**Advancing the Use of Artificial Intelligence and Machine Learning in the Development of Useable Tree-ring Proxy based Reconstruction of Streamflow**

**INSTRUCTIONS:** Synthesize at least 3 publications (e.g., journal articles, reports) into a 1-page summary of the project background. Publications referenced in the background must be from at least three different disciplinary perspectives. The background should present information explaining the context of the societal problem being addressed, synthesis of key research and practice documents describing past research accomplishments, and a plain terms explanation of the way the project will contribute to a practical solution.

**MADI:** Need for Project / Project Methods / SE United States Case Study / Hydrology

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>

Water resources and planning management decisions are generally based on observed historic records. Therefore it is crucial to investigate historical streamflow in order for water managers and planners to make better decisions on water resources. Although the Southeast United States is known to have more abundant water quantity than other regions in the US, having a long period of historical record can help stakeholders evaluate certain climatic factors that influence irregular periods of streamflow and water availability variability.

Proxies such as tree-rings from moisture sensitive trees can provide correlation with unimpaired streamflow. The issue in the Southeast US, especially in the Tennessee River Valley, is the lack of unimpaired streamflow gauges. TVA and the USACE operate numerous dams which impair the majority of the streamflow in the Tennessee River Valley. Although misconceptions still exist regarding the applicability of dendroclimatology in the southeast, tree rings in the region have been used to investigate the relationships between climate and tree-growth. This study reports the preliminary results from a statistical screening of tree-ring width records from the International Tree-Ring Data Bank (ITRDB), to evaluate the strength of the hydrological signal, in dendrochronological records from the Tennessee Valley. USGS streamflow data (observed) was used in correspondence with the ITRDB regional chronologies to analyze dendroclimatic potential and seasonal streamflow reconstructions within the valley. Results revealed the potential benefit of using tree-ring chronologies to reconstruct hydrological variables in the Southeastern U.S., by demonstrating the ability of proxy-based reconstructions to provide useful data beyond the instrumental record.

Utilizing methods similar to those found in this article and other articles by Tootle can be of great benefit for our NRT project, although it is clear from the article that utilizing tree-ring proxies poses a greater difficulty for the southeastern US as compared to other US regions and international studies. Methods include gathering USGS streamflow data (observed) for gauges in question (unimpaired) condensed into monthly or annual flows in million cubic meters (MCM). Using the ITRDB or another tree-ring proxy, the observed flows can be correlated with the proxy cells, in which those with significant correlation are retained (P <= 0.05). Additional pre screening methods include temporal stability analysis of the datasets. Stepwise linear regression is used on the prescreened data to develop the reconstructed streamflow vectors for each gauge. Skill statistics are calculated to ensure the validity of the results.

Future collections of new tree ring proxies would likely increase the statistical skill of the reconstructions and, perhaps, increase or lengthen the season (i.e., May–June–July) of the streamflow reconstruction, providing increased information on past water availability.

**MAHSA:** Dendroclimatology/ Dendrohydrology/ Streamflow Reconstruction

This project addresses the challenge of limited historical streamflow records, which often cover only short periods of time (typically 50 to 100 years). This limitation hinders the ability of water managers and planners to make informed decisions and develop effective policies. To overcome this challenge, the project proposes to use tree-ring proxies to extend streamflow records, offering a longer historical perspective on droughts and pluvial periods. Some of tree specious such as Baldcypress trees and oak trees are known to form false rings and flood rings respectively, which are anomalous growth patterns that can be used as proxies for investigating past hydrological and climate conditions. These growth anomalies offer insights into patterns of long-term climate trends and high flow periods. The project aims to develop streamflow reconstructions using both traditional regression techniques and innovative AI/ML approaches and evaluate extreme events. To provide a comprehensive understanding of the project's background, we'll synthesize information from a few different publications.

Tootle’s study focuses on the Sava River Basin in Europe, providing an overview of the region and its importance for various purposes, including energy production. The research utilizes tree-ring proxies and the Old-World Drought Atlas (OWDA) to reconstruct seasonal streamflow data for approximately 2000 years. It identifies both drought and pluvial periods in the reconstructed historical record, with a significant decrease in pluvial periods compared to the paleo record, likely due to increased human activities like the construction of water control structures. This knowledge is of paramount importance for water managers and planners seeking practical solutions for sustainable water resource management.

