi, D. Soil Physical Characteristics and Saturated Hydraulic Conductivity in the Landform of Barito Delta, Kalimantan, Indonesia. *Appl. Environ. Soil Sci.* **2022**, *2022*, 9118461.
49. Gupta, S.; Lehmann, P.; Bonetti, S.; Papritz, A.; Or, D. Global prediction of soil saturated hydraulic conductivity using random forest in a covariate-based geoTransfer function (CoGTF) framework. *J. Adv. Model. Earth Syst.* **2021**, *13*, e2020MS002242.

50. Beaudoin, H.; Rodell, M.; Getirana, A.; Li, B. *Groundwater and Soil Moisture Conditions from GRACE and GRACE-FO Data Assimilation L4 7-Days 0.125×0.125 Degree U.S. V4.0*; Goddard Earth Sciences Data and Information Services Center (GES DISC): Greenbelt, MD, USA, 2021. Available online: https://disc.gsfc.nasa.gov/datasets/GRACEDADM_CLSM0125US_7D_4.0/summary (accessed on 30 October 2023).
51. Dhungel, R.; Fiedler, F. Water balance to recharge calculation: Implications for watershed management using systems dynamics approach. *Hydrol.* **2016**, *3*, 13.
52. Ahmadi, A.; Olyaei, M.; Heydari, Z.; Emami, M.; Zeynolabedin, A.; Ghomlaghi, A.; Daccache, A.; Fogg, G.E.; Sadegh, M. Groundwater level modeling with machine learning: A systematic review and meta-analysis. *Water* **2022**, *14*, 949.
53. Ding, J.; Tarokh, V.; Yang, Y. Model selection techniques: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 16–34.
54. Vapnik, V.N. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **1999**, *10*, 988–999.
55. Awad, M.; Khanna, R. Support Vector Machines for Classification. In *Efficient Learning Machines*; Apress: Berkeley, CA, USA, 2015; p. 268.
56. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.; Vapnik, A. Support vector regression machines. *Adv. Neural. Inf. Process Syst.* **1996**, *9*, 161–226.
57. Amirkhalili, Y.S.; Aghsami, A.; Jolai, F. Comparison of Time Series ARIMA Model and Support Vector Regression. *Int. J. Hybrid Inf. Technol.* **2020**, *13*, 7–18.

58. Ranković, V.; Grujović, N.; Divac, D.; Milivojević, N. Development of support vector regression identification model for prediction of dam structural behaviour. *Struct. Saf.* **2014**, *48*, 33–39.
59. Ayodeji, A.; Liu, Y.K. SVR optimization with soft computing algorithms for incipient SGTR diagnosis. *Ann. Nucl. Energy* **2018**, *121*, 89–100.
60. Açıkkar, M.; Altunkol, Y. A novel hybrid PSO-and GS-based hyperparameter optimization algorithm for support vector regression. *Neural Comput.* **2023**, *35*, 19961–19977.
61. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.
62. Al-Anazi, A.F.; Gates, I.D. Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study. *Comput. Geosci.* **2010**, *36*, 1494–1503.
63. Kavousi-Fard, A.; Samet, H.; Marzbani, F. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert Syst. Appl.* **2014**, *41*, 6047–6056.
64. Karatzoglou, A.; Meyer, D.; Hornik, K. Support vector machines in R. *J. Stat. Softw.* **2006**, *15*, 1–28.
65. Kaneko, H.; Funatsu, K. Fast optimization of hyperparameters for support vector regression models with highly predictive ability. *Chemom. Intell. Lab. Syst.* **2015**, *142*, 64–69.
66. Tsirikoglou, P.; Abraham, S.; Contino, F.; Lacor, C.; Ghorbaniasl, G. A hyperparameters selection technique for support vector regression models. *Appl. Soft Comput.* **2017**, *61*, 139–148.

67. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
68. Wu, D.J.; Feng, T.; Naehrig, M.; Lauter, K. Privately evaluating decision trees and random forests. *Proc. Priv. Enh. Technol.* **2015**, *2016*, 335–355.
69. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
70. Rigatti, S.J. Random forest. *J. Insur. Med.* **2017**, *47*, 31–39.
71. Maliva, R. G. (2016). Geostatistical methods and applications. In *Aquifer Characterization Techniques*, Springer: Cham, 2016; No. 4, 595-617.
72. Li, Y.; Hernandez, J.H.; Aviles, M.; Knappett, P.S.; Giardino, J.R.; Miranda, R.; Puy, M.J.; Padilla, F. Morales, J. Empirical Bayesian Kriging method to evaluate inter-annual water-table evolution in the Cuenca Alta del Río Laja aquifer, Guanajuato, México. *J. Hydrol.* **2020**, *582*, p.124517.
73. Kanyama, Y. Application of Machine Learning Techniques in Predicting Groundwater Levels and Discharge Rates in the Northwest Aquifers. Master's Thesis, The University of the Witwatersrand, Johannesburg, South Africa, 22 April 2021.
74. Yin, W.; Fan, Z.; Tangdamrongsub, N.; Hu, L.; Zhang, M. Comparison of physical and data-driven models to forecast groundwater level changes with the inclusion of GRACE—A case study over the state of Victoria, Australia. *J. Hydrol.* **2021**, *602*, 126735.
75. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695.
76. Xu, Y.; Zhou, Y.; Sekula, P.; Ding, L. Machine learning in construction: From shallow to deep learning. *Dev. Built Environ.* **2021**, *6*, 100045.

77. Sk, M.; Ramanujam, N.; Champoil, V.; Biswas, S.K.; Rasool, Q.A.; Ojha, C. Identification of groundwater in hard rock terrain using 2D electrical resistivity tomography imaging technique: Securing water scarcity at the time of seasonal rainfall failure, South Andaman. *Int. J. Geosci.* **2018**, *9*, 59.
78. Arizona Department of Water Resources. Overview of the Arizona Groundwater Management Code. Available online: https://www.azwater.gov/sites/default/files/media/Arizona%20Groundwater_Code_1.pdf (accessed on 14 June 2024).
79. Hirt, P.; Snyder, R.; Hester, C.; Larson, K. Water consumption and sustainability in Arizona: A tale of two desert cities. *J. Southwest* **2017**, *59*, 264–301.
80. Megdal, S.B. Arizona groundwater management. *Water Rep.* **2012**, 104(9).

CHAPTER 4

CLIMATE VARIABILITY AND GROUNDWATER LEVELS: A CORRELATION
AND CAUSATION ANALYSIS ¹

¹ Zowam, F.J. and Milewski, A.M. To be submitted to the *Water* journal.

Abstract

Climate change (CC) is a global menace facing our planet today, where relatively short-term fluctuations in climate patterns (climate variability) indicate broader CC trends over longer timescales. Considering terrestrial water cycle intensity (WCI) and groundwater level (GWL) as indicators of CC and groundwater (GW) availability respectively, this study explored the dynamic relationship between WCI and GWL anomalies (WCIA and GWLA, respectively) in an arid region, based on an innovative approach to statistical correlation and causation analysis. Pearson correlation (r) assessed the strength and direction of a contemporaneous linear relationship between both variables, a cross-correlation function (ccf) determined the dynamic nature of those relationships considering monthly lags up to a predetermined maximum of 12 months, and Granger causality tests assessed the statistical significance of the lead variable for predicting the lagged variable. A contemporaneous linear relationship between both variables was mostly absent but appeared at various lags. At these lags, the strongest correlations were dominantly negative, with GWLA leading WCIA, as supported by the GC tests. This trend implies that the intensification of the water cycle reflects a decline in past GWLs in the affected areas, necessitating immediate water management actions.

Introduction

Climate change (CC), according to the Intergovernmental Panel on Climate Change (IPCC), is the statistically significant variation in average weather conditions lasting for decades or longer [1]. It is a global menace today, driven largely by human activities that release greenhouse gases into the atmosphere, altering global temperature and precipitation patterns [1–3]. Climate variability (CV), on the other hand, refers to short-term fluctuations associated with CC [4]. Initially, it described the fluctuations due to natural processes, but is now understood to reflect the anthropogenic influences that define CC [4,5], and is therefore sometimes referred to as CC [5].

The effects of CC vary among different groups of people. For instance: the health industry is faced with issues such as heat stress [6], injuries from extreme weather events [7,8], anxiety/depression and post-traumatic stress disorders associated with climate-induced disasters [6,9], behavioral disorders in elderly people and people with mental illnesses [9], and so on. Meteorological concerns include more frequent and intense extreme weather events such as floods, droughts, and hurricanes [2,8,10–12]. The severity and frequency of wildfires are also amplified [13,14]. In terms of agriculture, a positive, self-reinforcing feedback loop is likely because farmers not only contribute to CC but are also affected by it, and any attempts to compensate for these impacts exacerbate the effects on the climate [15].

