Paleo Streamflow Reconstruction Methods and Their Impact on Water Resources Management & Decision-Making

Prepared For: Dr. Shao, GY 552: Environmental Decision Making

Prepared By: Madi Brake

Date: December 12, 2023

Table of Contents

1. Introduction

Water resources and planning management decisions are generally based on observed records of climate variables such as temperature, precipitation, soil moisture, and streamflow. Consequently, water resources are oftentimes misallocated due to such short periods of observation, particularly with streamflow records. Numerous sectors from flood control to energy production are reliant upon historic streamflow data indicative of variation of water availability over time to make informed management decisions. In his article, "*Dendrochronolgy and links to streamflow*," D.M. Meko states that streamflow variability records on timescales of decades to centuries becomes increasingly critical for water management, as resource shortages are imposed by increasing demand and limited supply (Meko et al., 2010).

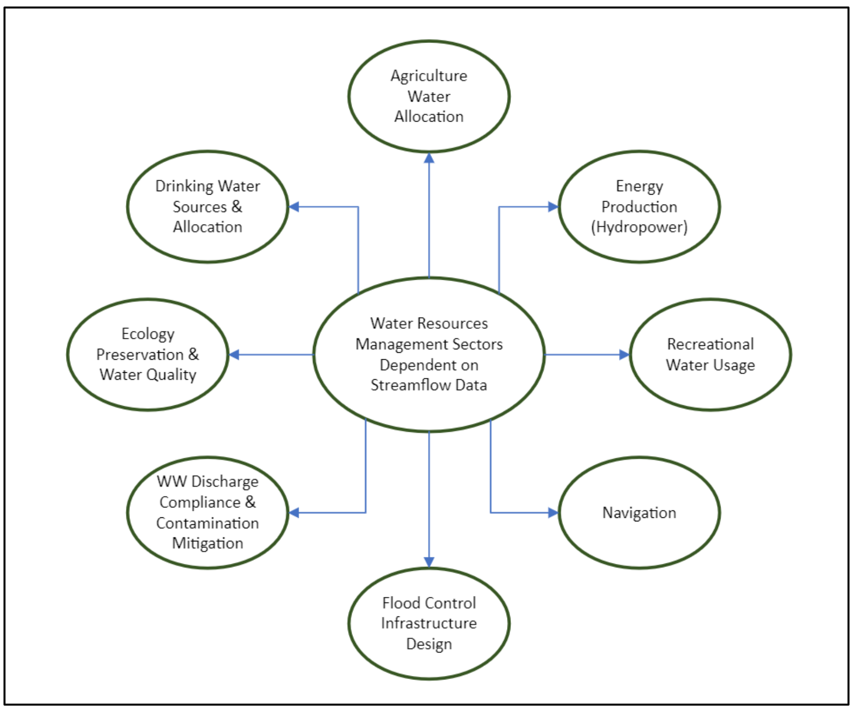

Figure 1: Water resources management sectors dependent on historic streamflow records and insight into water availability variation over time.

1.1 Research Background and Goals

I was first introduced to paleohydrology while studying abroad in Italy in the summer of 2023. Here, I explored traditional reconstruction techniques of various gauges within the Adige River Basin of northern Italy. Continuing this spatial focus for my master's degree, my current research focuses on the comparison between traditional reconstruction techniques and novel machine learning methods. By reviewing previous reconstructions of streamflow, the methods yielding success, and the methods not already explored, the goals of conducting my original research aligned as follows: analyze spatial and temporal variation of water availability over time, compare traditional reconstruction techniques with ML algorithms, and to share my findings for operational hydrologic decision processes.

2. Literature Review

Long-term streamflow records are necessary to inform decision makers on the possible range of variability over time. Reconstructions using paleoclimatic proxies are among the most important tools for studying past climates, one of the most widely used being dendroclimatic proxies.

2.1 Tree-Ring Based Proxies

Dendrochronology is the scientific method of dating tree rings to the exact year they were formed within a tree, while chronologies of moisture sensitive trees are best indicative of precipitation, or lack thereof, over a specified time period (Tootle et al., 2023). Precipitation-based tree-ring chronologies are used as proxies to form drought indices, such as the Palmer Drought Severity Index (PDSI), which date back thousands of years (Ho et al., 2016). The Old-World Drought Atlas (OWDA) provides self-calibrating Palmer Drought Severity Index (scPDSI) values for 5,414 gridpoints across Europe from 0 to 2012 AD (Formetta et al., 2021). Similarly, the Living Blended Drought Atlas (LBDA) provides gridded scPDSI values for the continental United States, for a maximum range of 365 to 2005 AD (Anderson et al., 2019). The self-calibrating PDSI modifies the way the PDSI is computed by automatically adjusting the climatic characteristic and calculating the duration factors based on the characteristics of the climate at a given location. Therefore, the scPDSI performs more consistently and allows for more accurate comparisons at different locations (Wells et al., 2004).

