Regionalization of hydrologic response in the Great Lakes basin: Considerations of temporal and spatial scales of analysis

Jonathan Kult[1,2], Lauren Fry[2], Andrew Gronewold[3], and Woonsup Choi[1]

[1]Department of Geography, University of Wisconsin-Milwaukee, Milwaukee, WI, Unites States
[2]Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, MI, United States
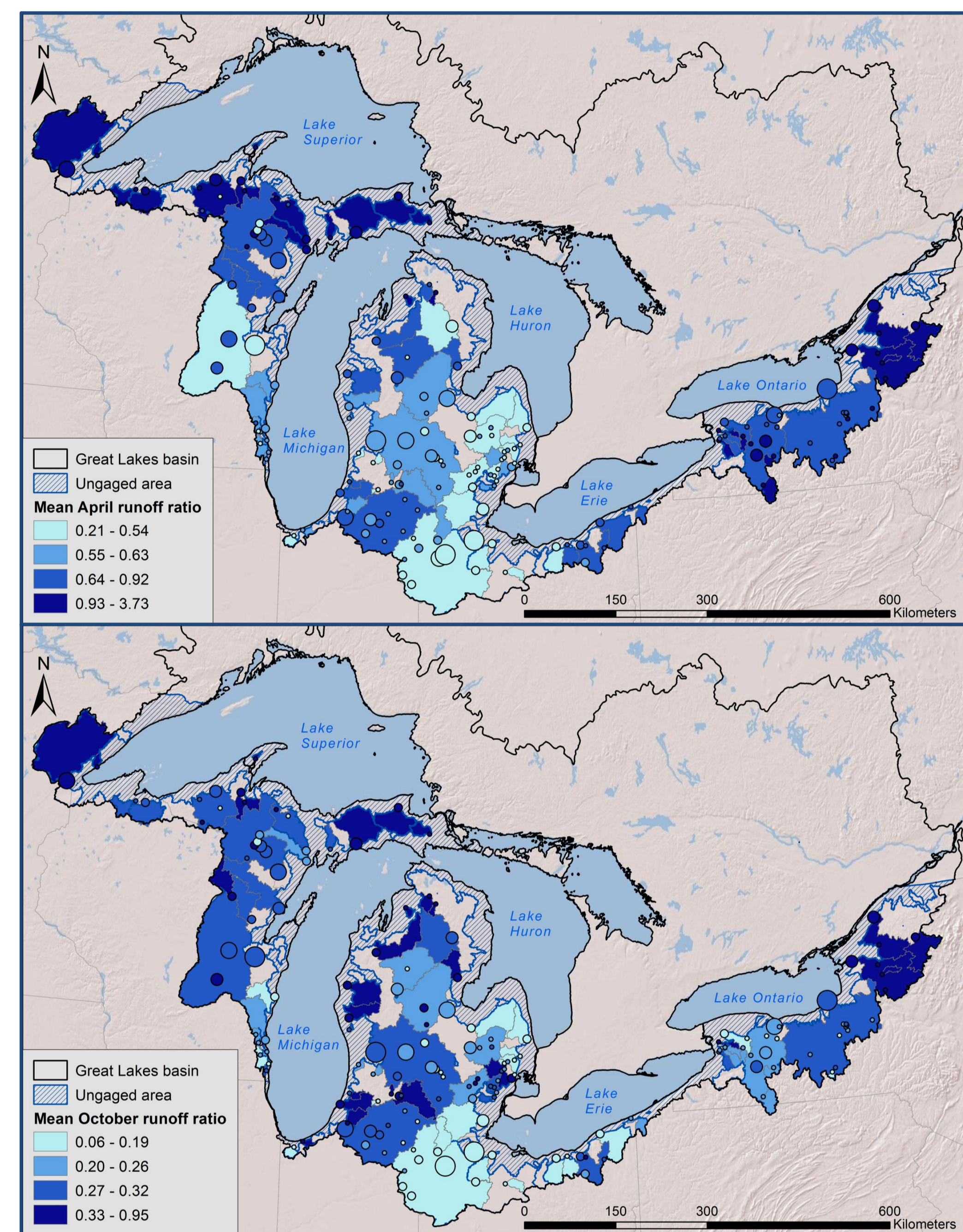[3]Great Lakes Environmental Research Laboratory, National Oceanic and Atmospheric Administration, Ann Arbor, MI, United States