However, hydrogeologists are concerned about the impacts on groundwater (GW) –water found in soils and rocks beneath the ground. GW is the largest reservoir of available freshwater in the world [16–19] and a critically important resource. It is controlled by climate conditions and is expected to be impacted by CC [1,19,20]. These impacts could

be direct or indirect, where direct impacts generally affect its quantity and indirect impacts affect quality. Specifically, direct impacts involve the natural replenishments or recharge and directly affect the availability of the resource [19]. Indirect impacts may include alterations in GW chemistry [21], increased mobility of geogenic contaminants [22], and leaching of water-soluble contaminants [23] following intense and frequent precipitation events. Additionally, during droughts, and as a result of altered redox conditions of aquifers, environmental pollutants such as nitrates and sulphates may be mobilized [22]. In coastal areas, these impacts include seawater intrusion due to reduced precipitation and GW recharge [1,24], and seawater flooding due to increased frequency and intensity of large coastal storms causing sea levels to rise [25]. In general, CC is causing alterations to key aspects of GW quality, through variations in concentrations of organic and inorganic compounds, dissolved oxygen levels, salinity, and pH [26].

Effectively managing GW resources now requires the inclusion of CC and variability impact assessments, which have previously been overlooked [27,28]. Scientists at the U. S Geological Survey (USGS) work with various tiers of national partners and international collaborators to understand how these factors impact groundwater availability in the United States [27]. However, because the variations in climate patterns would mean that broad generalizations may not adequately capture these interactions, there is a need for more localized studies that provide specific insights into how these variations impact GW. At the same time, arid regions are environmentally fragile and highly sensitive to global CC and variability, making CC assessments in such regions a hot topic in the climate science field [29]. Given this context, this study assesses local CC – GW interactions in an

arid region in the contiguous United States, with a focus on impacts related specifically to GW availability.

In examining these interactions, it is important to consider the following: (1) natural GW availability depends largely on precipitation, which is a major component of the global water cycle i.e. a dynamic system describing the continuous movement and exchange of water between the Earth's surface and atmosphere. (2) CC has significant direct impacts on this system [30,31]. With extreme weather becoming even more extreme, the disparity in precipitation between wet and dry areas is expected to intensify [26]. (3) These changes or shifts in the water cycle over any landscape or terrain can be quantified using the terrestrial water cycle intensity (WCI) metric [32–34], where WCI is defined as the sum of precipitation (P) and evapotranspiration (ET) averaged across a landscape unit over a specified time step [33], and P and ET are important components of the global water cycle.

Using satellite-based estimates of P and ET for the WCI computation, along with groundwater level (GWL) measurements from 59 monitoring wells, we explored the dynamic relationship between WCI and GWL anomalies at local scales across the predominantly arid U. S State of Arizona, employing an innovative approach to statistical correlation and causation analyses. We hypothesize that: (1) Monthly WCI and GWL anomalies over Arizona between January 2010 and December 2019 show a strong, negative correlation. (2) There is a strong lead-lag relationship between both variables, where the hypothesized lead variable (WCI anomaly) improves the prediction of the lag variable (GWL anomaly).

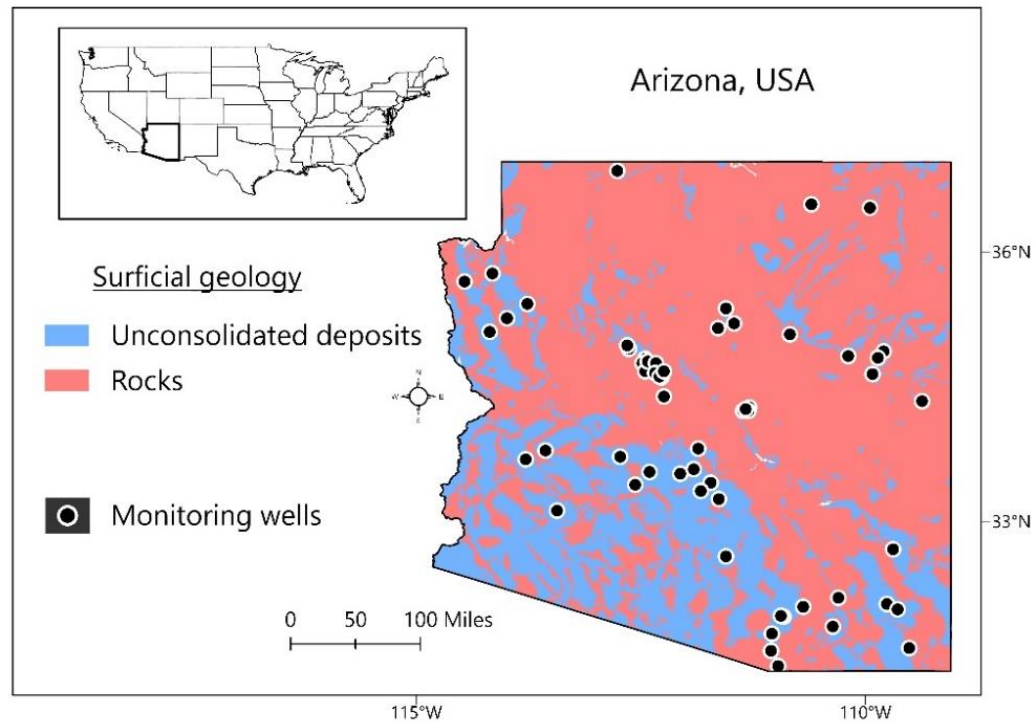


Figure 1. Study Area showing groundwater monitoring wells and surficial geology. Modified after: Zowam and Milewski (2024) [35].

Materials and Methods

While there is currently no universally accepted WCI indicator over land [33], Huntington et al. (2018) [33] introduced a terrestrial WCI framework that represents landscape processes more effectively than previous efforts and can be applied across all spatial and temporal scales of interest. The authors defined the terrestrial WCI as the sum of P and ET over a specific landscape unit and time interval, and calculated the WCI for the contiguous United States (CONUS) between 1945 and 2014, using ground-based P and ET measurements. In a recent article, Zowam et al. (2023) [34] validated the use of remote sensing for the WCI computation and extended the calculations for the CONUS to cover a

more recent period (2001 to 2019). In this study, a similar remote-sensing WCI approach is applied for a much smaller landscape area (Arizona) and period (2010 to 2019).

We utilized the final processing run of the Global Precipitation Measurement (GPM) dataset generated using the Integrated Multi-Satellite Retrievals for the GPM (IMERG) algorithm. IMERG is considered an excellent alternative to ground-based observations [36], and is particularly effective over semi-arid and arid regions [37,38]. The monthly $0.1^\circ \times 0.1^\circ$ resolution dataset was downloaded from the NASA data portal (<https://gpm.nasa.gov/data/directory>, accessed on 11 May 2021).

For ET, a finer grid resolution ($0.01^\circ \times 0.01^\circ$) synthesizing various global satellite ET measurement efforts was downloaded from the Harvard Dataverse repository (<https://doi.org/10.7910/DVN/ZGOUED>, accessed on 3 July 2021). The synthesized product outperformed local ET products in the United States, China, and the continent of Africa [39]. Monthly GWL data was obtained by averaging daily observations from 59 monitoring wells downloaded from the National Groundwater Monitoring Network (NGWMN) portal (<https://cida.usgs.gov/ngwmn/index.jsp>, accessed on 29 January 2023).

The P and ET rasters (Table 1) were resampled to a grid size of $0.125^\circ \times 0.125^\circ$ to ensure consistency with the Zowam and Milewski (2024) [35] study on gridded GWL prediction, conducted in the same study area. Anomalies for the resulting WCI rasters were calculated for each $0.125^\circ \times 0.125^\circ$ grid by subtracting each month's WCI from the annual average (2010 to 2019). Similarly, monthly GWL anomalies were determined at each well location by subtracting each month's averaged GWL from the annual average at that location.

Table 1. Summary of variables used in the study. All data was processed in ArcGIS Pro (version 2.9.3).

ID	Variable	Type	Resolution	Unit
1	P	Grid	0.1° monthly	mm
2	ET	Grid	0.01° monthly	mm
3	GWL	Point	— daily	feet

Note(s): P = Precipitation; ET = Evapotranspiration; GWL = Groundwater level. P and ET were resampled to $0.125^\circ \times 0.125^\circ$, and produced the WCI rasters.

Statistical Correlation

Differencing both datasets (WCI and GWL anomaly) eliminated any existing time correlation and reduced the number of observations by one. The resulting datasets satisfied the assumptions for Pearson correlation, following assessments based on strip charts, histograms, and bivariate plots, while considering only substantial departures from normality. Pearson correlation was performed simultaneously on 59 grids, each corresponding to the location of a monitoring well, to determine the strength and direction of the linear relationship between the variables (r).

$$r_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (1) [40]$$

Where r_{XY} is the Pearson correlation coefficient between the variables X and Y ranging from -1 (perfect negative) to $+1$ (perfect positive); X_i and Y_i are individual observations of variables X and Y respectively; $\bar{X}$ is the mean of variable X ; $\bar{Y}$ is the mean of variable Y [40].