Tootle, G., Oubeidillah, A., Elliott, E., Formetta, G., & Bezak, N. (2023). Streamflow Reconstructions Using Tree-Ring-Based Paleo Proxies for the Sava River Basin (Slovenia). *Hydrology*, *10*(7), 138.

Cleaveland’s article highlights the significance of long and homogeneous climate and streamflow records for accurate estimation of various statistical parameters. It emphasizes the impact of low flows caused by drought on water quality and associated challenges. The research employs baldcypress tree-ring data to reconstruct summer streamflow in the White River, Arkansas, covering an impressive 963-year period. The study demonstrates the potential for using tree-ring data to improve statistical estimates of flow parameters and emphasizes the importance of maintaining low flows for various purposes, including protecting fisheries and irrigation.

Cleaveland, M. K. (2000). A 963-year reconstruction of summer (JJA) stream flow in the White River, Arkansas, USA, from tree-rings. *The Holocene*, *10*(1), 33-41

Patskoski’s study introduces a hybrid approach that combines tree-ring chronologies and sea surface temperature (SST) data for reconstructing annual streamflow. It recognizes the limitations of traditional regression techniques in estimating high flow values. The proposed hybrid approach uses SST conditions from the tropical Pacific and tree-ring chronologies from watersheds in the southeastern United States. This approach overcomes the limitations of traditional methods, particularly in estimating high flow values. It employs Singular Spectrum Analysis (SSA) to extract periodic and non-periodic components from tree-ring chronologies, Nino3.4, and streamflow data, resulting in improved estimates of reconstructed annual streamflow for selected watersheds.

Patskoski, J., Sankarasubramanian, A., & Wang, H. (2015). Reconstructed streamflow using SST and tree-ring chronologies over the southeastern United States. *Journal of Hydrology*, *527*, 761-775.

In conclusion these articles provide historical context and accurate reconstructions of streamflow data, which are invaluable for guiding water resource management in regions facing complex challenges due to climate change and human interventions. Understanding the ecological and hydrological impact of changing water levels is essential for sustainable resource management, energy production, and ecological preservation.

**SPENCER:** AI / ML Streamflow Reconstruction

Having accurate and complete time series for hydrological data is crucial to making decisions around water resources. However, gaps in streamflow records and short records can make predicting water difficult for managers and decision makers. There have been some efforts to use machine learning techniques to fill in these gaps and to reconstruct records from other sources, such as tree ring data.

A method of machine learning called MissForest was used in Arriagada, et al to fill in gaps in daily streamflow reconstructions. The MissForest algorithm is an implementation of the random forest algorithm and uses the existing data to predict the missing data. This method works with large datasets with multiple variables. While this paper was done in Argentina, it seems like it could be easily applied to the southeast with the right data.

A paper by Zhao, er al. used an ensemble of different machine learning models to reconstruct summertime streamflow from 1788 to 2016 in the Kazakh Uplands using tree ring data as a proxy for streamflow. They also used data for temperature, and known dry and wet periods to further enhance the model. Using this they were able to find correlations with rainfall, temperature, and tree ring growth. They also believe that future models could bring in soil moisture as well for proxy indicators for reconstruction.

The ensemble model they created for this involved a random forest machine learning model, nearest neighbor algorithm, and multi-linear regression model. Using this they were able to identify hydrologic extremes of drought and high flow in the record. The periods that they identified with this method were correlated with historical records of famine or flooding in the region.

Overall stream flows are affected by multiple factors and reconstructions need to be done by taking into account an ensemble of variables. Machine learning techniques have been shown to be useful in reconstructing stream flows with these different variables and having better accuracy compared to more traditional regression. The use of these could help decision makers have a greater understanding of changes in the hydrologic cycle in a basin over time to better make decisions.

Arriagada, P., Karelovic, B., Link, O., 2021. Automatic gap-filling of daily streamflow time series in data-scarce regions using a machine learning algorithm. Journal of Hydrology 598, 126454.<https://doi.org/10.1016/j.jhydrol.2021.126454>

Zhao, X., Zhang, R., Chen, F., Maisupova, B., Kirillov, V., Mambetov, B., Yu, S., He, Q., Dosmanbetov, D., Kelgenbayev, N., 2022. Reconstructed summertime (June–July) streamflow dating back to 1788 CE in the Kazakh Uplands as inferred from tree rings. Journal of Hydrology: Regional Studies 40, 101007.<https://doi.org/10.1016/j.ejrh.2022.101007>