TRANSLATING SATELLITE-BASED TERRESTRIAL WATER STORAGE INTO IMPROVED RESERVOIR HEIGHT FORECASTING: A CASE STUDY IN THE UPPER PARANÁ RIVER BASIN

Jessica Besnier, Venkataraman Lakshmi

University of Virginia, School of Engineering and Applied Science, Department of Civil and Environmental Engineering

Abstract

Water levels in lakes, rivers, and oceans fluctuate due to the hydrological cycle. This study uses terrestrial water storage (TWS) estimates from GRACE satellite missions to predict reservoir operation in Brazil. Reservoir water elevations, derived from multi-satellite radar altimetry (RA) data, serve as a proxy for operation. Seventeen reservoirs in Southern Brazil were analyzed using two regression techniques and a machine learning (ML) model, incorporating geomorphologic and meteorologic factors such as precipitation and temperature. The ML model, optimized for a six-month forecast horizon, outperformed the others. The random forest regression model with 35 features reduced error by half compared to linear regression, predicting reservoir heights within 1.41 meters (MAE). These findings improve understanding of the relationship between TWS and RA heights in the Upper Paraná Basin, enhancing reservoir height prediction.

Introduction

Water levels fluctuate globally due to various factors like precipitation, runoff, and evaporation (Bates et al., 2008). Understanding these changes is crucial for managing water resources and mitigating the impacts of increasingly frequent and intense hydrological extremes. Reservoirs, critical infrastructure for water supply, energy production, and flood control, are particularly sensitive to these fluctuations (Nasar, 2015; Meng et al., 2021; Polomski and Wiatkowski, 2023). Accurately predicting reservoir height variations is essential for optimizing water allocation, ensuring energy security, and minimizing shortages, thus impacting numerous sectors from agriculture to urban development.

Previous studies have explored reservoir height prediction using machine learning (ML) models, often focusing on individual reservoirs and employing complex algorithms (Sun et al., 2021; Yin et al., 2023;Li et al. 2016 ; Sapitang et al., 2020). While these studies have provided valuable insights, gaps remain in multi-step forecasting, cross-reservoir applicability, and the use of simpler, more interpretable models. Specifically, the ability to predict reservoir heights at multiple future time steps, and the testing of models across varied reservoir systems, is needed. Furthermore, the use of more widely understood models would increase the usability of any outputs of such research.

To address these key research gaps, this study investigates the potential of GRACE TWS to assist in predicting reservoir height fluctuations across multiple reservoirs, using a comparative analysis of linear regression, polynomial regression, and a machine learning-based random forest approach. The goals of this work are to (1) explore the relationship between TWS and reservoir height, rather than focusing on inflows, outflows, or broader water storage dynamics, (2) compare the performance of simple regression models and random forest regression to determine the most effective approach for reservoir height prediction, (3) evaluate the models' ability to provide multi-step forecasts, enabling predictions of reservoir levels at different time horizons, and to (4) assess the generalizability of the models by applying them to a set of reservoirs withing the UPRB. By addressing these knowledge gaps, this study has the potential to advance the understanding of the complex linkages between TWS dynamics and reservoir height and to provide modeling tools for water resource management and decision-making.

Data and Methods

This study focused on seventeen reservoirs within the Upper Parana River Basin (UPRB), selected based on their geographic location, data availability, size, and relevance to the region. To ensure robust model training, reservoirs were included only if they had at least 10 years of overlapping monthly data for key variables: Radar Altimetry (RA) reservoir height (sourced from GREALM, DAHITI, and Hydroweb, spanning 2002-2022), GRACE/GRACE-FO Terrestrial Water Storage (TWS), IMERG precipitation, and GLDAS temperature. Additionally, reservoirs were chosen to be within distinct GRACE pixels, accounting for the dataset's spatial resolution.

Three predictive models were employed to analyze and forecast reservoir height: Linear Regression, Polynomial Regression, and Random Forest Regression. Linear Regression was used to examine the direct linear relationship between TWS and reservoir height. Polynomial Regression aimed to capture potential non-linear relationships by incorporating TWS, precipitation, and temperature as input variables. Random Forest Regression, an ensemble learning method, was selected for its ability to handle complex, non-linear data and high-dimensional inputs. This model was optimized through feature engineering and hyperparameter tuning to enhance predictive accuracy.

Model performance was assessed by comparing predicted reservoir height values against actual RA reservoir height data. To evaluate the models' forecasting capabilities, a 6-month multi-step forecast horizon was used. This approach allowed for the analysis of how well each model predicted reservoir heights at multiple points in the future.