The statistical correlation described above assumes a contemporaneous relationship between WCI and GWL anomalies, which may not necessarily be the case. Thus, we repeated the correlations at several lag intervals, up to a predetermined maximum of 12 months. This approach is known as cross-correlation and examines whether the fluctuations in one variable precede, are led by, or occur contemporaneously with fluctuations in the other variable [41]. The result of the cross-correlation function (ccf) analysis is a plot of correlation coefficients at the various examined lags, where the x-axis represents lag intervals, and y-axis indicates the correlation coefficients. The x-axis extends equally in both positive and negative directions from zero, meaning that a maximum lag of 12 months, for example, will show correlation coefficients at 24 different lags (−12 to +12). Considering how the analysis was set up, substantial correlation coefficients at negative lags indicate that the input time series precedes the output series, and vice-versa [42]. In this study, WCI anomaly (WCIA) is the input variable and GWL anomaly (GWLA) is the response variable of interest.

Granger Causality

Originally developed in the field of economics [43], Granger Causality (GC) has been applied across various disciplines including atmospheric and climate sciences [44,45], and recently, in hydrogeology, to understand groundwater patterns [46].

Given two stationary time series X and Y, we say X granger causes Y if our ability to predict future values of Y is enhanced when using all information except current values of X [43,46]. This is expressed mathematically as:

$$X_t = \sum_{j=1}^m a_j X_{t-j} + \sum_{j=1}^m b_j Y_{t-j} + \epsilon_t \quad (2) [43]$$

And:

$$Y_t = \sum_{j=1}^m c_j X_{t-j} + \sum_{j=1}^m d_j Y_{t-j} + n_t \quad (3) [43]$$

Where X_t and Y_t are two stationary time series, j is the current lag, m is the maximum number of lags considered, a_j , b_j , c_j and d_j are coefficients, and ϵ_t and n_t represent the two uncorrelated white noise series [43]. X_t will granger cause Y_t if c_j is not zero [43]. Similarly, Y_t will granger cause X_t if b_j is not zero [43]. If both events occur simultaneously, then a feedback relationship exists between both variables [43]. This is known as bidirectional Granger-causality [47,48,49].

GC tests were conducted on the significant lags identified in the ccf analysis, evaluating the influence of past values of one-time series on another. The null hypothesis that the past values of WCIA do not provide any useful information for predicting GWLA (and vice versa) was tested. The hypothesis for each test was rejected when a P-value less than 0.05 was returned.

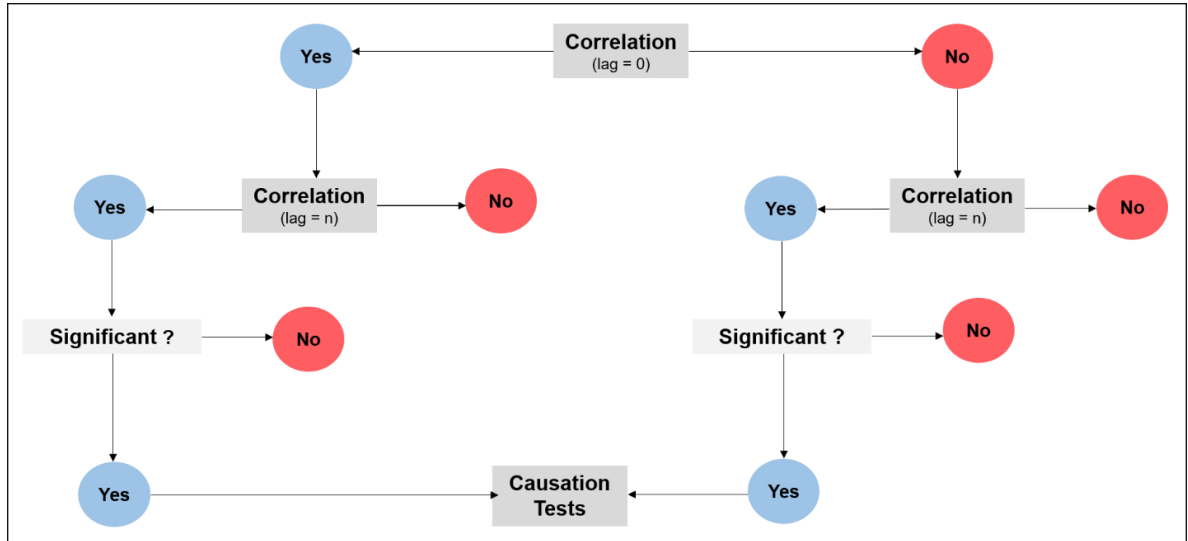


Figure 2. Methodological flow chart. The red (No) symbols imply that a correlation is close to zero. The blue (Yes) symbols indicate that a correlation is sufficiently different

from zero. A significant correlation is sufficiently different from zero and meets the threshold for statistical significance.

Results

All analyses were carried out in R (version 4.1.1). Statistical correlations were conducted at each monitoring well location, first at zero lags and then examined at lags up to 12 months (−12 to +12). The results (Table 2) show the correlation coefficients at zero lag, the maximum correlation coefficient (considering all lags) along with its corresponding lag, and an assessment of the significance of the correlation coefficients at both zero lag and the maximum correlation. Cross-correlation coefficients are normally distributed and considered statistically significant when they exceed the 95% confidence interval threshold, indicating that the observed correlation is unlikely due to random chance [50]. Maximum correlation (Max Cor) values greater than ± 0.4 are shown in **bold** (Table 2). The wells corresponding to these values (Figure 3) were the candidates for the GC test.

Table 2. Summary of correlation analysis at all monitoring well locations. ‘Cor’ refers to the correlation coefficient, and ‘Max Cor’ refers to the maximum correlation coefficient considering all lags. Negative (−ve) lags at the Max Cor suggest that WCIA precedes GWLA, while positive (+ve) lags indicate the reverse.

ID	Cor (lag = 0)	Significant? (Lag = 0)	Max Cor	Lag (Max Cor)	Significant? (Max Cor)
1	−0.07	No	+0.45 −0.44	−1 +1	Yes
2	+0.03	No	−0.23	−11	Yes
3	−0.09	No	−0.20	−3	Yes

4	+0.08	No	-0.33	-10	Yes
5	-0.25	Yes	-0.25	0	n/a
6	+0.00	No	+0.12	+3	No
7	-0.07	No	+0.22	+2	Yes
8	-0.16	No	+0.29	+4	Yes
9	-0.26	Yes	+0.35	+1	Yes
10	+0.19	Yes	-0.27	+1	Yes
11	-0.10	No	-0.26	-11	Yes
12	+0.09	No	+0.25	-6	Yes
13	-0.09	No	+0.30	+7	Yes
14	+0.12	No	-0.44	+1	Yes
15	+0.05	No	-0.25	-3	Yes
16	-0.21	Yes	+0.27	+3	Yes
17	-0.02	No	+0.26	-3	Yes
18	+0.03	No	-0.26	-2	Yes
19	-0.07	No	+0.21	+4	Yes
20	-0.03	No	-0.24	+1	Yes
21	+0.05	No	-0.32	+1	Yes
22	-0.38	Yes	-0.38	0	n/a
23	+0.03	No	+0.32	8	Yes
24	-0.06	No	-0.38	1	Yes
25	+0.19	Yes	+0.19	0	n/a
26	-0.37	Yes	-0.37	0	n/a
27	-0.08	No	+0.17	+7	No

28	−0.15	No	+0.37	+6	Yes
29	−0.19	Yes	+0.45	−5	Yes
30	+0.23	Yes	−0.24	−4	Yes
31	−0.20	Yes	−0.37	+7	Yes
32	+0.02	No	+0.21	+10	Yes
33	−0.00	No	−0.60	+1	Yes
34	−0.17	No	−0.17	0	n/a
35	−0.24	Yes	+0.25	+3	Yes
36	+0.00	No	−0.22	−9	Yes
37	−0.01	No	+0.24	+11	Yes
38	−0.05	No	+0.23	+8	Yes
39	−0.07	No	+0.29	−7	Yes
40	−0.12	No	+0.19	+3	Yes
41	+0.09	No	−0.42	+2	Yes
42	−0.11	No	−0.26	−2	Yes
43	+0.12	No	−0.43	+1	Yes
44	−0.07	No	+0.31	+3	Yes
45	−0.15	No	−0.29	−6	Yes
46	−0.12	No	−0.26	−6	Yes
47	−0.11	No	−0.31	1	Yes
48	−0.05	No	−0.35	1	Yes
49	−0.35	Yes	−0.35	0	n/a
50	+0.18	No	−0.42	+1	Yes
51	−0.31	Yes	+0.38	−8	Yes