Contact for details:
Jonathan Kult (jmkult@uwm.edu)

©NOAA-GLERL

INTRODUCTION

Numerous water resource research and management objectives require knowledge of hydrological processes occurring in ungaged watersheds (Sivapalan et al., 2003). Methods for predicting streamflow in ungaged watersheds typically invoke the concept of regionalization, whereby knowledge pertaining to gaged catchments is transferred to ungaged catchments. Wagener and Montanari (2011) review emerging methods wherein hydrologic response indices (rather than streamflow observations) in gaged watersheds are employed to establish a regionalization scheme.

Hydrologic response can be defined as the translation of rainfall into runoff via watershed routing, storage and loss processes. A wealth of indices (see Olden and Poff, 2003) have been derived to implicitly quantify these processes, such as watershed input-output relationships (e.g. runoff ratio), hydrograph analytics (e.g. rising limb density) and probabilistic characterizations of the magnitude, frequency, duration and timing of flow events (e.g. baseflow index and flood frequency).

OBJECTIVES

The objectives of this study are to:

+ Develop regression-based regionalization models relating watershed physical characteristics (Table 1) and monthly runoff ratio (MRR; monthly runoff divided by monthly precipitation) as an index of hydrologic response;

+ Use the regionalization models to predict streamflow in ungaged catchments at a monthly time step;

+ Demonstrate the effects of defining hydrologic response at different temporal scales in terms of model skill and applicability to water resource management objectives; and

+ Assess the potential for regression tree models for hydrologic modeling alongside a commonly used linear regression model.

Table 1: Significance of watershed physical characteristics based on fitting monthly linear models. Highlighted characteristics were retained for developing the regionalization schemes.

Figure 1

Figure 1: Seasonal and spatial variability of monthly runoff ratio (MRR) in the Great Lakes basin, shown as quartiles of MRR_m for April (top) and October (bottom). Study gages are shown as circles graduated by drainage area. Filled polygons are the largest of any nested watersheds. Hatched areas denote catchments with historically limited or nonexistent streamflow records. Note the sizable difference in range for MRR_m between the spring and autumn months.

STUDY AREA

The Great Lakes basin drains 521,830 km² of land in the United States and Canada, spanning a variety of climates, land uses, subsurface properties and human activities. Great Lakes water levels are important at local and regional scales for recreation, commerce, ecosystems, hydropower, transportation, and consumptive supply.

As a study area, the Great Lakes basin poses unique challenges related to its large size, transnational regulation and data coordination efforts, and a clear siting bias in the gaging network, with coastal areas primarily ungaged and inland areas predominantly gaged (Figure 1). Each of these hydrologic modeling challenges are variously manifested worldwide, commonly in the regions most susceptible to hydrologic impacts from expanding populations and changes in land use and climate (Sivapalan et al., 2003).

Figure 2: Location and results from predictions in nine validation watersheds. Median E and R^2 values of the four regionalization schemes are given in the callout boxes. Watersheds are color-coded according to mean absolute difference of volume, D_v. Boxplots give the range in E and R^2 values over the nine watersheds and indicate that the rpart models perform at least as well as a traditional multiple linear regression.

Figure 3: Model-simulated versus observed monthly runoff (in mm) for three representative validation watersheds during WY 2001-2010.

DATA AND METHODS

The following data were obtained for the study:

+ Watershed physical characteristics in Geographic Information System format from GAGES-II dataset (USGS, 2011)

+ Precipitation and temperature data from NOAA-GLERL

+ Streamflow data from UGSG for w = 163 gaged watersheds spanning the Great Lakes basin with continuous flow records during WY 1981-2010

For a gaged watershed w with observations covering t years, we define monthly runoff ratio (MRR) at two temporal scales, as presented in Table 2.