Results

Initial Analysis

Feature correlation between the reservoir height anomalies and precipitation, temperature, and TWS. The Pearson correlations are reported as -0.04, -0.21, and 0.45 for precipitation, temperature, and TWS respectively, thus indicating a stronger correlation between the TWS and reservoir height anomalies than the temperature and precipitation.

3.2 Accuracy of Model Prediction

The linear regression model provided insights into the relationship between TWS and reservoir height anomalies. Figure 1 (a) illustrates the linear relationship along with the regression line equation $y= 0.01x+1.19$. Figure 1(b) shows the actual vs. predicted reservoir height anomalies, with a MAE of 2.81m.

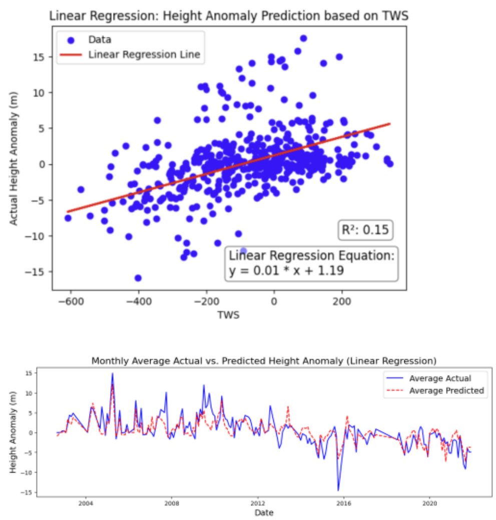

Figure 1. Linear regression was run and (a)the relationship between the TWS and the RA height for all reservoirs for different monthly measurements and the linear line and (b) time series of the actual vs. predicted values over the test period.

Polynomial Regression

Polynomial Regression builds upon linear regression as it also considers the relationships between one target variable and multiple other variables simultaneously, to see how several predictors influence the outcome together. Figure 2 (a) shows the actual vs. predicted height anomalies from the polynomial regression model and the subsequent error values when the degrees of freedom are changed. When changing the degrees of freedom, the error value changes exponentially with three being the lowest error value.

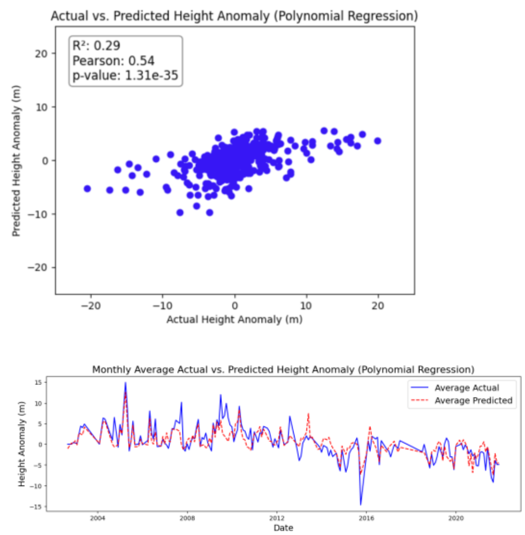

Figure 2: Polynomial Regression Results: (a) The plot of the predicted reservoir height anomalies and the actual height anomalies of the polynomial regression model with degrees = 3. (c) The time series of monthly average test values for the polynomial regression model from all of the reservoirs.

The MAE values tend to increase as the degrees increase, showing the lowest is degrees at 3, indicating that this problem is more linearly related than polynomial as the error barely decreased when switching from linear regression to polynomial regression. The overall MAE for this model is 2.47m, indicating a better predictor than the linear regression, which is expected due to the polynomial capabilities of this model. While linear regression requires the data to be linear, polynomial regression does not and allows for higher degrees of relationships between the datasets. However, there is no significant increase as the data might be linearly related as the decrease in error is not substantial, as it is only a modest improvement at 12.1% from the MAE for linear regression.

Random Forest Regression

To optimize the Random Forest regression model, we systematically examined feature dimensionality, testing configurations from *k*=5, five features to *k*=140,140 features. Through analysis of error metrics and computational efficiency, *k*=35 emerged as the optimal feature count. Figure 3 shows the MAE across different forecast horizons for the different permutations of the model. As the MAE increases, that indicates that the model is predicting worse and further away from the actual values.