52	-0.14	No	-0.33	-12	Yes
53	-0.19	Yes	-0.26	+12	Yes
54	-0.16	No	+0.36	-9	Yes
55	+0.05	No	+0.15	+10	No
56	+0.06	No	+0.24	-6	Yes
57	-0.15	No	-0.35	+7	Yes
58	+0.03	No	+0.31	-4	Yes
59	-0.06	No	-0.20	-2	Yes

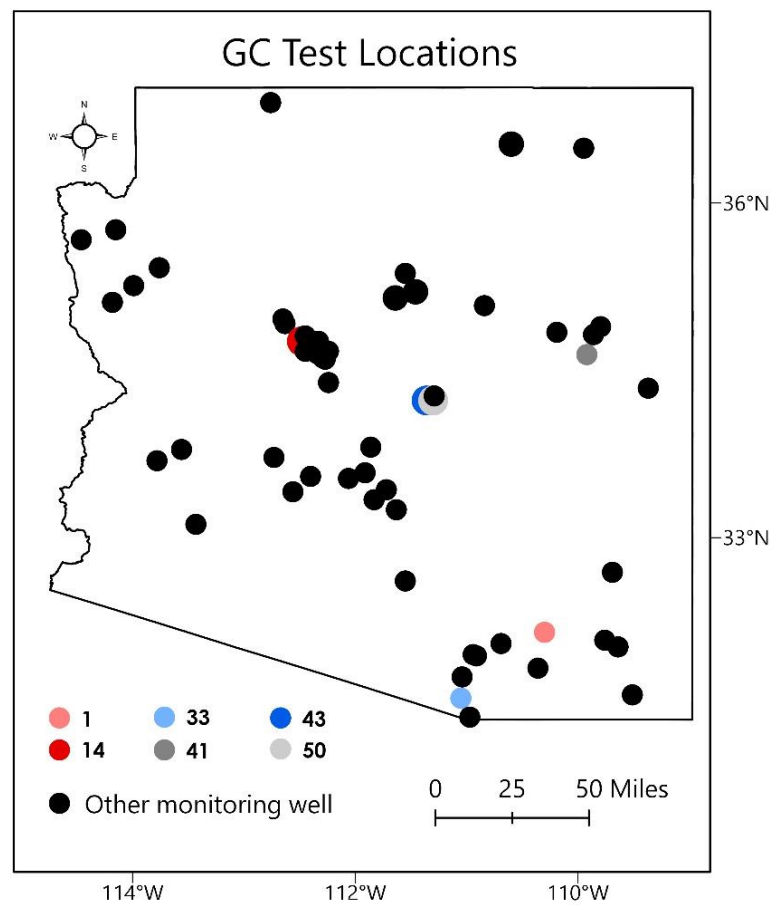


Figure 3. Location of the GC test wells showing the dominant trend. The strongest correlations occurred at the locations of the colored circles.

Examining Well 33 and considering the way the analysis was set up, WCIA preceding GWLA is indicated with red lines, and the reverse is shown in blue (Figure 4). In this example, the variables are not contemporaneously correlated since the maximum correlation occurs at a non-zero lag (+1).

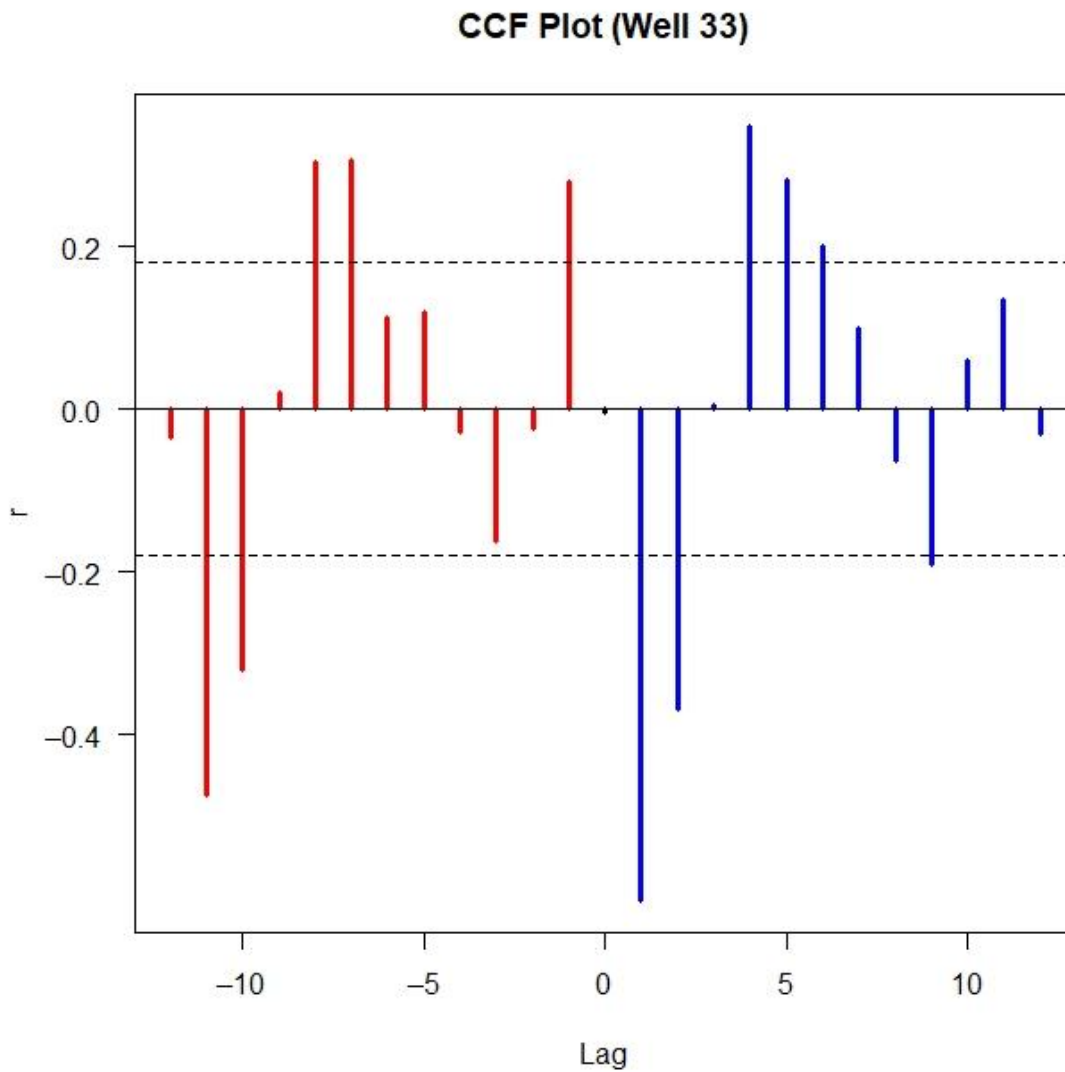


Figure 4. Ccf plot for Well 43. Negative lags (red) suggest that WCIA precedes GWLA,

and positive lags (blue) show GWLA preceding WCIA. The correlation at Zero lag is shown in the solid black line, and the black dotted lines represent the lower and upper confidence limits beyond which a correlation coefficient (y-axis) is statistically significant.

GC tests are forward-looking time-series analyses that determine whether past values of a predictor variable are useful for predicting current values of the target variable beyond the information obtained from past values of the target variable itself. Based on the P-Values (Table 3) and a 5% significance level, we reject the null hypothesis that past values of GWLA do not contain useful information for predicting WCIA values at all six wells. Similarly, we reject the null hypothesis that past values of WCIA do not contain useful information for predicting GWLA values at wells 1 and 33, where bidirectional causality may be occurring.

Table 3. Results of the GC tests conducted at the wells with the largest lagged correlation coefficients. $WCIA \sim GWLA$ implies that GWLA is treated as the independent variable, as determined by the ccf analysis for the respective wells. $GWLA \sim WCIA$ implies that WCIA is treated as the independent variable, to evaluate the presence of bidirectional causality.

ID	Lag (Max Cor)	P-Value (WCIA ~ GWLA)	P-Value (GWLA ~ WCIA)
1	+1	0.000000062	0.000000053
14	+1	0.0000016	0.827
33	+1	0.00000000000038	0.001
41	+2	0.0000000065	0.148
43	+1	0.0000012	0.08
50	+1	0.0000054	0.06

Some of the GC test wells showed a close association with Active Management Areas (AMAs) (Figure 5). AMAs are areas experiencing groundwater overdraft issues due to their strong reliance on groundwater, prompting the active management of the resource [51]. For instance, the relationship between WCIA and GWLA is particularly notable at Well 33, as indicated by the results of the correlation analysis (Table 2), GC tests (Table 2), its association with an AMA (Figure 5), and the scatter plots of contemporaneous and lagged correlations (Figure 6).