Such information from tree rings, like the gridded scPDSI data, has been used to reliably reconstruct a number of climate variables including streamflow on centennial to millennial scales (Cook et al., 2015). Additionally, dendroclimatic reconstructions provide information at higher temporal resolution than most other proxies, allowing modern climatic events to be placed into a longer-term context (Robeson et al., 2020).

2.2 Traditional Reconstruction Methods

Numerous streamflow reconstruction studies have been conducted using tree-ring based proxies with varying spatial and temporal extents. In her article, *"Can a paleodrought record be used to reconstruct streamflow?: A case study for the Missouri River Basin,"* Michelle Ho first explored using the LBDA, a gridded scPDSI dataset for the continental United States, to reconstruct streamflow. This method would yield effective reconstructions over the Missouri River Basin and would go on to influence streamflow reconstruction studies across the United States and eventually into Europe using the OWDA. The process begins with the collection of streamflow data, aggregated to the average discharge monthly or annually, depending on seasonality preference of the study, along with the gridded scPDSI cells within a 450-kilometer radius of the centroid of the study extents. Prescreening efforts include correlation and temporal stability analysis of the observed streamflow data against the scPDSI cells. The scPDSI cells passing >99% correlation and stability are then used as the predictors in a stepwise linear regression (SLR) or log-linear model, where streamflow is used as the predicted (Tootle et al., 2023).

Recent progress in this research has determined further sharpening of reconstruction skill with the use of quantile-mapping bias correction of the reconstructed streamflow vectors (Robeson et al., 2020). Robeson's quantile-mapping bias correction technique of paleoclimatic reconstructions was applied to the upper Colorado River, in which his results showed that several extreme events from the reconstructed streamflow were even more intense than formerly thought. Robeson concluded that bias correction should be applied wherever paleoclimatic reconstructions are compared directly to observations.

Past studies using the traditional reconstruction methodology from Ho's 2016 research have yielded highly skillful results for rivers across the globe. An example is the Sava River Basin in Slovenia, in which April-September seasonal streamflow was reconstructed from 0 to 2012 AD using scPDSI data and monthly aggregated streamflow observations from 2 gauges in the basin (Tootle et al., 2023). The results showed a high variability in streamflow over time for those gauges, giving insight into historical drought and pluvial durations and frequencies as compared to modern climate events. The variation of streamflow over time also allows water managers to evaluate the possible anthropogenic causes of modern events if unprecedented.

3. Original Research

The Adige River Basin (ARB) is the third largest catchment and the second longest river in Italy, encompassing two Italian regions (Trentino Alto-Adige and Veneto). The upper ARB is in northeastern Italy, where precipitation is highly variable across the basin, ranging from 500 to 1600 mm annually. Given the influence of the Italian Alps, runoff (streamflow) is snow-dominated with minimal flows during the winter months and higher flows during the spring-summer months due to snowmelt. The ARB is a vital water source in Italy for varying demands including energy [over 30 large (>3MW) hydropower plants] and agriculture (grapes and wine production). Given the recent drought impacting watersheds across northern Italy, the optimal allocation of water is an ongoing challenge in the ARB. An additional challenge is the lack of streamflow records, both spatially and temporally.

3.1 Data and Methods

Recent research efforts developed a calibrated hydrologic model of the ARB in which annual modeled streamflow was generated for 33 gauges with a period of record of 1980 to 2018. The instrumental period of 1980-2012 was used for the reconstructions, along with the self-calibrating Palmer Drought Severity Index (scPDSI) cell data within a 450-kilometer radius of the centroid of the ARB. Pre-screening was performed such that poorly correlated (less than 99%) scPDSI cells were eliminated. Additional pre-screening was done to ensure that correlation was stable for the overlapping period of record. The PDSI cells passing pre-screening were used as predictors (independent variables) while observed streamflow was used as the predicted (dependent variable) applying forward and backward Stepwise Linear Regression (SLR).

The ML reconstructions were performed through RapidMiner Studio. For each gauge, the instrumental record and 450-km radius scPDSI values were utilized for training and cross-validation, split into 2 random folds. The ML algorithms used to reconstruct streamflow at each of the 11 gauges were Generalized Linear Model (GLM) and Random Forest (RF). The output from RapidMiner Studio included the ML reconstructed streamflow and performance vectors including coefficient of correlation (R^2) and root mean squared error (RMSE). The SLR, GLM, and RF vectors were bias-corrected using quantile-mapping R code, developed by Giuseppe Formetta of the University of Trento.

