Assessing the Potential of Integrating Hydrological Modelling Outputs with Logan River Observatory Gauge Datasets to Improve Anomaly Detection of Streamflow Data in the Franklin Sub-watershed

Ehsan Kahrizi[1]

[1]*Department of Civil and Environmental Engineering, Utah Water Research Laboratory, Utah State University (Ehsan.kahrizi@usu.edu)*

Abstract: This study assesses the potential for integrating hydrological modeling outputs with gauge-based observational datasets to enhance anomaly detection within aquatic sensor records. Using the Advanced Terrestrial Simulator (ATS), a hydrological simulation was performed for the Franklin sub-watershed of the Logan River watershed. Publicly available datasets supported detailed surface and subsurface configurations, including Daymet meteorological forcing, SSURGO soils, GLHYMPS geology, and NLCD land cover. Although convergence issues restricted simulation beyond the first year, the modeled results captured the general seasonal trends in streamflow dynamics, including timing and magnitude of peak flows related to snowmelt and precipitation. Comparisons between observed and simulated discharges demonstrated the ability of the model to fill observational gaps and validate spikes, suggesting its utility in distinguishing hydrologic events from sensor-induced anomalies. Flow duration curve comparisons revealed that the model replicates the overall flow regime structure, with systematic overestimations at mid- and high-flow ranges, partially attributed to simplified subsurface representations that do not account for the karstic characteristics of the watershed. Quantitative performance metrics, including R^2, NSE, RMSE, PBIAS, and KGE, indicated that the simulation would require further calibration for precise forecasting; however, even with a moderate statistical agreement, the model verified valuable for supporting quality control processes. The results show that hydrological models can operate as a physically based benchmark for validating sensor data and improving the reliability of observational datasets critical for hydrologic forecasting and management.

Keywords*: Hydrological Modelling; Anomaly Detection; Gauge Datasets; Logan River Observatory; Franklin Sub-watershed*

1 Introduction

Hydrological models are important in understanding and managing water resources by simulating water bodies and hydrological processes and forecasting hydrological events. The accuracy and reliability of these models rely on the quality and precision of their input datasets, including forcing data, land cover, soil texture, hydroclimate data, topography, bedrock properties, and, significantly, observational data from in-situ sensors (Gibert et al., 2018; Reitan, 2022). Across the United States, many gauges and monitoring stations continually collect diverse hydrological variables, which are essential both as direct inputs to hydrological models and for calibration purposes (Horsburgh et al., 2015). Therefore, observational datasets from these gauges are essential for ensuring the reliability of hydrological model outputs.

Due to advances in sensor networking and software system infrastructures, the availability of aquatic sensor data is rapidly growing (Adu-Manu et al., 2020; Jones, 2024; Pellerin et al., 2016; Rode et al., 2016). While the increasing diversity, volume, and resolution of datasets can enhance our understanding of hydrologic processes, they simultaneously introduce complexities in data management, interpretation, and utilization in modelling (Pellerin et al., 2016; Shen, 2018). Additionally, many of these raw datasets contain uncertainties, errors, or anomalies due to sensor calibration issues, maintenance artifacts, or data collection protocols. Involving such unreliable datasets in hydrological models can introduce substantial uncertainty, resulting in invalid or misleading predictions. This underscores the high demand for strong methodologies that ensure the reliability and accuracy of input data used in the hydrological modelling workflows.

Quality Control (QC) is one of the components of the data management workflow, ensuring reliable observational data is involved in subsequent analysis and modelling processes (Horsburgh et al., 2015). However, an inherent challenge in QC is differentiating sensor-induced anomalies from natural hydrological events. High-resolution and complex raw datasets from in-situ sensors often contain a mixture of both accurate hydrological patterns (e.g., natural variations, flood events, droughts) and sensor-related anomalies (e.g., calibration drifts, sensor fouling, infrastructural failures) (Jones et al., 2022; Shen, 2018). While several studies have explored methods for anomaly detection within hydrological datasets, fewer efforts have effectively addressed the challenge of distinguishing between sensor-induced anomalies and accurate hydrological signals in observational datasets. The lack of straightforward methods and studies for differentiating anomalies from true hydrological events hinders the development of reliable hydrological models and limits the utility of the models' outputs for environmental management and decision-making.

In response to these challenges, this study aims to evaluate the potential feasibility of integrating hydrological modelling outputs with observational datasets from the Logan River Observatory (LRO), specifically at the Franklin monitoring site within the Logan River watershed. The primary objective is to perform hydrological model simulations to better distinguish between sensor anomalies and genuine hydrological events. By employing the Advanced Terrestrial Simulator (ATS) model, this study will simulate expected natural streamflow conditions and compare these predictions against raw observational data collected by in situ sensors. Discrepancies between modeled streamflow predictions and sensor observations, identified through predefined threshold bands, can assist as indicators to distinguish anomalies in observational data from the hydrological processes effectively.

The anticipated outcomes of this research can potentially enhance the QC processes for aquatic observational datasets. By systematically distinguishing sensor-induced anomalies from natural hydrological patterns, the proposed methodology aims to improve the reliability and accuracy of observational datasets. Ultimately, this enhanced data reliability can substantially improve hydrological modelling outcomes, thereby increasing the precision and robustness of hydrological forecasts and simulations.