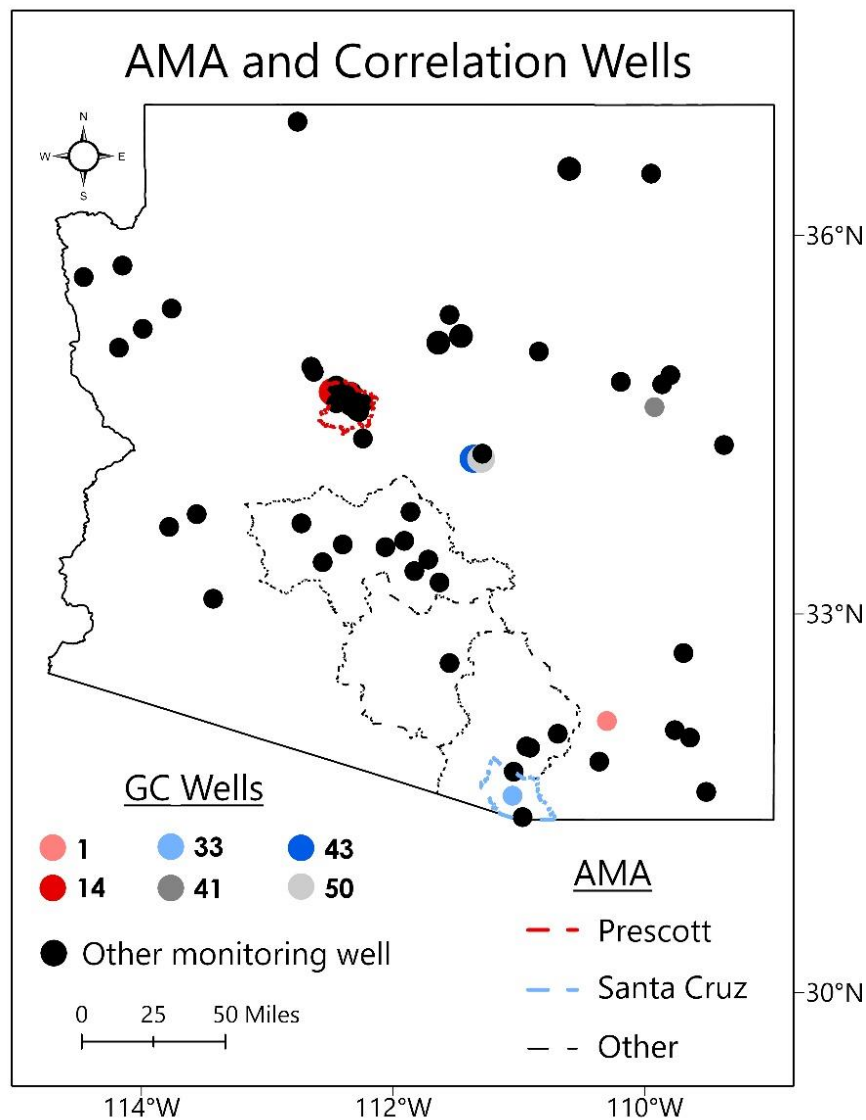


Figure 5. AMAs and correlation wells. The GC test wells represent the locations of the largest correlation coefficients in the study area. Some of these wells were located within AMAs (Prescott and Santa Cruz).

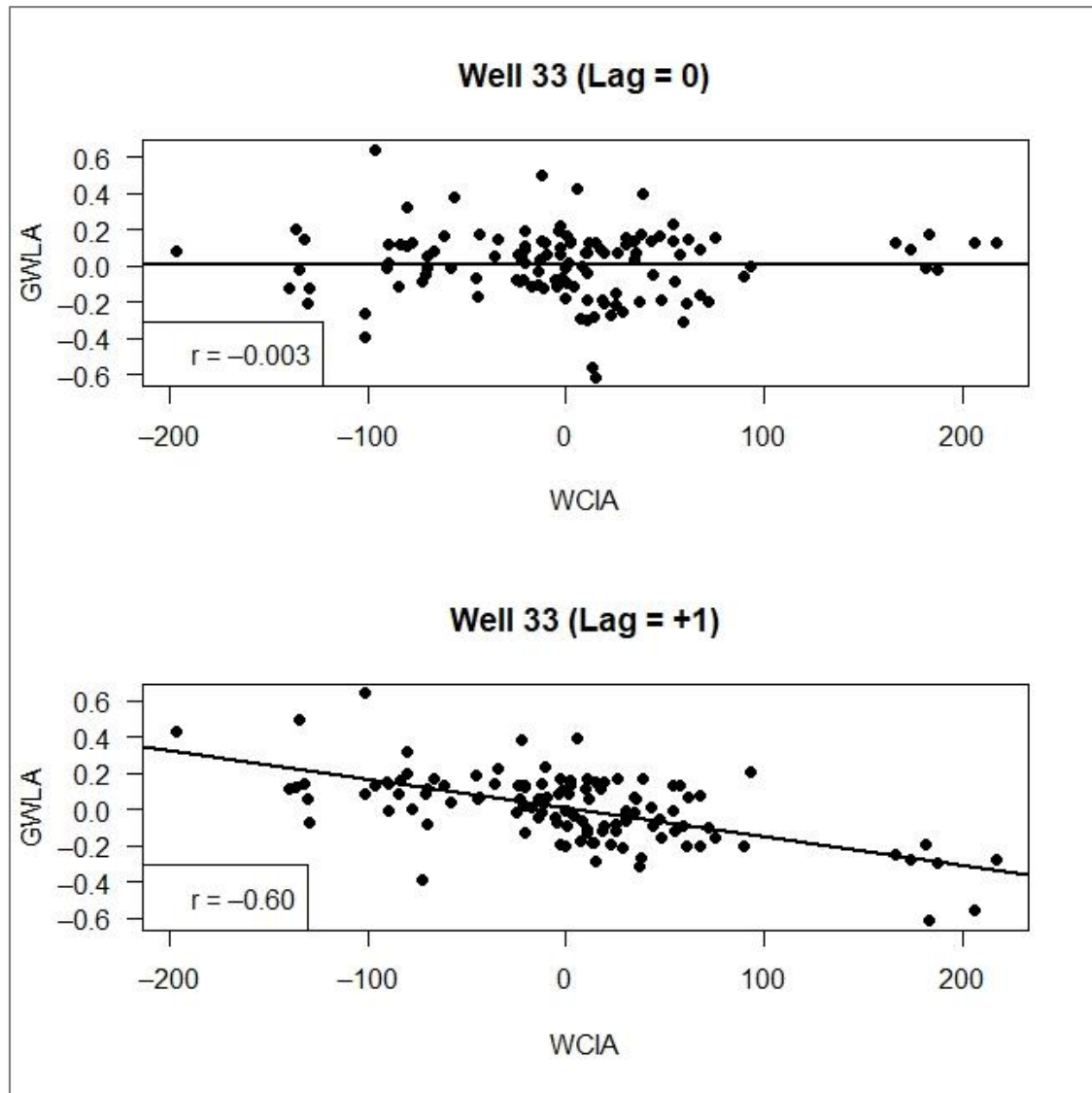


Figure 6. Scatter plots for Well 33 illustrate contemporaneous (lag = 0) and lagged (lag = +1) relationships between WCI and GWL anomalies.

The ccf analysis, initially conducted on 59 monitoring well sites, was extended to 59 ungauged sites (Figure 7) using the predicted GWLA data from Chapter Three. None

of these ungauged sites showed a correlation up to the threshold magnitude of 0.4 (Table 4).