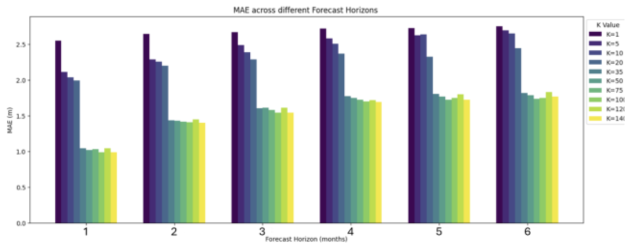

Figure 3: Random Forest Regression Error comparison from *k*=5 to *k*=140. This figure shows the Mean Absolute error across the forecast horizon (6 months) for each day for many different *k*-values.

The rationale for selecting about 35 features was twofold: (1) it provided the best balance between predictive accuracy and computational efficiency and (2) the features at *k*=35, as revealed by importance analysis, included the most significant predictors across forecast horizons. Specifically, the reservoir height from the previous month and five months prior, as well as the previous month's TWS, ranked highest in importance. These predictors were consistently impactful due to the strong temporal autocorrelation in reservoir levels, which supports the hypothesis that recent and lagged reservoir states are critical forecasting.

<u>Model Comparison</u>

After the three of the models were run and the subsequent error values were calculated. They were all recorded and compared against each other. Table 1 shows the error values for Linear regression, Polynomial Regression, and Random Forest Regression with *k*= 35. Based on these results, random forest regression has the lowest error results

Table 1: Error Value Table: This table displays the three models run in the study and their subsequent error values (MSE, RMSE, and MAE). The RF regression error value is the average error across the 6-month forecast horizon for *k*=35

Model	MSE (m^2)	RMSE (m)	MAE (m)
Linear Regression	17.08	4.13	2.81
Polynomial Regression	13.49	3.67	2.47
Random Forest Regression	6.088	2.47	1.41

Based on these error metrics, linear regression has the greatest error with a MAE of 2.81m, followed by polynomial regression with a MAE of 2.47m, then random forest regression with *k*=35 with a MAE of 1.41m. Although the random forest models require the most computing time compared to linear and polynomial regression, the error values are significantly improved. For MAE and RMSE, the units are the same as the output, reservoir height anomalies in meters. In comparison to linear regression, the polynomial regression decreased by a percentage of 12.1% and the random forest regression decreased by 49.8%, or almost half regarding MAE.

For RMSE, linear regression also had the highest error with 4.13m, followed by polynomial regression at 3.67m, and then random forest regression with *k*=35 at 2.47 m. These patterns are also seen in MSE, as that is the square of these values. In comparison to linear regression, polynomial regression decreased the error by 11.1% and random forest regression decreased this error by 40.1%. While the RMSE results are not as drastic as the MAE differences between the different methods, they both express that random forest regression significantly improved the error results for all reservoirs compared to the linear regression.

Model Performance and Model Uncertainty

Compared to other studies, our random forest regression model demonstrated superior performance. Ibanez et al., (2021) finds with an MAE of 2.9, 5.1, and 6.7m for 30-day, 90-day, and 180-days ahead using deep neural networks-multivariate method. When compared to this study, random forest regression has an error (average first month MAE across all reservoirs) of 1m (1.41m average across 6-months) a 65.5% increased improvement in the first month prediction alone. There are also significant improvements for the 90-day forecast when compared to the 3-month forecast horizon and the 120-day forecast when compared to the 4-month forecast horizon (1.7m vs. 5.1m and 1.8 vs. 6.7m) for a 66.7% and 73.1% increase in water level prediction. The size, location, and other attributes of the reservoirs could be a reason for the large variation of error values as well as the diversity in methods to carry out the predictions.

There are several factors that contribute to model uncertainty including variability in satellite performance, complexity of hydrological interactions, and the sensitivity of the model to its input parameters. While the random forest regression showed promising results with low MAE, it is paramount to investigate potential sources of uncertainty. For all three regression techniques, data variability, natural or unnatural, introduces uncertainty especially over longer forecast horizons. Additionally, model sensitivity is critical to understanding the reliability of predictions.

The results of a sensitivity analysis for the random forest regression model with k=35 highlight the model's robustness and susceptibility to input data perturbations. When the input variables (reservoir heigh anomaly, TWS, precipitation, and temperature) were perturbed by +/- 5%, +/- 10%, +/- 20%, the models performance exhibited varied degrees of degradation when measured by MAE. Table 5 showcases the outputs from the sensitivity analysis. A 5% perturbation of the input resulted in a 22-25% increase in the baseline MAE, 1.41m. This impact escalated with greater perturbations, showing a 37-43% increase for 10% perturbations. Finally, the greatest perturbations showed the greatest increases with MAE reaching 2.50m and a large increase of 69-77% for 20%. These findings indicate that the model is relatively sensitive to changes in the overall input data, before it is tabulated. When perturbated after being tabulated the MAE only changed to 1.45 m with a maximum of a 2.8% increase for 20% due to having so many features and the numbers being very small and the model being robust to these small changes.