2 Materials and Methods

2.1 Case Study Area

The Franklin sub-watershed, as part of the Logan River watershed, was selected as the case study for this research (Figure 1). The Logan River originates in the Bear River Mountains, with headwaters near

the Utah-Idaho border at an elevation of approximately 2,900 meters. It flows through forest and rangeland and is impounded by three small reservoirs as it makes its way through Logan Canyon. The area of the Franklin sub-watershed is around 70 km^2. The mean elevation in this sun-watershed is around 2110 meters. The location of the observational gauge (latitude and longitude) is 41.9502 and -111.580553, respectively. The average daily discharge of the Logan River, measured near the outlet of Logan Canyon at the USGS gage (USGS 10109000 Logan River Above State Dam, near Logan, Utah), is 6.51 cubic meters per second (Jones et al., 2017; Neilson et al., 2021; Oleksy et al., 2022).

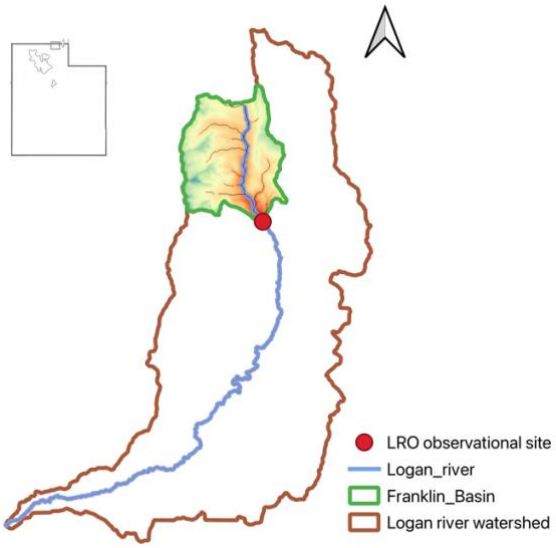


Figure 1 Logan River site map

2.2 Observational Datasets and Variables

The observational data, which are publicly available on the Observatory (LRO) website (http://lro.usu.edu), were used for this study. The data collection and preparation process includes recording raw data from in situ sensors and post-processing by technicians performed at the Utah Water Research Laboratory (UWRL). Sensors are connected to dataloggers that continuously record measurements every 15 minutes for aquatic and climate monitoring sites. Aquatic monitoring sites record various environmental variables and additional diagnostic information, such as battery voltage, to ensure proper sensor functionality. Similarly, climate monitoring sites record multiple meteorological variables and diagnostic information to monitor operational status. To provide easy access, both raw and corrected data are systematically organized into comma-separated values (.csv) files and posted to the HydroShare repository. Each CSV file contains metadata that provides contextual information about the recorded values.

The installed in-situ sensors collect different variables, such as water temperature, velocity, pH, dissolved oxygen, water level, and turbidity. While measuring the streamflow directly is impossible, the velocity and stage (water level) measurements are used to calculate river discharge (Figure 2). Discharge, representing the volume of water flowing through a river cross-section per unit time, can be estimated using a rating curve (Figure 3). The rating curve is an empirical relationship between the observed stage and the measured discharge, which will be created, and the discharge can be obtained (Turnipseed et al., 2010). After initializing the rating curve, the corresponding discharge can be accurately determined by continuously monitoring the water stage with *in-situ* sensors and referencing the established rating curve. This method provides an efficient and reliable way to quantify streamflow, especially when direct cross-sectional area measurements are not feasible at every time step.

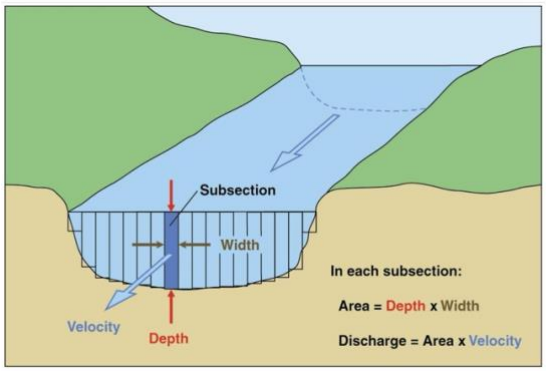


Figure 2 Cross-section of a river (Turnipseed et al., 2010)