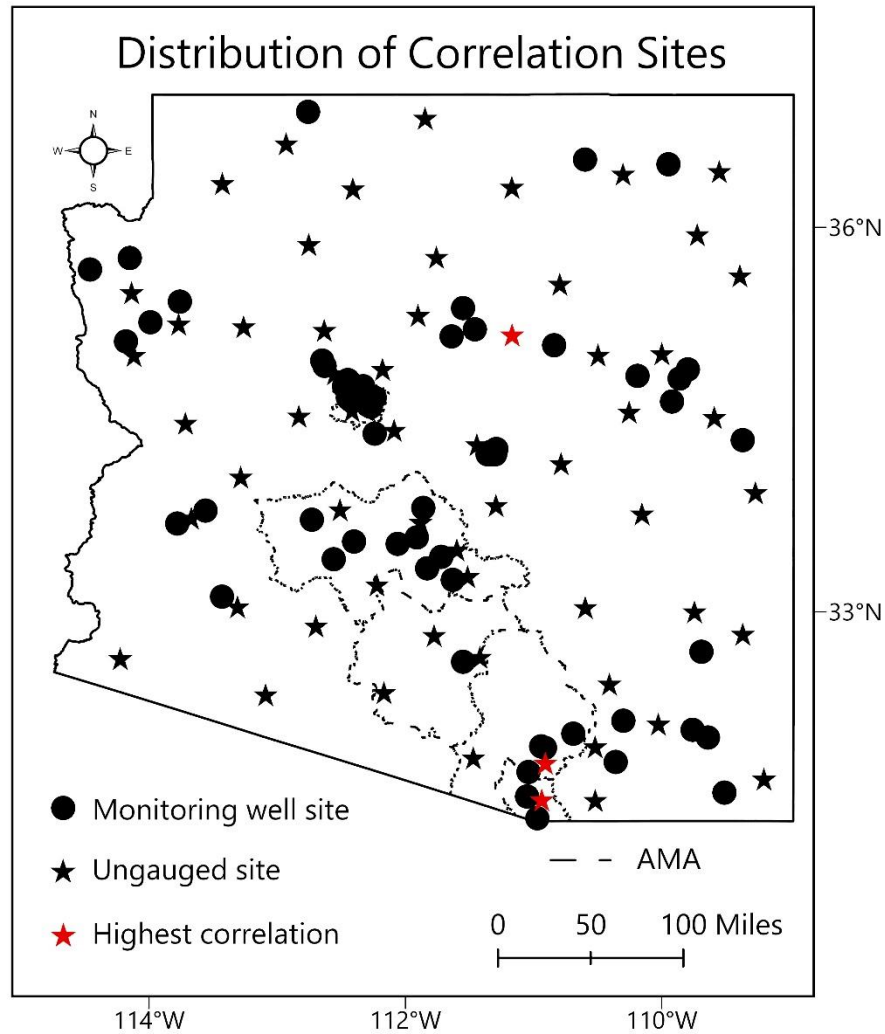


Figure 7. Locations where ccf analysis between GWLA and WCIA were conducted. The black circles indicate the original monitoring well locations. The black stars represent ungauged sites corresponding to the locations of the WCIA (chapter Two) and ML-predicted GWLA (chapter Three) grids. The red stars indicate the locations where the black stars showed the highest correlation coefficients.

Table 4. Summary of ccf analysis between GWLA and WCIA at the ungauged sites. The wells with lagged correlation coefficients greater than ± 0.3 are shown in **bold**.

ID	Cor (lag = 0)	Significant? (Lag = 0)	Max Cor	Lag (Max Cor)	Significant? (Max Cor)
1	-0.05	No	-0.23	+8	Yes
2	-0.04	No	+0.15	-6	No
3	-0.01	No	+0.18	+5	No
4	+0.12	No	+0.26	+2	Yes
5	-0.10	No	-0.21	-4	No
6	-0.03	No	-0.20	+2	Yes
7	+0.19	Yes	-0.33	+4	Yes
8	+0.12	No	-0.33	+4	Yes
9	+0.13	No	-0.26	+4	Yes
10	+0.07	No	-0.29	-8	Yes
11	-0.07	No	+0.29	-8	Yes
12	-0.07	No	-0.19	+1	Yes
13	+0.00	No	-0.26	+1	Yes
14	+0.17	Yes	+0.17	0	n/a
15	+0.11	No	-0.19	-9	Yes
16	-0.26	Yes	+0.33	+2	Yes
17	+0.02	No	+0.15	+10	No
18	-0.08	No	-0.17	+10	No
19	-0.13	No	-0.19	+9	Yes
20	-0.12	No	+0.25	-1	Yes
21	+0.15	No	-0.22	-1	Yes

22	+0.24	Yes	+0.24	0	n/a
23	−0.05	No	−0.17	−5	No
24	−0.07	No	+0.20	−5	Yes
25	−0.09	No	+0.22	+11	Yes
26	−0.09	No	−0.26	+8	Yes
27	+0.05	No	−0.18	+8	No
28	+0.08	No	+0.16	−5	No
29	+0.15	No	+0.16	−5	No
30	−0.12	No	+0.22	+6	Yes
31	+0.21	Yes	+0.21	0	n/a
32	+0.12	No	+0.29	+12	Yes
33	−0.08	No	+0.16	+1	No
34	−0.07	No	−0.18	+7	No
35	+0.03	No	+0.26	+4	Yes
36	+0.10	No	+0.16	−4	No
37	+0.05	No	−0.20	−2	Yes
38	−0.17	No	−0.26	+4	Yes
39	+0.02	No	−0.18	−3	No
40	+0.09	No	+0.21	+4	Yes
41	+0.03	No	+0.18	−5	No
42	−0.06	No	+0.20	+11	Yes
43	+0.20	Yes	+0.20	+8	Yes
44	−0.09	No	+0.23	+1	Yes
45	+0.14	No	+0.17	+8	No

46	-0.10	No	-0.24	-8	Yes
47	-0.10	No	-0.21	-4	Yes
48	+0.02	No	+0.19	-10	Yes
49	-0.04	No	-0.16	+10	No
50	-0.05	No	-0.17	+8	No
51	-0.01	No	+0.23	-8	Yes
52	-0.01	No	+0.21	+4	Yes
53	+0.03	No	-0.24	+4	Yes
54	+0.06	No	+0.23	-12	Yes
55	-0.02	No	-0.25	-8	Yes
56	+0.10	No	-0.20	+2	No
57	+0.04	No	+0.16	+1	No
58	+0.12	No	+0.23	-10	Yes
59	+0.09	No	+0.22	-3	Yes

Discussion

The relationship between GWLs and driving factors is typically complex and nonlinear [35,52–55], but if precipitation is typically linearly correlated with GWL [56–58], likewise ET [58], we expect the WCI, a variable obtained from precipitation and ET, to also show a linear relationship with GWL. This was mostly not the case for the contemporaneous relationship (lag = 0) between GWLA and WCIA examined in this study. However, a moderate to strong linear relationship was observed at six monitoring well

locations (1, 14, 33, 41, 43, and 50) when WCIA at time $t + k$ was compared with GWLA, where k represents the lag (Table 2). This correlation occurred at positive lags, implying that changes in WCIA lagged GWLA, and was negative at all six sites, indicating that an increase in lagged values of WCIA corresponds to a decrease in GWLA values, and vice versa (Table 2). Groundwater influences climate in various ways, including contributions to soil moisture [19]. Aquifers are also often hydrologically connected to surface water bodies that they feed or that feed them, which may also influence the climate and the water cycle by providing additional moisture (in the case of a gaining stream). The findings in this study corroborate some of these known interactions and present a new outlook for climate and groundwater relationships in the study area.

Contemporaneous correlation ($\text{lag} = 0$) was observed at wells 5, 22, 25, 26, 34, and 49 (Table 2), because at those wells, the maximum correlation appeared at zero lag, indicating that the observed fluctuations between the two variables occurred within the same time period (without any delay between them). This correlation was not significant at well 34, but at wells 5 and 25, it was the only statistically significant correlation. Therefore, at wells 5 and 25, we accept the null hypothesis that there is no significant lead-lag relationship between both variables at the examined lags. Also, the maximum correlations at non-zero lags were not statistically significant at wells 6, 27, and 55 (Table 2). At these wells (as well as at well 34), there is no evidence of a relationship between WCIA and GWLA.

The strongest lagged correlation (-0.60) was observed at well 33 (Table 2) and occurred at a lag of $+1$, indicating that GWLA in the current month had a strong linear relationship with WCIA in the following month (and vice-versa). GWLA also showed the

strongest relationships with precipitation (-0.58) and evapotranspiration (-0.43) anomalies, all three occurring at the same lag (+1), suggesting that the influence of GWL on climate at that location occurs at a lag of +1. GC tests returned statistically significant results at all six wells based on P-values (Table 3), implying that the past values of GWLA provide useful information for predicting WCIA values beyond what is provided by the autoregressive structure of the WCIA variable. But at wells 1 and 33, WCIA also provided useful information for predicting GWLA (Table 3). This suggests that bidirectional causality exists at those locations (and lag). However, at well 33 specifically, the much larger F-statistic implied stronger evidence for the former (GWLA granger causing WCIA). F-statistic measures the ratio of explained variance to unexplained variance in statistical analyses [59], and the corresponding p-value represents the probability of observing the statistic or one more extreme, if the null hypothesis, is true [60].

Out of the five AMAs – Santa Cruz, Prescott, Phoenix, Pinal, and Tucson [51], Prescott and Santa Cruz contained monitoring wells with some of the strongest lagged correlations, including well 33 (Figure 5). This suggests that these AMAs might be hydrogeologically sensitive to shifts in the terrestrial water cycle. At the ungauged sites, we lowered the substantial correlation threshold to a magnitude of 0.3 to account for uncertainties in the GWLA predictions. Two of the three ungauged sites where the highest correlation coefficients were observed showed a “GWLA leading WCIA” relationship (Table 4). One of the them happened to also be in the same AMA as Well 33 (Table 4, Well 7; Figure 7). Given that several other factors simultaneously influence GWLs, future research must expand on this insight to understand how each of these factors influences the WCI-GWL relationships at the AMAs and other vulnerable areas.