Table 2 (below): Definition and description of MRR at two temporal scales of analysis. Y represents the total number of MRR observations for 12 months over all years t for all watersheds w contributing to the regionalization scheme. The block of "Applications" spanning both temporal scales lists research and management objectives for which the selection of the temporal characterization of hydrologic response should be considered carefully.
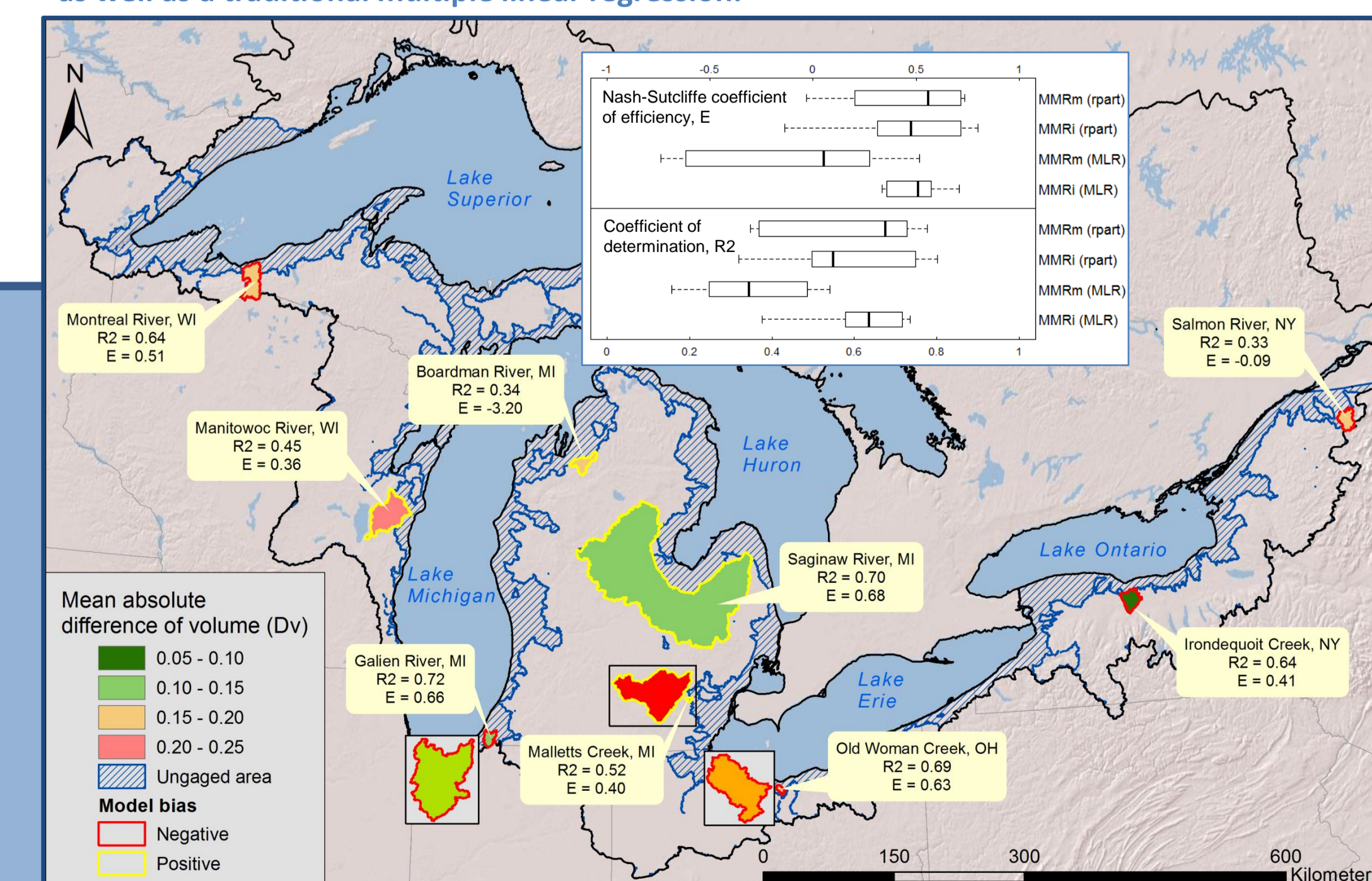
Monthly runoff ratio (MRR) temporal scale	MRR_i	MRR_m
MMR derivation	12 monthly runoff ratios computed for each year i in (i = 1,2,...t) MRR = monthly runoff divided by monthly precipitation	Mean of t observed MRR_i values
MMR contribution to model	Each observed MRR contributes to the model ($Y = 12w$)	A single mean MRR contributes to the model ($Y = 12wt$)
Interpretation as independent variable, y, for modeling	y : full variability of watershed hydrologic response	y : average watershed hydrologic response
Applications	Short-term lake level forecasts; extreme/rare event predictions	Near-shore wetland habitats, freight traffic, hydropower potential, municipal and agricultural consuption, synthetic runoff series for calibrating hydrologic models	Long-term lake level forecasts; beach and recreational areas
Limitations	Potential for overly conservative management decisions	Knowledge of model uncertainty may be insufficient for sensitive applications or use with calibrating rainfall-runoff models
Usage in contemporary PUB research	Rarely used in PUB research	Frequently used: hydrologic response indices are typically based on long-term averages of their parameters