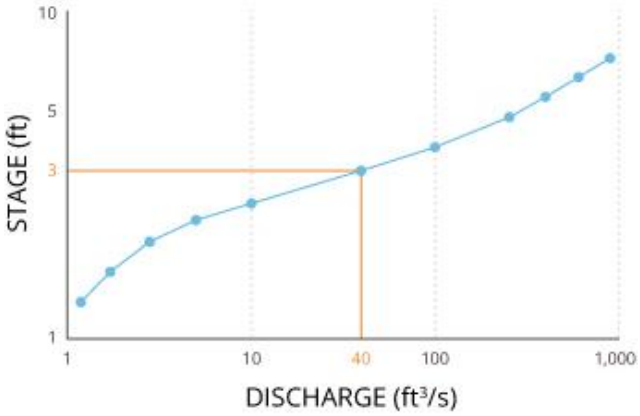


Figure 3 Stage-discharge rating curve

2.3 Simulation Inputs and Settings

2.3.1 Forcing datasets and available products

To implement the hydrological simulations in this study, different public datasets were employed to provide the necessary forcing and input data for my model setup. Meteorological inputs, which include temperature and precipitation, were retrieved from the Daymet Version 4 dataset hosted by the Oak Ridge National Laboratory (https://daymet.ornl.gov). The Leaf Area Index (LAI) data, which is required for characterizing vegetation cover and evapotranspiration processes, was obtained from the NASA's MODIS product, which is available on the EarthData portal (https://earthdata.nasa.gov). Also, the soil properties, like soil texture and hydraulic characteristics, were retrieved from the USDA's National Resources Conservation Service (NRCS) SSURGO database (https://datagateway.nrcs.usda.gov). To consider the subsurface geological layer, the required geologic data were retrieved from the Global Hydrogeology Maps (GLHYMPS) project (https://doi.org/10.5683/SP2/DLGXYO), but the soil thickness information was extracted from the SoilGRIDS global soil information system (https://www.isric.org/explore/soilgrids).
Additionally, the topographic data for model discretization, like watershed delineation, was derived from the National Elevation Dataset (NED) provided by the U.S. Geological Survey (USGS) (https://apps.nationalmap.gov/downloader/#/). Further, the stream and river network geometries were mixed from two different sources, the USGS National Hydrography Dataset (NHD) and NHDPlus High-Resolution datasets (https://apps.nationalmap.gov/downloader/#/). Land cover characteristics were retrieved from the National Land Cover Database (NLCD) (https://www.mrlc.gov), which allows the assignment of vegetation types and associated surface parameters in my model. These datasets provided the necessary environmental, geological, and meteorological information to analyze and run my model.

2.3.2 Temporal configuration

The Franklin sub-watershed within the Logan River Basin was delineated based on the Hydrologic Unit Code (HUC) 16010203, located within HUC4 region 1601. Based on this HUC information, the workflow was streamlined by avoiding the need to download and search through all national HUC datasets. To define the temporal domain of the simulations, a two-year transient simulation was selected, which ranging from October 1, 2021, to October 1, 2023. The initial year was treated as a spin-up period and discarded, which ensured that only stabilized model outputs from the second year were used for analysis. Also, the simulation origin date was set to January 1, 1980, which corresponds to the starting date of the Daymet forcing data in order to maintain consistency across input datasets. Also, a four-year cyclic spin-up was considered to initialize subsurface states and improve model stability before the transient simulation period.

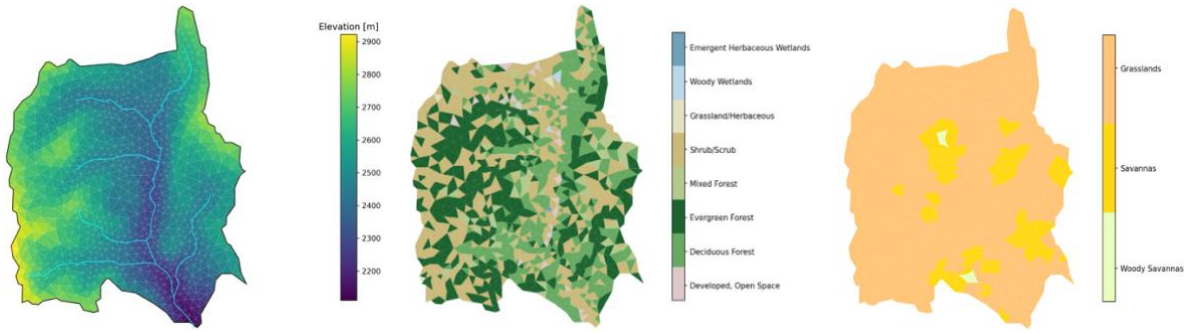
2.3.3 Surface configuration

Different steps were performed to create the surface mesh, as follows. The river network was extracted within the watershed bounds using the NHDPlusHR, and subsequently simplified to remove minor

tributaries and divergence artifacts. Small river networks with fewer than two reaches and reaches contributing less than 5% of the watershed area were pruned to maintain a cleaner hydrological network focused on the primary flow paths. The watershed boundary was simplified during the river network preparation, and river segments were snapped to watershed edges to ensure topological consistency. The meshing strategy applied distance-based refinement, wherein mesh elements closer to river reaches (within 300 meters) were refined to an average area of approximately 45,000 m², while elements farther than 700 meters were allowed to grow to around 125,000 m². The triangulation was generated while enforcing a minimum interior triangle angle of 32 degrees to avoid poorly shaped elements. The resulting mesh consisted of 884 points and 1640 triangular elements. The final visualization of the mesh is shown in Figure 4.I.

The NLCD raster was clipped to the watershed boundary and resampled onto the mesh element centroids. Each triangle in the mesh was assigned a corresponding land cover classification based on its centroid location. The mapped land cover types included categories like Grassland/Herbaceous, Evergreen Forest, Deciduous Forest, Shrub/Scrub, and Developed areas. An official NLCD color scheme was applied to visualize the spatial distribution of land cover classes across the watershed (Figure 4.II), providing essential input for defining vegetation and surface characteristics in the model simulations.

