Can Statistical Models Provide Improved Performance over DAR Methods for Estimating Flow-duration Curves and Daily Streamflows in Ungauged Catchments?

Shubham Jain [1,2], *Michael Schramm* [1] *(michael.schramm@ag.tamu.edu), Dhruva Kathuria* [1,2]

1. Texas Water Resources Institute, Texas A&M AgriLife Research and Extension Service, College Station, TX
2. Texas A&M College of Agriculture and Life Sciences, College Station, TX

Abstract

Flow-duration curves (FDCs) are used to estimate and describe streamflow exceedance values, development of pollutant load duration curves, and estiamtion of daily streamflow values. Development of FDCs at ungaged sites if often accomplished using the drainage-area ration method which equates the ratio of of streamflow at the ungaged site with a nearby gaged site with the ratio of the corresponding areas (Asquith, Roussel, and Vrabel 2006). This method while easily applicable, has certain limitations especially when the two watersheds differ in land cover, soil type, topography or when the stations are far apart. In this study, an alternate approach is proposed which uses watershed attributes from GAGES II dataset as explanatory variables and models the flow-duration curve using random forest regression.

Methods

We subset gages and potential covariates using the USGS GAGES II dataset (Falcone 2011). Discharges at 17 streamflow percentiles at each gage were developed using DAR and random forest regressions. Random forests models were produced using leave one out cross validation. DARs calculations were produced using all candidate streamgages within 100 km. Performance was evaluated using NSE and r^2 across all gages and along the predicted streamflow exceedance percentiles.

Results

DARs have a high variance in performance, with 20% of gages performing worse than the mean (NSE < 0). 53% of gages had NSE >= 0.7, indicating good performance for select gages. The DAR approach appears to perform most poorly at low flows.

Random forests provide improved performance across all flow quantiles compared to the DAR approach. Random forest had good performance at 64% of gages and fairly for 12% of gauges.

Conclusion

DARs are simple to apply but require specific conditions to maximize performance (Asquith, Roussel, and Vrabel 2006). The random forest method exhibits acceptable performance across a wide range of streamflow quantiles and can be used when conditions for DARs are not met. FDC transfer methods enable the reconstruction of daily streamflows timeseries at the ungaged basin. The next step in this project is to evaluate criteria necessary for FDC transfer methods. An R package is also in development to facilitate estimating FDCs in ungaged basins using random forest models and the GAGES II dataset.

Random forest models **accurately predict** flow-duration curves using **readily available landscape data**.

Both random forest models and drainage-area approaches can accurately predict flow-duration curves in ungaged basins. Random forest models are suitable where nearby similar gaged watersheds are present and drainage-area ratios don't perform well.

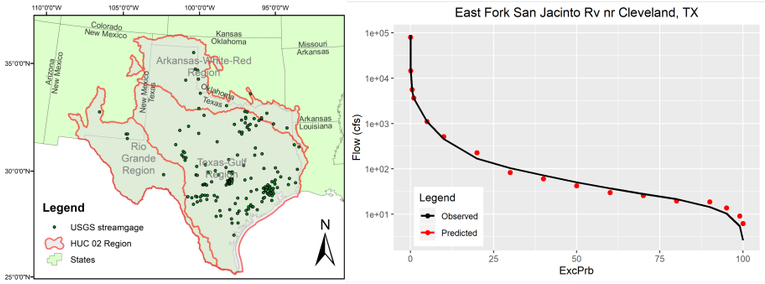

Figure 1: (A) Selected streamflow gages used to predict and validate methods and (B) an example of a predicted versus observed FDC as a single site.

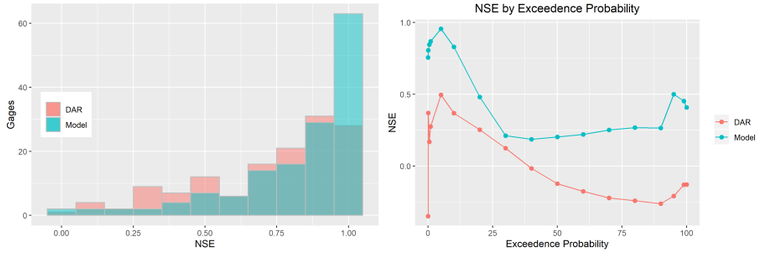

Figure 2: (A) Distribution of per-gage NSE values by DAR and random forest methods and (B) comparison of NSE by exceedance percentile for DAR and random forest methods.

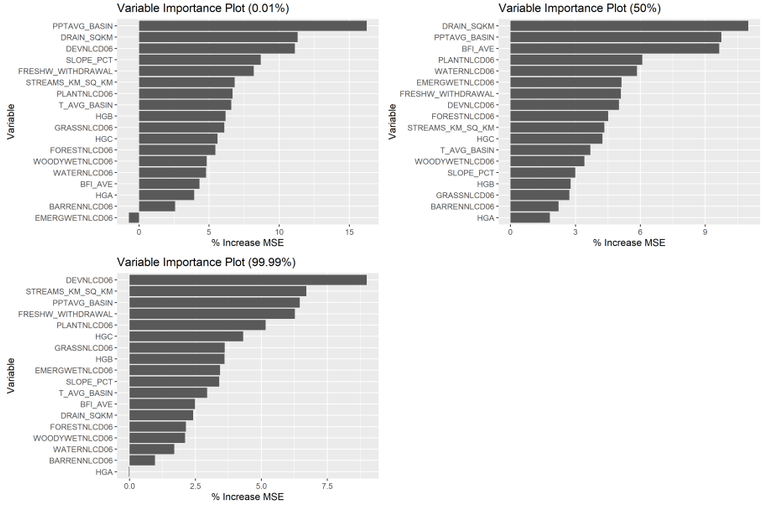

Figure 3: Variable importance plots for random forest models at selected flow exceedance percentiles.

References

Asquith, William H., Meghan C. Roussel, and Joseph Vrabel. 2006. "Statewide Analysis of the Drainage-Area Ratio Method for 34 Streamflow Percentile Ranges in Texas." 2006-5286. U.S. Geological Survey. https://pubs.usgs.gov/sir/2006/5286/pdf/sir2006-5286.pdf.

Falcone, James. 2011. "GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow." Digital spatial dataset. Reston, VA: U.S. Geological Survey. https://water.usgs.gov/lookup/getspatial?gagesII_Sept2011.

Texas Water Resources Institute
make every drop count
TEXAS A&M AgriLife RESEARCH | EXTENSION