Conclusion

The dynamic relationship between WCIs and GWLs at local scales across the predominantly arid U.S. State of Arizona was evaluated using statistical correlation and causation analyses. Circling back to the hypotheses outlined in the introduction: (1) GWLA did not show a strong negative contemporaneous correlation with WCIA. The correlation was weak or absent, and not necessarily negative across all wells. (2) Although the study identified a lead-lag relationship between both variables, this relationship was also not strong across all 59 wells and where it was strong, GWLA generally led WCIA. Therefore, we reject both hypotheses.

The strongest lagged correlation coefficients across all (gauged and ungauged) sites were dominantly negative. This trend implies that a continuous intensification of the water cycle reflects a decline in past GWLs in the affected areas. This backward interpretation may help determine when immediate management responses and swift interventions are necessary. The connections with AMAs reinforce the need for the continuous monitoring and effective management of groundwater in vulnerable areas.

Reference

1. Kumar, C.P. Climate change and its impact on groundwater resources. *Int. J. Eng. Sci.* **2012**, 1(5), 43-60.
2. Trenberth, K.E. Changes in precipitation with climate change. *Clim. Res.* **2011**, 47(1-2), 123-138.

3. Trenberth, K.E. Climate change caused by human activities is happening and it already has major consequences. *J. Energy Nat. Resour. Law.* **2018**, 36(4), 463-481.
4. Pisor, A.C.; Touma, D.; Singh, D.; Jones, J.H. To understand climate change adaptation, we must characterize climate variability: Here's how. *One Earth* **2023**, 6(12), 1665-1676.
5. von der Heydt, A. S., Ashwin, P., Camp, C. D., Crucifix, M., Dijkstra, H. A., Ditlevsen, P., & Lenton, T. M. Quantification and interpretation of the climate variability record. *Glob. Planet. Change* **2021**, 197, 103399.
6. Crowley, R.A.; Health and Public Policy Committee of the American College of Physicians. Climate change and health: a position paper of the American College of Physicians. *Ann. Intern. Med.* **2016**, 164(9), 608-610.
7. McMichael, A.J.; Lindgren, E. Climate change: present and future risks to health, and necessary responses. *J. Intern. Med.* **2011**, 270(5), 401-413.
8. Ebi, K.L.; Hess, J.J. Health Risks Due to Climate Change: Inequity in Causes and Consequences: Study examines health risks due to climate change. *Health Aff.* **2020**, 39(12), 2056-2062.
9. Chinthrajah, S.; Garcia, E.; Hasan, Z.; Hy, A.; Wong, L. Climate Change Health Effects and What You Can Do. *Am. J. Respir. Crit. Care Med.* **2022**, 205(1), P1-P2.
10. Ebi, K.L.; Vanos, J.; Baldwin, J.W.; Bell, J.E.; Hondula, D.M.; Errett, N.A.; Hayes, K.; Reid, C.E.; Saha, S.; Spector, J.; Berry, P. Extreme weather and climate

- change: population health and health system implications. *Annu. Rev. Public Health* **2021**, 42(1), pp.293-315.
11. Clarke, B.; Otto, F.; Stuart-Smith, R.; Harrington, L. Extreme weather impacts of climate change: an attribution perspective. *Environ. Res.: Climate* **2022**, 1(1), 012001.
 12. Greenough, G.; McGeehin, M.; Bernard, S.M.; Trtanj, J.; Riad, J.; Engelberg, D. The potential impacts of climate variability and change on health impacts of extreme weather events in the United States. *Environ. Health Perspect.* **2001**, 109(suppl 2), 191-198.
 13. Bolan, S.; Padhye, L.P.; Jasemizad, T.; Govarthanan, M.; Karmegam, N.; Wijesekara, H.; Amarasiri, D.; Hou, D.; Zhou, P.; Biswal, B.K.; Balasubramanian, R. Impacts of climate change on the fate of contaminants through extreme weather events. *Sci. Total Environ* **2023**, p.168388.
 14. Fried, J.S.; Torn, M.S.; Mills, E. The impact of climate change on wildfire severity: a regional forecast for northern California. *Clim. Change* **2004**, 64(1), 169-191.
 15. Bajželj, B.; Richards, K.S. The positive feedback loop between the impacts of climate change and agricultural expansion and relocation. *Land* **2014**, 3(3), 898-916.
 16. Lall, U.; Josset, L.; Russo, T. A snapshot of the world's groundwater challenges. *Annu. Rev. Environ. Resour.* **2020**, 45(1), 171-194.
 17. Kundzewicz, Z. W.; Doell, P. Will groundwater ease freshwater stress under climate change? *Hydrol. Sci. J.* **2009**, 54(4), 665-675.

18. Ahlfeld, D.P.; Dripps, W.R. Groundwater Issues. In *Water: Science, Policy, and Management: Challenges and Opportunities.*; American Geophysical Union: Washington, DC, United States of America, 2003, 16, 79-98.
19. Taylor, R.G.; Scanlon, B.; Döll, P.; Rodell, M.; Van Beek, R.; Wada, Y.; Longuevergne, L.; Leblanc, M.; Famiglietti, J.S.; Edmunds, M.; Konikow, L. Ground water and climate change. *Nat. Clim. Change* **2013**, 3(4), pp.322-329.
20. Loáiciga, H. A. Climate change and ground water. *ANN ASSOC AM GEOGR* **2003**, 93(1), 30-41.
21. Barbieri, M.; Barberio, M.D.; Banzato, F.; Billi, A.; Boschetti, T.; Franchini, S.; Gori, F.; Petitta, M. Climate change and its effect on groundwater quality. *Environ Geochem Health* **2023**, 45(4), pp.1133-1144.
22. Aziz, Z. 2023. Potential Impacts of Climate Change on Groundwater Quality. Division of Science and Research. New Jersey Department of Environmental Protection. Trenton, NJ. 56 pages. Available online: <https://dspace.njstatelib.org/handle/10929/112368> (assessed 7 August 2024)
23. Kløve, B.; Ala-Aho, P.; Bertrand, G.; Gurdak, J.J.; Kupfersberger, H.; Kværner, J.; Muotka, T.; Mykrä, H.; Preda, E.; Rossi, P.; Uvo, C.B. Climate change impacts on groundwater and dependent ecosystems. *J. Hydrol.* 2014, 518, pp.250-266.
24. Ranjan, P.; Kazama, S.; Sawamoto, M. Effects of climate change on coastal fresh groundwater resources. *Glob. Environ. Change* **2006**, 16(4), 388-399.
25. Cantelon, J.A.; Guimond, J.A.; Robinson, C.E.; Michael, H.A.; Kurylyk, B.L. Vertical saltwater intrusion in coastal aquifers driven by episodic flooding: a review. *Water Resour. Res.* **2022**, 58(11), e2022WR032614.

26. Dao, P.U.; Heuzard, A.G.; Le, T.X.H.; Zhao, J.; Yin, R.; Shang, C.; Fan, C. The impacts of climate change on groundwater quality: A review. *Sci. Total Environ.* **2023**, 169241.
27. Gurdak, J.S.; Hanson, R.T.; Green, T.R. Effects of climate variability and change on groundwater resources of the United States. *US Geological Survey* **2009**, No. 2009-3074.
28. Hanson, R. T.; Newhouse, M. W.; Dettinger, M. D. A methodology to assess relations between climatic variability and variations in hydrologic time series in the southwestern United States. *J. Hydrol.* **2004**, 287(1-4), 252-269.
29. Yan, X.; Cheng, P.; Zhang, Q.; Li, X.; He, J.; Yan, X.; Zhao, W.; Wang, L. Comparisons of climate change characteristics in typical arid regions of the Northern Hemisphere. *Front. environ. sci.* **2022**, 10, p.1033326.
30. Trenberth, K. E. Water cycles and climate change. *Glob. Environ. Change* **2014**, 31-37.
31. Malinowski, Ł.; Skoczko, I. Impacts of climate change on hydrological regime and water resources management of the Narew River in Poland. *J. Ecol. Eng.* **2018**, 19(4).
32. Bosilovich, M.G.; Schubert, S.D.; Walker, G.K. Global changes of the water cycle intensity. *J. Clim.* **2005**, 18(10), 1591-1608.
33. Huntington, T.G.; Weiskel, P.K.; Wolock, D.M.; McCabe, G.J. A new indicator framework for quantifying the intensity of the terrestrial water cycle. *J. Hydrol.* **2018**, 559, 361-372.