Table 2: Sensitivity Analysis. This table shows the percentage that the initial input variables perturbation, the resulting MAE, and the percent increase from 1.41m

Input Perturbation	Updated MAE	Increase in MAE
-20%	2.39	69%
-10%	1.93	37%
-5%	1.72	22%
5%	1.77	25%
10%	2.02	43%
20%	2.50	77%

For the overall changes, the increasing trend in MAE with larger perturbations suggests that the model relies heavily on the accuracy of input variables to maintain performance, thus emphasizing the importance of ensuring high-quality data to receive reliable predictions. This analysis reveals potential vulnerabilities in the performance of the model under different scenarios of data, highlighting the need for rigorous data validation.

Accuracy of Forecast

Upon determining that *k*=35 is the optimal Random Forest model, the average MAE across the entire forecast horizon (6 months) is 1.41m. This value is significantly lower than both the linear regression and polynomial regression and will be further explored in the next section. Additionally, the error for the first forecast horizon (1 month) is 1m, indicating that for the first month, the model predicted the reservoir height anomalies within one meter. As the forecast extends, the error also grows, reaching approximately 1.75m when forecasting six months ahead for the model.

Figure 4 shows the three different models plotted across the six-month forecast horizon for the various reservoirs. For the Frunas reservoir, month 1 is 7/2021 and month 2 is 8/2021 and so on. These results show the model forecast based on the training data the models were given and then the figure shows the output from what the models predicted to the actual values showing a testing aspect of the models and their capacity to predict across the six-month forecast horizon.

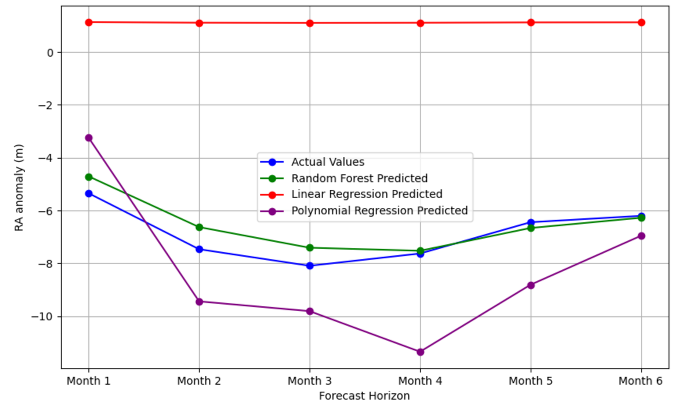

Figure 4b: Actual vs. Predicted values for the Frunas Reservoir: Individual reservoir model comparison plots. This figure shows the three different model outputs for each reservoir compared to the actual values. The plots for all the reservoirs are in supplementary materials and represent the last 6 months of validated data for each reservoir. This plot shows values for 7/2021 to 12/2021.

The random forest regression model most accurately predicted the height anomaly compared to the polynomial regression and the linear regression model.

Discussion

This study explores the use of terrestrial water storage (TWS) from GRACE to predict reservoir operations in the Upper Paraná River Basin. Analyzing 17 reservoirs, we compared three regression approaches, finding that random forest regression significantly outperformed linear and polynomial models by capturing complex relationships between TWS and reservoir height.

While linear and polynomial regression offer interpretability, they assume fixed relationships that limit their performance. Random forest regression, despite requiring more computational resources, proved more accurate, reducing mean absolute error (MAE) by nearly 50% compared to linear regression. Feature engineering, incorporating precipitation, temperature, and lagged reservoir heights, enhanced model performance. Error increased with forecast horizon, from 1m in the first month to 1.9m at six months, highlighting the challenge of long-term predictions. Feature analysis confirmed that previous reservoir height and TWS are key predictors, though feature importance varies by reservoir and time span. Future improvements could involve dynamic feature selection and additional data sources, such as land use or socio-economic factors.

This study demonstrates the potential of integrating TWS data for multi-step reservoir forecasting, supporting optimized water management, flood control, and sustainable resource use. Further refinement with real-time data and advanced modeling techniques could enhance predictive accuracy.