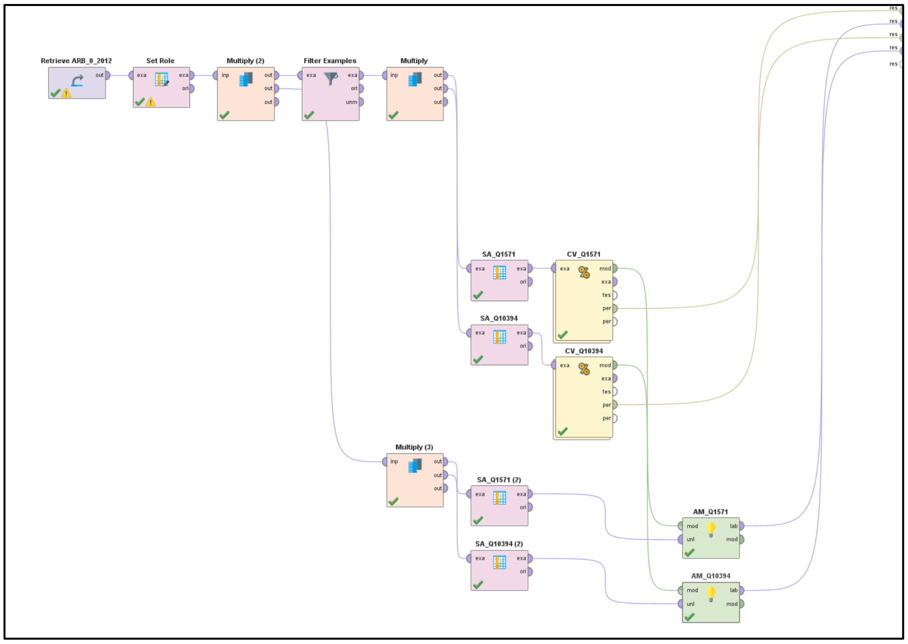

Figure 2: RapidMiner machine learning schematic used to apply GLM and RF algorithms for streamflow reconstruction of the ARB.

3.3 Results

Two gauges were chosen for the evaluation of method performance comparison, gauge 1571 and gauge 10394.

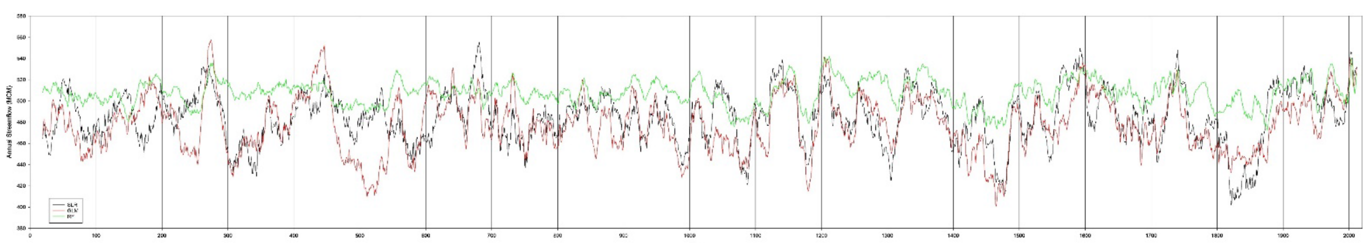

Figure 3: Gauge 1571- SLR (black) vs GLM (red) vs RF (green), with 20-year end filter (0-2012 AD)

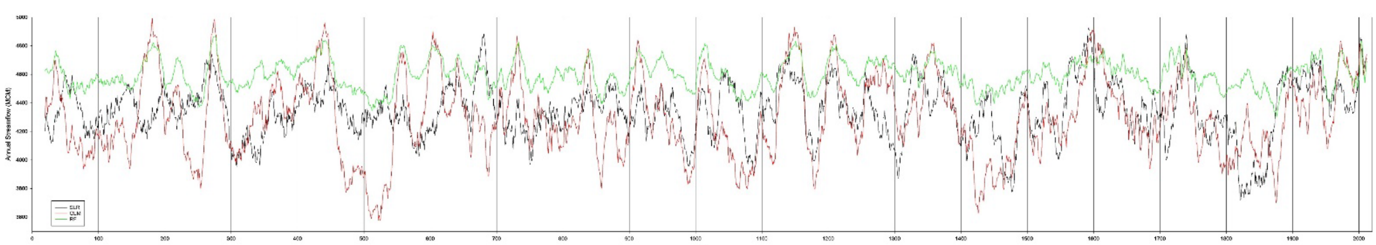

Figure 4: Gauge 10394 - SLR (black) vs GLM (red) vs RF (green) with 20-year end year filter (0-2012 AD)

The RF ML algorithm performed the worst, not capturing extremes or the range of variability that SLR and GLM convey. Comparing skill statistics for gauge 1571, the traditional method of streamflow reconstructed outperformed either machine learning algorithm drastically, with an R^2 of 0.61, compared to 0.168 and 0.137 for GLM and RF, respectively. The same dramatic skill increase was observed for SLR compared to GLM and RF for gauge 10394.