MODIS-based Land Use Land Cover (LULC) and Leaf Area Index (LAI) datasets were retrieved for land surface characterization. Both datasets were clipped to the watershed boundary and resampled onto the mesh element centroids. A custom MODIS land cover color map was applied to visualize the spatial distribution of dominant vegetation types, including Grasslands, Savannas, and Woody Savannas (Figure 4.III), which were used for model parameterization of surface processes.



I) Surface meshes

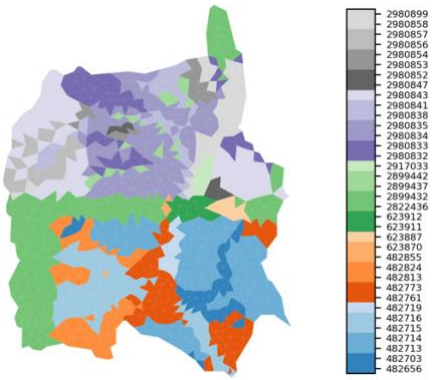
II) Land Cover Classification Mapping

III) MODIS Land Cover and LAI Mapping

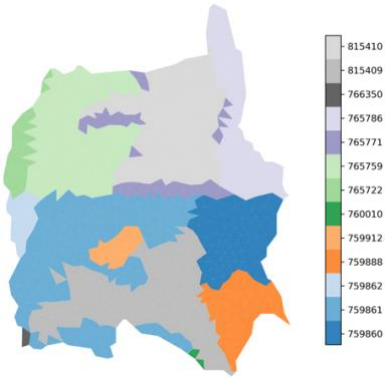
Figure 4 Surface configuration

2.3.4 Sub-surface configuration

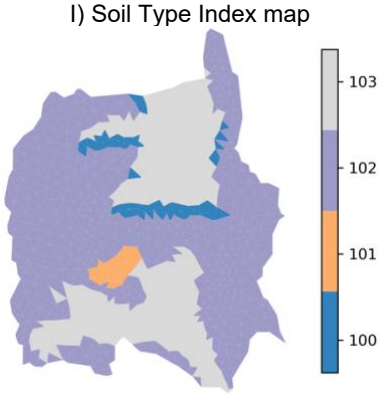
The subsurface configuration was developed by integrating soil and geological datasets to define the structure and material properties beneath the surface mesh. Soil characteristics were extracted from the SSURGO database, which assigned each mesh element a unique soil type based on its NRCS "mukey" attribute, which results in the spatial distribution shown in the Figure 5.I. Where SSURGO data was incomplete, geological formations from the GLHYMPS v2 dataset were used to supplement the subsurface properties, capturing deeper subsurface conditions through permeability and porosity assignments, as illustrated in the Figure 5.I. After cleaning and re-indexing the geology classifications to align with modeling requirements, the Figure 5.III was generated to represent the final geologic material structure across the watershed. Finally, depth-to-bedrock (DTB) information was used from the SoilGrids database, which maps the vertical extent of soil columns before encountering bedrock and visualized in the Figure 5.IV. These layers represented the subsurface conditions for my model.



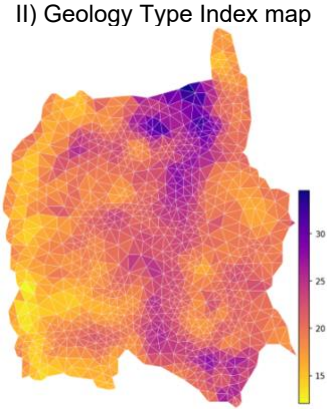
I) Soil Type Index map



II) Geology Type Index map



III) Geologic Type Index map

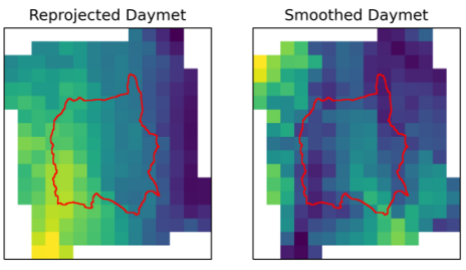


IV) Bedrock Depth (DTB) map

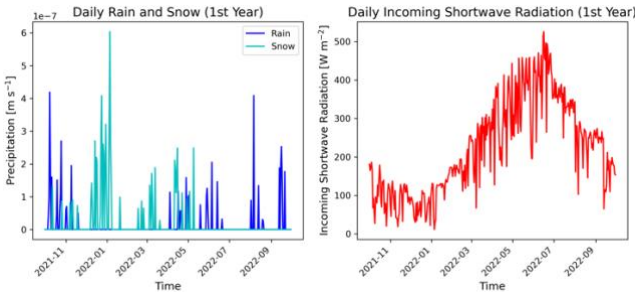
Figure 5 Sub-surface configuration

2.3.5 Meteorological Forcing Data

The spatial and temporal meteorological forcing data used in the simulations were evaluated prior to model input. The reprojected and smoothed Daymet datasets are shown in Figure 6.I. The spatial smoothing was applied to reduce grid noise and improve consistency across the watershed boundary. Also, to assess the temporal patterns, Figure 6.II presents the daily precipitation partitioned into rain and snow components and the daily incoming shortwave radiation during the first simulation year. These plots highlight the seasonal variation in meteorological inputs with snow events concentrated in winter months and peak solar radiation occurring during the summer.