Regionalization schemes relating MRR_m and MRR_i with physical watershed characteristics are established using traditional multiple linear regression (MLR) and a method rarely employed in hydrology—regression tree analysis, implemented in R with package rpart (Therneau and Atkinson, 2011). Trees in rpart are grown so as to maximize differences in watershed characteristics at each branching in a simple analysis of variance. The resulting trees are then pruned to minimize the risk of misclassifying an observation, while avoiding excessive model complexity or overfitting.

The 12 watershed physical characteristics used for the regionalization schemes are the highlighted rows in Table 1. Criteria for inclusion in this set of characteristics include:

+ Variable significance at the $p < 0.05$ level over most of the year based on fitting monthly linear models;

+ Minimal correlation between variables, based on Pearson correlation coefficients (results not shown);

+ Representation of climate, soils, land cover and geomorphology variables;

+ Availability of data with continuous coverage over all ungaged areas of the Great Lakes basin.

RESULTS

October and April MRR_m values are displayed in Figure 1. Figure 4 presents the full range of seasonal variability in observed MRR_m and MRR_i as well as differences in the distributions of the two temporal scales. Both figures illustrate the seasonal variability of hydrologic response over the Great Lakes basin.

Monthly runoff ratios are predicted from the MLR and rpart models for nine validation watersheds (Figure 2) for WY 2001-2010. Multiplication by monthly observed precipitation yields a time series of monthly runoff predictions by each model for each watershed. Figure 3 displays hydrographs of predicted versus observed runoff for three of the nine validation watersheds. In Figure 2, goodness-of-fit statistics (R^2 and the Nash-Sutcliffe coefficient of efficiency, E) are given in the callout boxes. Watersheds are color-coded by mean absolute difference of volume (D_v), a measure of model bias.
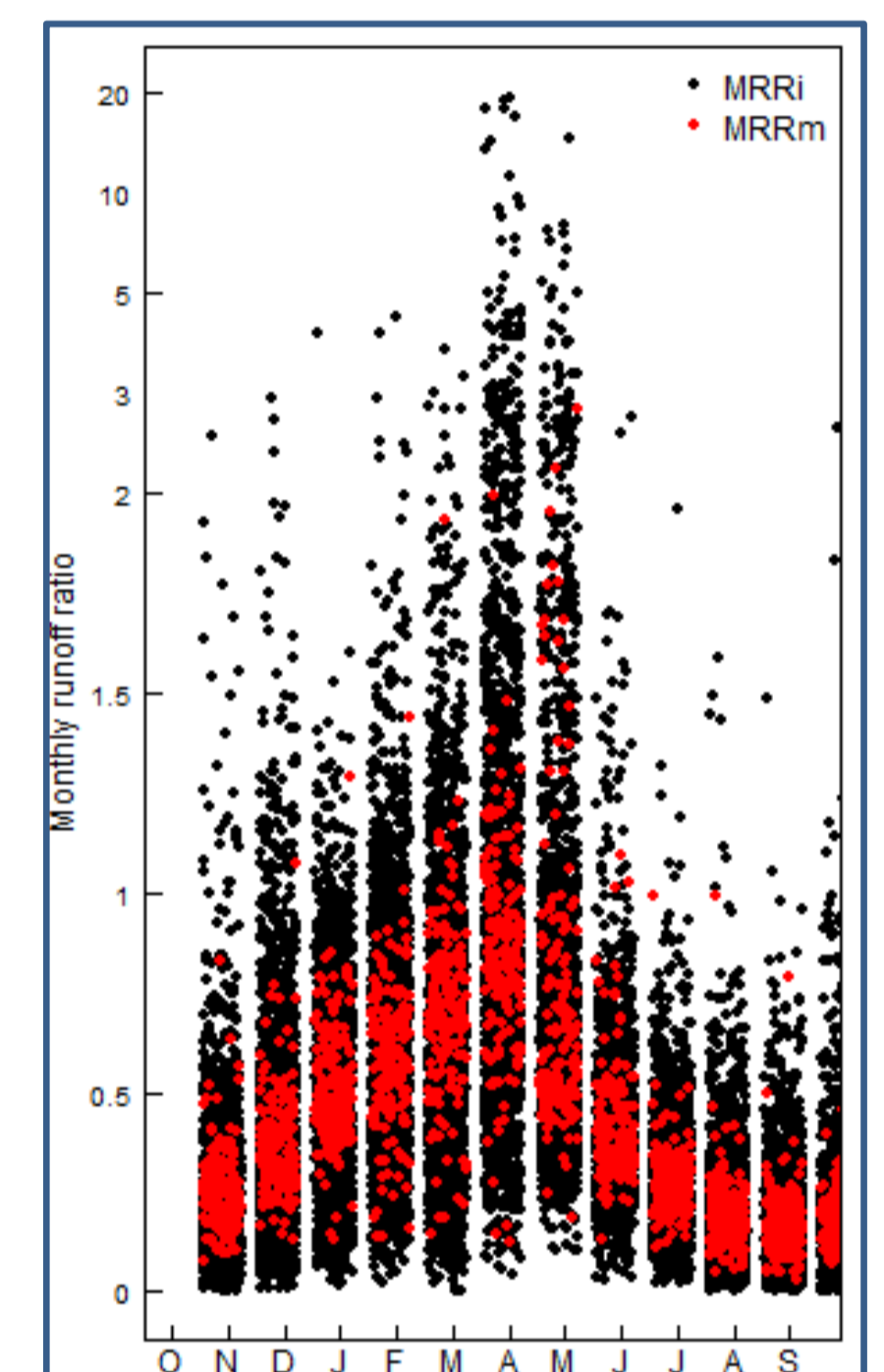