4. Discussion

The results from traditional SLR reconstructions yielded highly skillful reconstructions of streamflow in the Adige River Basin from 0-2012 AD. Both machine learning methods, GLM and RF, showed poor performance when compared to SLR. GLM captured the historical pluvial and drought extremes similar to those captured using SLR, whereas RF did not. The performance of the ML methods is largely attributed to the small dataset of observed streamflow data, as the machine learning algorithms were trained and cross validated on a dataset that did not contain as extreme of variation as the paleo reconstructed record. Traditional SLR reconstructions seem to be most accurate in reflecting streamflow variation over time while capturing extremes, in the extent and available observed data for the Adige River Basin.

From the literature review, paleo streamflow reconstructions have been done successfully using the traditional stepwise linear regression or log-linear modeling-based methods. The use of machine learning algorithms to conduct paleo reconstructions of climate variables has not been widely studied, especially for streamflow. The use of machine learning algorithms for streamflow reconstruction must consider the robustness of the instrumental record to provide a skillful paleo reconstruction, consequently, the preliminary research comparing SLR vs ML techniques in the ARB heavily favored traditional methods.

4.1 Stakeholder Involvement

Water resources decisions, as mentioned previously, must consider a variety of sectors including risk management and water quality. These decisions made on water resources are based upon streamflow records that indicate variation in water availability. Reconstructed streamflow records offer insight into past pluvial and drought periods, putting modern climate events into a historical perspective. Stakeholders that utilize historic streamflow records in day-to-day operation include civil and environmental engineers, hydropower and flood infrastructure control authorities, and river forecasting centers. For example, wastewater discharge is only environmentally compliant if the receiving stream or creek meets a minimum baseflow, determined from historical records and long-term variability analysis. Additionally, flood control authorities rely heavily on historic streamflow data and paleo streamflow reconstructions for operational flood guides, balancing curves, and dam safety risk assessments. River and hydrologic forecasting centers utilize dendroclimatic reconstructions of streamflow to identify vulnerabilities in long-term reservoir operations. In summary, extending streamflow records past the instrumental period is crucial for informing stakeholders that make decisions on water resources.

5. Conclusions

Given the importance of river systems for varying demands, information about past drought and pluvial periods provide important information to water managers and planners. Streamflow reconstructions provide hundreds of years of insight into hydrologic cycling and climate factors, which constitute critical knowledge when making water policy decisions. Traditional reconstruction methods using tree-ring based proxies have been successful using SLR or log-linear modeling with bias correction. The use and implementation of novel ML techniques can provide a convenient and effective streamflow reconstruction method for water managers and planners, but with limitations like dataset size, seasonality preference, and reconstruction skill.

6. References

Anderson, S.; Ogle, R.; Tootle, G.; Oubeidillah, A. Tree-Ring Reconstructions of Streamflow for the Tennessee Valley. Hydrology 2019, 6, 34. https://doi.org/10.3390/hydrology6020034

D.M. Meko, C.A. Woodhouse, K. Morino, Dendrochronology and links to streamflow, Journal of Hydrology, Volumes 412–413, 2012, Pages 200-209, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2010.11.041.

Edward R. Cook et al., Old World megadroughts and pluvials during the Common Era.Sci. Adv.1,e1500561(2015).DOI:10.1126/sciadv.1500561

Formetta, G.; Tootle, G.; Bertoldi, G. Streamflow Reconstructions Using Tree-Ring Based Paleo Proxies for the Upper Adige River Basin (Italy). Hydrology 2022, 9, 8. https://doi.org/10.3390/hydrology9010008

Ho, M., Lall, U., and Cook, E. R. (2016), Can a paleodrought record be used to reconstruct streamflow?: A case study for the Missouri River Basin, Water Resour. Res., 52, 5195–5212, doi:10.1002/2015WR018444.

Robeson, S. M., Maxwell, J. T., & Ficklin, D. L. (2020). Bias correction of paleoclimatic reconstructions: A new look at 1,200+ years of Upper Colorado River flow. Geophysical Research Letters, 47, e2019GL086689. https://doi.org/10.1029/2019GL086689

Tootle, G.; Oubeidillah, A.; Elliott, E.; Formetta, G.; Bezak, N. Streamflow Reconstructions Using Tree-Ring-Based Paleo Proxies for the Sava River Basin (Slovenia). Hydrology 2023, 10, 138. https://doi.org/10.3390/hydrology10070138

Wells, N., S. Goddard, and M. J. Hayes, 2004: A Self-Calibrating Palmer Drought Severity Index. J. Climate, 17, 2335–2351, https://doi.org/10.1175/1520-0442(2004)017<2335:ASPDSI>2.0.CO;2.