I) Daymet dataset



II) Daily precipitation

Figure 6 Meteorological dataset

2.4 Metrics

Several widely used hydrological performance metrics were computed to evaluate the agreement between the observed and simulated streamflow data. The coefficient of determination (R^2) was used to assess the proportion of variance in observed discharge explained by the simulation (Equation 1). Also, the Nash-Sutcliffe Efficiency (NSE) evaluated the predictive skill of the model relative to the observed mean discharge, where values close to 1 indicate high model performance and negative

values suggest that the observed mean provides a better estimate than the model (Equation 2). Additionally, Root Mean Square Error (RMSE) was calculated to quantify the average error magnitude between simulated and observed discharges, emphasizing more significant errors (Equation 3). Percent Bias (PBIAS) measured the average tendency of the simulated data to overestimate or underestimate observations, with positive values indicating overprediction (Equation 4). Finally, the Kling-Gupta Efficiency (KGE) provided a composite metric incorporating correlation, bias, and variability errors, offering a holistic view of model accuracy (Equation 5).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Q_{\text{obs},i} - Q_{\text{sim},i}\right)^2}{\sum_{i=1}^{n}\left(Q_{\text{obs},i} - \overline{Q_{\text{obs}}}\right)^2} \quad \text{Equation 1}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Q_{\text{obs},i} - Q_{\text{sim},i}\right)^2}{\sum_{i=1}^{n}\left(Q_{\text{obs},i} - \overline{Q_{\text{obs}}}\right)^2} \quad \text{Equation 2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_{\text{obs},i} - Q_{\text{sim},i}\right)^2} \quad \text{Equation 3}$$

$$PBIAS = 100 \times \frac{\sum_{i=1}^{n}\left(Q_{\text{sim},i} - Q_{\text{obs},i}\right)}{\sum_{i=1}^{n} Q_{\text{obs},i}} \quad \text{Equation 4}$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad \text{Equation 5}$$

Where; $Q_{\text{obs},i}$ = observed discharge at time step i, $Q_{\text{sim},i}$ = simulated discharge at time step i, $\overline{Q_{\text{obs}}}$ = mean observed discharge, r = correlation coefficient, β = bias ratio, and γ = variability ratio.

3 Results and Discussions

The hydrological simulation, which was conducted using the ATS over the Franklin sub-watershed, provided valuable insights into both model performance and its role in improving anomaly detection within observational datasets. Notably, after many attempts, my model could not run for the second year of my study due to the convergence issue, and I presented my results only for the first year. A comparison of observed and simulated daily discharge (Figure 7) demonstrates that the model reasonably captures the general seasonal trends of streamflow dynamics. The simulated data successfully reflects the timing and magnitude of peak flows during spring and early summer, aligning with snowmelt and precipitation events. This agreement at the seasonal scale indicates that the surface and subsurface configurations and meteorological forcings effectively represented the watershed's hydrological processes.

However, several discrepancies between observed and simulated discharge were noted. During periods of missing observational data (highlighted in gray), the simulation provides a continuous discharge estimate, offering a valuable tool for gap-filling and improving the quality control (QC) process for sensor datasets. Moreover, specific sharp anomalies, such as spikes visible in the observed discharge data (highlighted in red), were visible in the simulated discharge, meaning that such spikes are likely due to hydrological responses like high-intensity precipitation rather than sensor malfunctions. Thus, hydrological simulations can serve as a benchmark to distinguish true events from observational artifacts, enhancing the reliability of corrected datasets for future analyses.

The flow duration curve shown in Figure 8 further evaluates the performance of the model by comparing the discharge exceedance probabilities of observed and simulated data. Both curves display similar general shapes, suggesting the model captures the flow regime structure reasonably well across different percentiles. However, systematic overestimation of flow volumes is visible, particularly in the mid- and high-flow ranges. This bias could stem from several factors, such as uncertainties in the meteorological forcing, imperfect representation of snowmelt processes, or subsurface parameterization. Nevertheless, the model's ability to replicate the general discharge variability makes it useful for qualitative interpretation and anomaly validation.

Further, in terms of the magnitude, the simulated results in the range of the peak flow (e.g., 2022-05-01) are overestimated. In contrast, those for the ranges after the peak flow (e.g., 2022-07-01) are underestimated. The reason for underestimation is related to the fact that the texture of the soil in the

simulation is considered heterogeneous, while in reality, the texture of the mountain in the Logan River watershed is karstic, which is characterized by numerous caves and complex underground structures.

Quantitative model performance metrics were assessed using a one-to-one scatter plot between observed and simulated discharge (Figure 9). The results show an R^2 of -1.48 and a Nash-Sutcliffe Efficiency (NSE) of -1.48, indicating poor statistical agreement on an event-by-event basis. The RMSE was relatively high (194,629 m^3/d), with a Percent Bias (PBIAS) of 73.9%, showing the model systematically overpredicts discharge. Kling-Gupta Efficiency (KGE) also showed negative performance (-0.13). Although these statistical measures suggest that further calibration and refinement would be needed for predictive applications, the model's broader seasonal patterns and discharge magnitudes still offer significant value for anomaly detection, which may be used to validate observational gaps and interpret uncertain periods in the sensor record.

