Use-Inspired Problem Statement

1. **Research Challenge (SPENCER)**

Flooding results in loss of life and causes billions of dollars in damages annually in the United States. Flooding is expected to increase in temperate areas from increased precipitation due to human induced climate change and land use changes (Alfieri, et al. 2017). This increase in precipitation can have many effects on the management of streams and water infrastructure that were not built for these changes. The way that water infrastructure is designed currently is based on streamflow from gage stations. This data, however, can have many limitations on time or a continuous record. One potential solution to this is using tree rings as a proxy for stream flow. This can allow for extending and filling in gaps in the stream flow record to allow for better planning and flood prediction in areas with incomplete or limited records. Advances in machine learning can also aid in this area allowing for better and more completed extending of stream flow records than typical regression techniques.

1. **Scholarly Support**

Formetta et al. 2021 focused on the Adige River Basin (ARB), a vital water source for northern Italy, that supports irrigation, hydropower, and municipal water supply. The mountainous terrain and intricate hydrological processes of the Adige River Basin result in highly variable streamflow. Comprehending the historical variability of streamflow in the Adige River Basin is crucial for effective water resource management and planning in light of climate change. Traditional streamflow reconstruction methods utilize tree-ring proxies, which depend on the link between tree-ring growth and precipitation. These methods, however, are constrained by the availability of tree-ring data and the precision of the proxies. A novel approach employing the Old World Drought Atlas (OWDA) scPDSI as a reconstruction proxy has been developed and effectively applied in other European watersheds. The OWDA scPDSI, a self-calibrating drought index derived from tree-ring chronologies, is accessible over a wide spatial and temporal range.

Machine-Learning-Based Precipitation Reconstructions: A Study on Slovenia’s Sava River Basin (Molina et al, 2023) **MADI**

The implementation of machine learning for paleo streamflow reconstructions has not been largely explored, however, successful ML reconstructions of other climate variables have been completed in recent years. Paleo reconstructions of precipitation in the Sava River Basin were recently published using AI/ML/DL techniques which showed skill statistics higher than that of traditional reconstruction methods (Molina et al, 2023).

This preliminary literature investigation helps build context of the previous studies on streamflow reconstructions using dendroclimatic proxies. Having accurate and complete time series for hydrological data is crucial to making decisions about water resources. Our NRT project aims to extend historical streamflow records with tree-ring proxies in the spatial extent of the Southeast United States using both traditional regression methods and novel AI/ML/DL techniques.

1. **Research Need Mahsa**

The use of tree-ring proxy-based streamflow reconstruction in the southeast US has the potential to provide valuable insights into past hydrological and climate conditions, which can inform water resource management decisions and help mitigate the impacts of climate change. However, there are several gaps, concerns, and opportunities that need to be addressed in order to fully realize the potential of this research.

One of the key gaps is the limited availability of streamflow records in the southeast US. This makes it difficult to assess long-term trends and variability in water resources, and to develop accurate and reliable streamflow reconstructions (Patskoski et al. 2015). Another gap is the need for further development and evaluation of AI/ML approaches for streamflow reconstruction. While these methods have the potential to overcome some of the limitations of traditional regression techniques, they may also be more sensitive to the quality of the data used to train them. It is important to develop robust AI/ML models that can accurately capture complex relationships in streamflow data and that can be used to reconstruct streamflow records for a variety of conditions. Additionally, there's a lack of literature review or integration of more advanced machine learning techniques beyond regression models to enhance the accuracy or predictive capabilities of streamflow reconstructions. In addition to these gaps, there are also several concerns that need to be addressed. One concern is that traditional regression techniques may not be able to capture complex relationships in streamflow data (Gharib, A., & Davies, E. G. (2021). A workflow to address pitfalls and challenges in applying machine learning models to hydrology. *Advances in Water Resources*, *152*, 103920.

