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. The drainage-area ratio (DAR) method is often used to develop FDCs for ungaged sites. The method is easily applied but is limited to nearby watersheds with similar hydrology. We assessed the performance of random forest regressions and parametric approaches for predicting multiple stream exceendace percentiles using watershed attributes from the GAGES II dataset (Falcone 2011) and compared them to the DAR method.

Methods

We subset gages and candidate covariates using the USGS GAGES II dataset (Figure 1). Discharges at 17 streamflow percentiles at each gage were developed using DAR, parametric models, and random forest regressions. Parameters for the parametric approach, using an extended Burr II distribution (Shao et al. 2009), were estimated using multiple regression. Random forests models were fit using leave one out cross validation. DAR calculations were developed using the nearest candidate streamgage within 100 km. Performance was evaluated using NSE 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 NSE >= 0.7 at 64% of gages and 0.4 < NSE < 0.7 for 12% of gauges. The parametric approach generally failed to produce shape parameters with physical meaning.

Conclusion

DARs are simple to apply but require specific conditions to maximize performance. Random forest 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.

Texas Water Resources Institute
make every drop count

Texas A&M AgriLife Research | Extension

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.

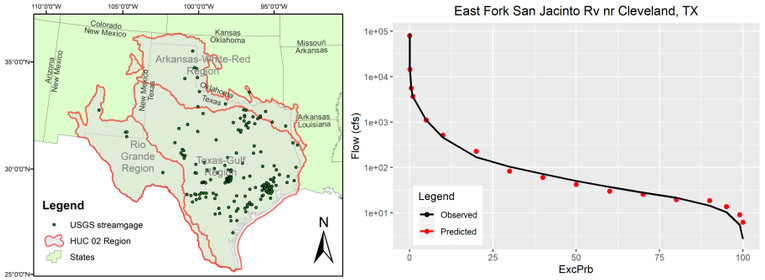

Figure 1: (A) Selected streamflow gages used to predict and validate methods and (B) an example of a predicted versus observed FDC as a single site.

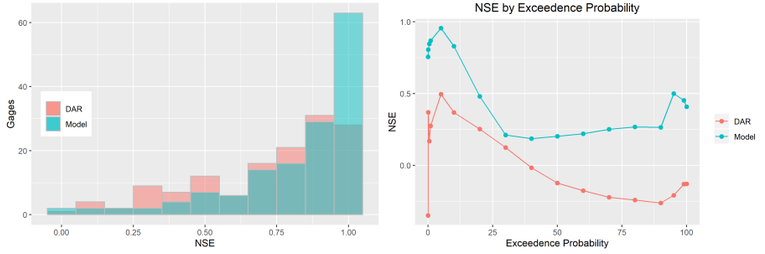

Figure 2: (A) Distribution of per-gage NSE values by DAR and random forest methods and (B) comparison of NSE by exceedance percentile for DAR and random forest methods.

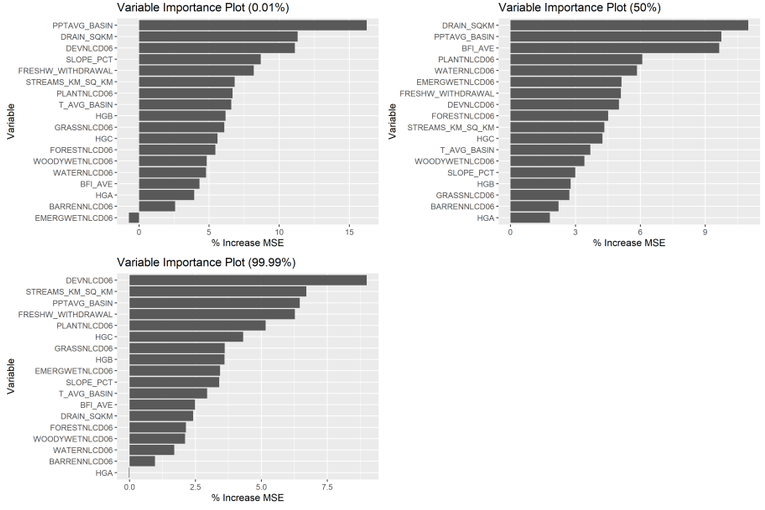

Figure 3: Variable importance plots for random forest models at selected flow exceedance percentiles.

References

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.

Shao, Quanxi, Lu Zhang, Yongqin D Chen, and Vijay P Singh. 2009. "A New Method for Modelling Flow Duration Curves and Predicting Streamflow Regimes Under Altered Land-Use Conditions/Une Nouvelle Méthode de Modélisation Des Courbes de débits Classés et de Prévision Des Régimes d'écoulement Sous Conditions Modifiées d'occupation Du Sol." *Hydrological Sciences Journal* 54 (3): 606–22.