Overall, integrating hydrological model outputs with gauge-based observational datasets can enhance the ability to validate, interpret, and correct anomalies. Even with imperfect simulation accuracy, hydrological modeling offers a physically based expectation against which suspicious observational data points can be evaluated, ultimately strengthening the reliability of environmental sensor networks for research and management purposes.

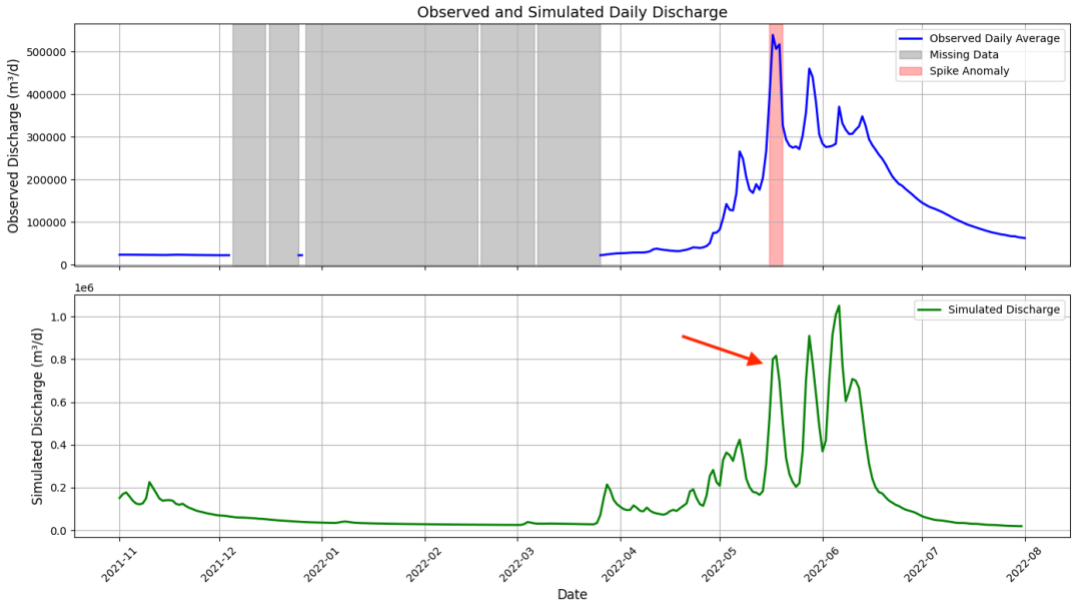


Figure 7 Comparison of observed and simulated daily discharge, with missing data periods (gray) and spike anomalies (red) highlighted in the observed dataset

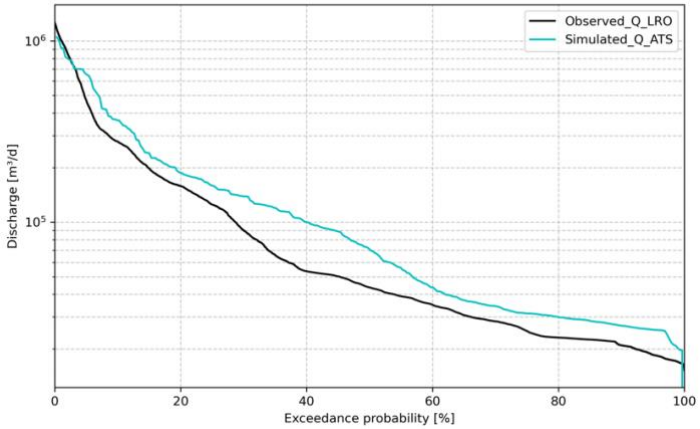


Figure 8 Flow duration curve comparing observed and simulated daily discharge based on exceedance probability for the Franklin sub-watershed.

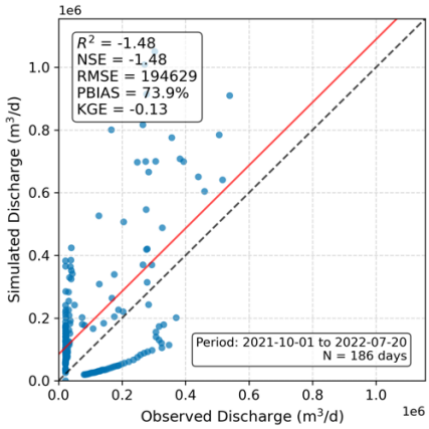


Figure 9 One-to-one scatter plot comparing simulated and observed daily discharge values, including model performance metrics (R^2, NSE, RMSE, PBIAS, KGE) for the evaluation period at the Franklin sub-watershed.

4 Conclusions

This project assessed the feasibility of integrating hydrological model outputs with gauge-based observational datasets to improve anomaly detection and data reliability for the Franklin sub-watershed. Using the ATS model, the simulated streamflow successfully captured the general seasonal patterns and peak timing observed in the field data. However, notable flow magnitudes and statistical metrics discrepancies. The hydrological simulation demonstrated significance for validating missing data periods and distinguishing hydrologically plausible anomalies, like sharp peaks related to natural precipitation events, from sensor-induced errors.