Ridgeway et al. 1999), which could lead to inaccuracies in streamflow reconstructions. Another concern is that machine learning techniques may be overfitted if the models are not properly trained (Gharib & Davies 2021). It is important to develop methods for evaluating the performance of different streamflow reconstruction methods under a variety of conditions and to identify the most effective methods for different applications. Moreover, the development and application of AI/ML models require large amounts of high-quality data. Ensuring the availability and quality of tree-ring proxy data is crucial for the success of AI/ML-based streamflow reconstruction projects. Despite these gaps and concerns, there are also several opportunities for the use of tree-ring proxy-based streamflow reconstruction in the southeast US. Streamflow reconstructions can be used to inform water resource management decisions by providing insights into past hydrological and climate conditions, as well as potential future trends. They can also be used to identify areas that are at risk of drought or flooding, and to develop strategies for managing these risks (Anderson et al. 2019). Stakeholders involved in water resource management, such as agencies like the Tennessee Valley Authority (TVA), will benefit from this research. The extended historical streamflow records will enable better decision-making regarding water allocation and management strategies. Insights derived from the data aid in optimizing resource utilization and resilience planning against climate-induced variability. resource utilization and resilience planning against climate-induced variability.

1. **Goal and Objectives MADI**

The goal of conducting this research is to extend historical streamflow records with tree-ring proxies, provide past information on water availability and variability, and assist water managers in making better water resources decisions in the southeast United States. The research objective is to utilize and compare AI/ML/DL models for streamflow reconstruction to traditional methods within the Tennessee Valley watershed using tree-ring based proxies. Because AI/ML/DL reconstruction techniques have not been employed on streamflow, our research is original in exploring these emerging techniques on a climate variable critical for making water resource management decisions.

1. **Anticipated Impact (SPENCER)**

Reconstructing stream flow from tree ring data using machine learning will allow for the TVA to have longer and more complete records. This would better inform their decisions when it comes to making improvements and repairs to infrastructure that manages water resources and flood control. This research could also aid the research and scientific community in understanding the changes in the water levels over time in the Tennessee Valley area. Leading to more complete research on ecological and hydrological impacts on resource management, energy production and ecological preservation.

With the threat of floods potentially increasing due to climate change, having a better understanding of historical streamflow in watersheds is more important than ever. Especially with people most at risk and with less means to mitigate their own risk being statistically overrepresented in flood plains, generating more complete streamflow records will allow for better planning to potentially avoid greater disasters (Tate et al., 2021).

Benefits of the research

Consequences of not doing the research

**References**:

Formetta, G., Tootle, G., & Bertoldi, G. (2021). Streamflow Reconstructions Using Tree-Ring Based Paleo Proxies for the Upper Adige River Basin (Italy). *Hydrology*, *9*(1), 8.

Ridgeway, G., Madigan, D., & Richardson, T. S. (1999, January). Boosting methodology for regression problems. In *Seventh International Workshop on Artificial Intelligence and Statistics*. PMLR.

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.

Gharib, A., & Davies, E. G. (2021). A workflow to address pitfalls and challenges in applying machine learning models to hydrology. *Advances in Water Resources*, *152*, 103920.

Anderson, S., Ogle, R., Tootle, G., & Oubeidillah, A. (2019). Tree-ring reconstructions of streamflow for the Tennessee Valley. *Hydrology*, *6*(2), 34.

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K., Feyen, L.,

2017. Global projections of river flood risk in a warmer world. Earth’s Future 5, 171–182.

<https://doi.org/10.1002/2016EF000485>

Tate, E., Rahman, M.A., Emrich, C.T., Sampson, C.C., 2021. Flood exposure and social vulnerability

in the United States. Nat Hazards 106, 435–457. <https://doi.org/10.1007/s11069-020-04470-2>

Ramírez Molina, A.A.; Bezak, N.; Tootle, G.; Wang, C.; Gong, J. Machine-Learning-Based Precipitation Reconstructions: A Study on Slovenia’s Sava River Basin. *Hydrology* **2023**, *10*, 207. <https://doi.org/10.3390/hydrology10110207>