Conclusions

This study improves reservoir height predictions by integrating terrestrial water storage (TWS) estimates with multi-step forecasting across various reservoirs. Using GRACE TWS, IMERG precipitation, GLDAS temperature, Hydrolakes attributes, and radar altimetry data, we demonstrated that TWS effectively predicts reservoir height in the Upper Paraná Basin. The methodology can be applied globally. Machine learning models benefit from feature selection and importance analysis, which enhance interpretability and improve prediction accuracy. Understanding key features aids in refining model performance and supporting reservoir management decisions. Despite advancements, predicting reservoir height remains challenging due to human-driven operations. Future research should refine models by incorporating real-time data, expanding datasets, and exploring advanced techniques to improve predictions of storage volume, inflow, and outflow. Machine learning-based forecasting can enhance water resource management, flood control, and drought preparedness, contributing to sustainable water governance.

Acknowledgments

There are no conflicts of interest by the authors. We sincerely acknowledge funding from the NASA terrestrial hydrology program (Manager Jared Entin) and NSF GRFP (Grant # 1842490), and the Virginia Space Grant Consortium (VSGC).The data used in this study is publicly available for everyone and can be accessed at the links in table 1. The codes used can be made available on request. Please reach out to corresponding author, Jessica Besnier (jb8qv@virginia.edu), if interested. GREALM can be accessed at https://ipad.fas.usda.gov/cropexplorer/global_reservoir/ , DAHITI can be accessed at https://dahiti.dgfi.tum.de/en/products/water-level-altimetry/ , Hydroweb can be accessed at https://hydroweb.theia-land.fr/ , GRACE can be accessed at (https://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/ , GPM IMERG can be accessed at https://gpm.nasa.gov/data/directory , GLDAS can be accessed at https://disc.gsfc.nasa.gov/datasets/GLDAS_CLSM025_DA1_D_2.2/summary , and Hydrolakes at https://www.hydrosheds.org/products/hydrolakes#downloads . We would like to thank Christian O Leary (Nimbus Research Center) for his assistance with choosing and improving machine learning models and for general assistance and advice throughout the project. We would also like to thank the Hydrosense research group (University of Virginia) for their questions and critiques of the work through all stages that contributed to making the study more relevant.

Open Research

All the data used in this study is publicly available and outlined in the data table with links provided in the acknowledgements. The codes needed to reproduce this work are available and can be shared on request.

References

Bates, B. (2008, June 1). *Climate change and water*. http://www.taccire.sua.ac.tz/handle/123456789/552

Ibañez, S. C., Dajac, C. V. G., Liponhay, M. P., Legara, E. F. T., Esteban, J. M. H., & Monterola, C. P. (2021). Forecasting reservoir water levels using deep neural networks: a case study of Angat Dam in the Philippines. *Water*, *14*(1), 34. https://doi.org/10.3390/w14010034

Li, B., Yang, G., Wan, R., Dai, X., & Zhang, Y. (2016). Comparison of random forests and other statistical methods for the prediction of lake water level: a case study of the Poyang Lake in China. *Nordic Hydrology*, *47*(S1), 69–83. https://doi.org/10.2166/nh.2016.264

Meng, Y., Zhang, H., Jiang, P., Guan, X., & Yan, D. (2021). Quantitative assessment of safety, society and economy, sustainability benefits from the combined use of reservoirs. Journal of Cleaner Production, 324, 129242. https://doi.org/10.1016/j.jclepro.2021.129242

Połomski, M., & Wiatkowski, M. (2023). Impounding Reservoirs, Benefits and Risks: A review of environmental and technical aspects of construction and operation. Sustainability, 15(22), 16020. https://doi.org/10.3390/su152216020

Sapitang, M., Ridwan, W. M., Kushiar, K. F., Ahmed, A. N., & El-Shafie, A. (2020). Machine Learning application in Reservoir Water level Forecasting for Sustainable Hydropower Generation Strategy. *Sustainability*, *12*(15), 6121. https://doi.org/10.3390/su12156121

Sun, A. Y., Scanlon, B. R., Save, H., & Rateb, A. (2021). Reconstruction of GRACE total water storage through automated machine learning. Water Resources Research, 57(2). https://doi.org/10.1029/2020wr028666

Yin, J., Slater, L. J., Khouakhi, A., Yu, L., Liu, P., Li, F., Pokhrel, Y., & Gentine, P. (2023). GTWS-MLrec: global terrestrial water storage reconstruction by machine learning from 1940 to present. *Earth System Science Data*, *15*(12), 5597–5615. https://doi.org/10.5194/essd-15-5597-2023