While the model tended to overestimate discharge volumes systematically and showed poor statistical agreement on an event-by-event basis, it provided a continuous physical reference that enhanced quality control workflows. Limitations included convergence challenges restricting the simulation period to one year and soil structure assumptions that did not fully represent the karstic nature of the watershed. Future improvements could focus on refining subsurface characterization, enhancing forcing data accuracy, and implementing model calibration procedures.

In general, despite these limitations, integrating hydrological modeling with observational data indicates significant potential for supporting real-time sensor network validation, improving dataset reliability, and ultimately advancing hydrological research and watershed management efforts.

Data Availability Statement

The datasets and software used in this study are publicly available to ensure transparency and reproducibility. The Python programming language was utilized within Jupyter Notebooks for data processing, plotting, and analysis. Additionally, the simulations were performed using the Advanced ATS model. The complete codes and processed datasets supporting this study are available on GitHub at: https://github.com/Ehsankahrizi/hydrologic-modeling-course-2024/tree/capstone_project/capstone-projects/Ehsan_Kahrizi
Additionally, larger complementary files and simulation outputs have been deposited and are accessible through Zenodo. The Zenodo link is provided within the GitHub repository for public access.

Acknowledgments

I want to thank Dr. Jeff Horsburgh for his valuable guidance in selecting the appropriate case study for this study and for his insightful support in shaping the project's core idea.

References

Adu-Manu, K. S., Katsriku, F. A., Abdulai, J. D., & Engmann, F. (2020). Smart River Monitoring Using Wireless Sensor Networks. *Wireless Communications and Mobile Computing*, *2020*. https://doi.org/10.1155/2020/8897126

Gibert, K., Horsburgh, J. S., Athanasiadis, I. N., & Holmes, G. (2018). Environmental Data Science. *Environmental Modelling and Software*, *106*, 4–12. https://doi.org/10.1016/j.envsoft.2018.04.005

Horsburgh, J. S., Reeder, S. L., Jones, A. S., & Meline, J. (2015). Open source software for visualization and quality control of continuous hydrologic and water quality sensor data. *Environmental Modelling and Software*, *70*, 32–44. https://doi.org/10.1016/j.envsoft.2015.04.002

Jones, A. S. (2024). *Water Data Science: Data Driven Techniques, Training, and Tools Water Data Science: Data Driven Techniques, Training, and Tools for Improved Management of High Frequency Water Resources for Improved Management of High Frequency Water Resources Data Data Recommended Citation Recommended Citation*. Utah State University. https://digitalcommons.usu.edu/etd2023/134

Jones, A. S., Aanderud, Z. T., Horsburgh, J. S., Eiriksson, D. P., Dastrup, D., Cox, C., Jones, S. B., Bowling, D. R., Carlisle, J., Carling, G. T., & Baker, M. A. (2017). Designing and Implementing a Network for Sensing Water Quality and Hydrology across Mountain to Urban Transitions. *Journal of the American Water Resources Association*, *53*(5), 1095–1120. https://doi.org/10.1111/1752-1688.12557

Jones, A. S., Jones, T. L., & Horsburgh, J. S. (2022). Toward automating post processing of aquatic sensor data. *Environmental Modelling and Software*, *151*. https://doi.org/10.1016/j.envsoft.2022.105364

Neilson, B. T., Tennant, H., Strong, P. A., & Horsburgh, J. S. (2021). Detailed streamflow data for understanding hydrologic responses in the Logan River Observatory. *Hydrological Processes*, *35*(8). https://doi.org/10.1002/hyp.14268

Oleksy, I. A., Jones, S. E., & Solomon, C. T. (2022). Hydrologic Setting Dictates the Sensitivity of Ecosystem Metabolism to Climate Variability in Lakes. *Ecosystems*, *25*, 1328–1345. https://doi.org/10.1007/s10021-021-0071

Pellerin, B. A., Stauffer, B. A., Young, D. A., Sullivan, D. J., Bricker, S. B., Walbridge, M. R., Clyde, G. A., & Shaw, D. M. (2016). Emerging Tools for Continuous Nutrient Monitoring Networks: Sensors Advancing Science and Water Resources Protection. *Journal of the American Water Resources Association*, *52*(4), 993–1008. https://doi.org/10.1111/1752-1688.12386

Reitan, T. (2022). *Review of methods for automation of quality control on hydrologic time series and considerations for a research approach at NVE*. Norwegian Water Resources and Energy Directorate. https://publikasjoner.nve.no/rapport/2022/rapport2022_15.pdf

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., Arhonditsis, G. B., Jordan, P., Kronvang, B., Halliday, S. J., Skeffington, R. A., Rozemeijer, J. C., Aubert, A. H., Rinke, K., & Jomaa, S. (2016). Sensors in the Stream: The High-Frequency Wave of the Present. *Environmental Science and Technology*, *50*(19), 10297–10307. https://doi.org/10.1021/acs.est.6b02155

Shen, C. (2018). A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. In *Water Resources Research* (Vol. 54, Issue 11, pp. 8558–8593). Blackwell Publishing Ltd. https://doi.org/10.1029/2018WR022643