34. Zowam, F.J.; Milewski, A.M.; Richards IV, D.F. A Satellite-Based Approach for Quantifying Terrestrial Water Cycle Intensity. *Remote Sens.* **2023**, *15*(14), 3632.
35. Zowam, F.J.; Milewski, A.M. Groundwater Level Prediction Using Machine Learning and Geospatial Interpolation Models. *Water* **2024**, *16*(19), 2771.
36. Pradhan, R.K.; Markonis, Y.; Godoy, M.R.V.; Villalba-Pradas, A.; Andreadis, K.M.; Nikolopoulos, E.I.; Papalexiou, S.M.; Rahim, A.; Tapiador, F.J.; Hanel, M. Review of GPM IMERG performance: A global perspective. *Remote Sens. Environ.* **2022**, *268*, 112754.
37. Mohammed, S.A.; Hamouda, M.A.; Mahmoud, M.T.; Mohamed, M.M. Performance of GPM-IMERG precipitation products under diverse topographical features and multiple-intensity rainfall in an arid region. *Hydrol. Earth Syst. Sci.* **2020**, *2020*, 1-27.
38. Morsy, M.; Scholten, T.; Michaelides, S.; Borg, E.; Sherief, Y.; Dietrich, P. Comparative analysis of TMPA and IMERG precipitation datasets in the arid environment of El-Qaa plain, Sinai. *Remote Sens.* **2021**, *13*(4), 588.
39. Elnashar, A.; Wang, L.; Wu, B.; Zhu, W.; Zeng, H. Synthesis of global actual evapotranspiration from 1982 to 2019. *Earth Syst. Sci. Data* **2021**, *13*(2), 447-480.
40. Lee Rodgers, J.; Nicewander, W.A. Thirteen ways to look at the correlation coefficient. *Am. Stat.* **1988**, *42*(1), 59-66.
41. Jackson, J.C.; Caluori, N.; Abrams, S.; Beckman, E.; Gelfand, M.; Gray, K. Tight cultures and vengeful gods: How culture shapes religious belief. *J. Exp. Psychol. Gen.* **2021**, *150*(10), 2057.

42. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time series analysis: forecasting and control*; John Wiley & Sons: Hoboken, New Jersey, U.S.A, 2015; ISBN 978-1-118-67502-1.
43. Granger, C.W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society* **1969**, 424-438.
44. McGraw, M.C.; Barnes, E.A. Memory matters: A case for Granger causality in climate variability studies. *J. Clim.* **2018**, 31(8), 3289-3300.
45. Attanasio, A.; Pasini, A.; Triacca, U. Granger causality analyses for climatic attribution. *Atmospheric and Climate Sciences* **2013**, 3(4) pp. 515-522.
46. Singh, N. K., & Borrok, D. M. A Granger causality analysis of groundwater patterns over a half-century. *Sci. Rep.* **2019**, 9(1), 12828.
47. Maziarz, M. A review of the Granger-causality fallacy. *J. Philos. Econ.* **2015**, 8(2), 86-105.
48. Lopez, L.; Weber, S. Testing for Granger causality in panel data. *Stata J.* **2017**, 17(4), 972-984.
49. Chvosteková, M.; Jakubík, J.; Krakovská, A. Granger causality on forward and reversed time series. *Entropy* **2021**, 23(4), 409.
50. Fathian, F.; Fakheri-Fard, A.; Modarres, R.; Van Gelder, P. H. A. J. M. Regional scale rainfall-runoff modeling using VARX-MGARCH approach. *Stoch. Environ. Res. Risk Assess.* **2018**, 32, 999-1016.
51. Majumdar, S.; Smith, R.; Conway, B.D.; Lakshmi, V. Advancing remote sensing and machine learning-driven frameworks for groundwater withdrawal estimation

- in Arizona: Linking land subsidence to groundwater withdrawals. *Hydrol. Process.* **2022**, 36(11), e14757.
52. Haider, A.; Lee, G.; Jafri, T.H.; Yoon, P.; Piao, J.; Jhang, K. Enhancing Accuracy of Groundwater Level Forecasting with Minimal Computational Complexity Using Temporal Convolutional Network. *Water* **2023**, 15(23), 4041.
 53. Tao, H.; Hameed, M.M.; Marhoon, H.A.; Zounemat-Kermani, M.; Heddami, S.; Kim, S.; Sulaiman, S.O.; Tan, M.L.; Sa'adi, Z.; Mehr, A.D.; Allawi, M.F.; Yaseen, Z.M. Groundwater level prediction using machine learning models: A comprehensive review. *Neurocomputing* **2022**, 489, 271-308.
 54. Ardana, P.D.H.; Redana, I.W.; Yekti, M.I.; Simpen, I.N. Groundwater Level Forecasting Using Multiple Linear Regression and Artificial Neural Network Approaches. *Civ. Eng. Archit.* **2022**, 10(3), 784-799.
 55. Najafabadipour, A.; Kamali, G.; Nezamabadi-Pour, H. Application of Artificial Intelligence Techniques for the Determination of Groundwater Level Using Spatio-Temporal Parameters. *ACS Omega* **2022**, 7(12), 10751-10764.
 56. Hussain, F.; Wu, R.S.; Shih, D.S. Water table response to rainfall and groundwater simulation using physics-based numerical model: WASH123D. *J. Hydrol. Reg. Stud.* **2022**, 39, 100988.
 57. Zhang, M.; Singh, H.V.; Migliaccio, K.W.; Kisekka, I. Evaluating water table response to rainfall events in a shallow aquifer and canal system. *Hydrol. Process.* **2017**, 31(22), 3907-3919.
 58. Yan, S.F.; Yu, S.E.; Wu, Y.B.; Pan, D.F.; Dong, J.G. Understanding groundwater table using a statistical model. *Water Sci. Eng.* **2018**, 11(1), 1-7.

59. Kissell, R.; Poserina, J. Chapter 2—Regression Models. In *Optimal Sports Math, Statistics, and Fantasy*; Kissell, R., Poserina, J., Eds.; Academic Press: Chennai, India, 2017; 39-67.
60. Holland, S. Data Analysis in the Geosciences. Available online: <http://strata.uga.edu/8370/lecturenotes/pvaluesConfidenceIntervals.html> (accessed on 7 September 2021).

CHAPTER 5

CONCLUSION

In this study, the analyses progressed from a broad-scale national focus (chapter two) to local scales (chapters three and four), to understand climate change impacts on groundwater levels at local scales within a vulnerable semiarid/arid region. This chapter provides a summary of the key findings and reinforces their implications for water security and management.

First, the water cycle is speeding up in many parts of the contiguous United States (CONUS) particularly the west, and the El Niño-Southern Oscillation (ENSO), the most prominent naturally occurring climate variability pattern, impacts this intensification by bringing more water to the southeast, west, and parts of the north, and reduced moisture to the northeast region [1]. In particular, the state of Arizona might be experiencing more pronounced intensification compared to other arid regions of the CONUS, and was the focus of the remainder of the study [1].

Machine learning methods were combined with advanced geospatial interpolation models to develop the first statewide groundwater level anomaly (GWLA) prediction for the US State of Arizona across its two distinct aquifer types: the unconsolidated sand and gravel aquifer, and rock aquifer [2]. The multi-model approach, which utilized only remotely sensed input variables, demonstrated satisfactory performance, and can be replicated for similar climates and hydrologically data-sparse and remote areas of the world [2].

Eventually, a detailed statistical correlation analysis between water cycle intensity and groundwater level anomalies (WCIA, GWLA) was conducted at known monitoring well locations and several ungauged sites across the state of Arizona. The dominant (statistically significant) relationship was a negative lead-lag correlation between WCIA and GWLA, observed at eight of the sites (combining gauged and ungauged sites), where an increase in current WCIA would result in a decrease in past GWLA values, and vice-versa [3]. Some of these sites coincided with areas previously designated as Active Management Areas (AMAs), where the strictest groundwater management regulations are enforced [3]. Notably, the strongest negative lead-lag relationship (-0.6) in this study was observed within one of these AMAs [3].

Given that the changes in WCIA lag behind changes in GWLA, an intensified water cycle today may signal an already depleting groundwater resource at the affected sites. Therefore, to mitigate climate change impacts on groundwater availability across the study area, relevant authorities must take cues from studies like this to inform strategic groundwater management decisions in a timely and efficient manner and strengthen groundwater monitoring efforts.

References

1. Zowam, F.J.; Milewski, A.M.; Richards IV, D.F. A Satellite-Based Approach for Quantifying Terrestrial Water Cycle Intensity. *Remote Sens.* 2023, 15(14), 3632.
2. Zowam, F.J.; Milewski, A.M. Groundwater Level Prediction Using Machine Learning and Geospatial Interpolation Models. *Water.* 2024, 16(19), 2771.

3. Zowam, F.J.; Milewski, A.M. Climate Variability and Groundwater Levels: A Correlation and Causation Analysis. *Water*. 2024, to be submitted.