Figure 4 (right): Seasonal variability in observed monthly runoff ratios at two temporal scales: MRR_m (red dots, n = 163) and MRR_i (black dots, n = 1630). Values greater than 1 indicate a net monthly watershed surplus. Surplus conditions are rare during summer and autumn months, but quite prevalent during winter and early spring months in response to snowpack accumulation and ablation processes.

CONCLUSIONS

Our results suggest the following conclusions, each of which contributes to contemporary research involving hydrologic predictions in ungaged watersheds:

+ Monthly runoff in ungaged watersheds can be predicted with reasonable skill using regionalization relationships between runoff ratio and watershed physical characteristics;

+ Regression tree models using rpart demonstrate comparable skill alongside a commonly used linear regression model;

+ Predictions in ungaged watersheds are highly influenced by the temporal scale used to condition the models.

This last result is particularly important given the numerous applications of hydrologic response indices in contemporary research for making predictions in ungaged watersheds. These predictions are essential for water resource management in the Great Lakes basin and worldwide.

ACKNOWLEDGEMENTS

REFERENCES

Olden, J. D., & Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. River Research and Applications, 19(2), 101-121.

Sivapalan, M., Takeuchi, K., Franks, S., Gupta, V., Karambiri, H., Lakshmi, V., . . . Zehe, E. (2003). IAHS decade on predictions in ungauged basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. Hydrological Sciences Journal, 48(6), 857-880.

Therneau, T. M., & Atkinson, E. J. (1997). An Introduction to Recursive Partitioning using the RPART Routines. Technical Report 61, Section of Biostatistics, Mayo Clinic, Rochester.

USGS (2011). GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. Available at http://water.usgs.gov/lookup/getspatial?gagesII_Sept2011.

Wagener, T., & Montanari, A. (2011). Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. Water Resources Research, 47, W06301.