Turnipseed, D. P., Sauer, V. B., & Survey, U. S. G. (2010). Discharge measurements at gaging stations. In *Techniques and Methods*. https://doi.org/10.3133/tm3A8

Appendix (plots)

Further outputs of the simulation for the Franklin sub-watershed:

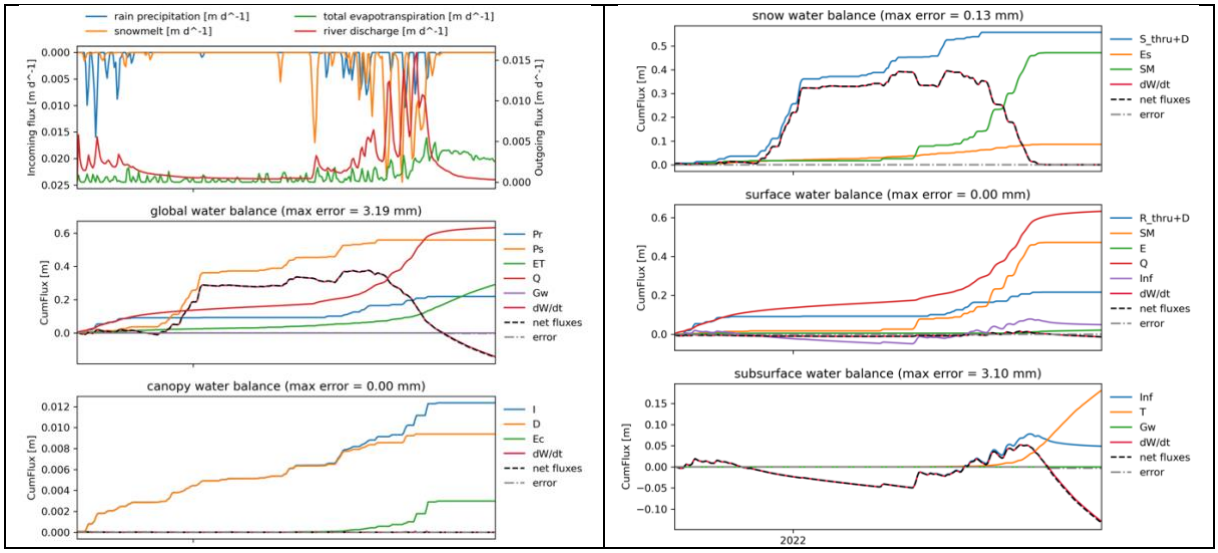
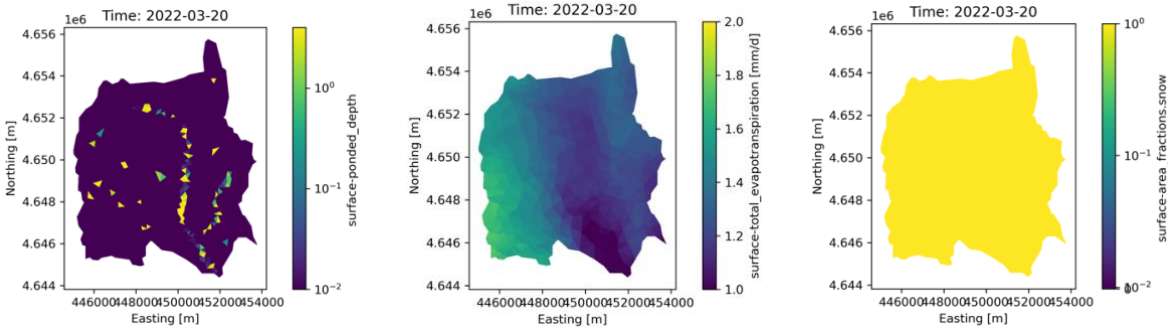
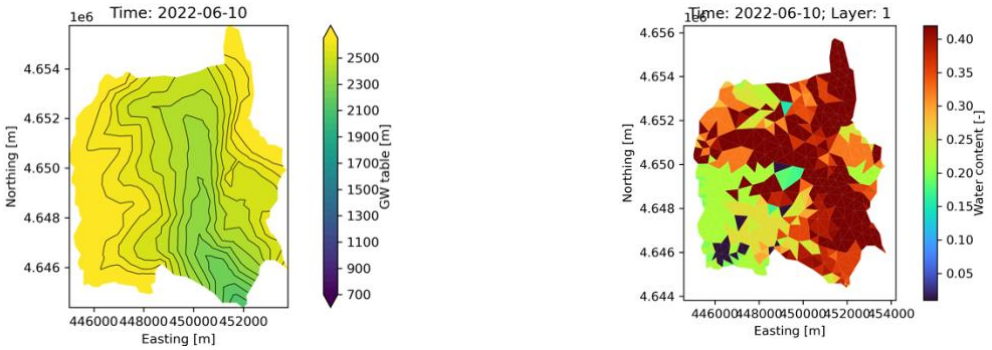


Figure 10 Water balance plots



I) Surface ponded depth II) Surface-total evapotranspiration III) Snow cover

Figure 11 Surface plots



I) Groundwater surface II) Volumetric water content

Figure 12 Sub-surface plots