# Assessment of Differences in Watershed Physical Characteristics Between Gaged and Ungaged Portions of the Great Lakes Basin

Lauren M. Fry[1,2], Jonathan M. Kult[1,3], Andrew D. Gronewold[2], and Timothy Hunter[2]

[1]Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, MI, USA (lauren.fry@noaa.gov),
[2]NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor , MI, USA; [3]University of Wisconsin-Milwaukee, Milwaukee, WI, USA

Abstract No. H51I-1464

## Introduction

Methods for predicting streamflow in basins with limited or nonexistent streamflow records typically invoke the concept of regionalization, whereby knowledge pertaining to gaged catchments is transferred to ungaged catchments (Vogel 2006). Regionalization approaches relating watershed physical characteristics to calibrated model parameters or hydrologic signatures have been important contributions to what Wagener and Montanari (2011) consider a "convergence of approaches" toward understanding watershed behavior in ungaged catchments. However, differences in watershed physical characteristics between gaged and ungaged catchments constitute a common obstacle for any regionalization approach. For example, predictions from approaches based on spatial proximity may be of limited value or reliability if there are substantially different climate, soils, land cover, or topographic regimes in the gaged and ungaged areas. Additionally, regression models between gaged watersheds' physical characteristics and their hydrologic response may not be valid across ungaged basins if the distributions of gaged catchments' characteristics do not adequately represent those of ungaged catchments. Understanding the degree of spatial heterogeneity among such characteristics is therefore essential for selecting an appropriate regionalization scheme.

The stream gage network in the Great Lakes basin, as with many coastal regions, exhibits a clear siting bias: coastal areas are largely ungaged, while inland areas are predominantly gaged (Figure 1). We hypothesize that this siting bias may result in important differences between gaged and ungaged areas, complicating regionalization schemes in the basin.

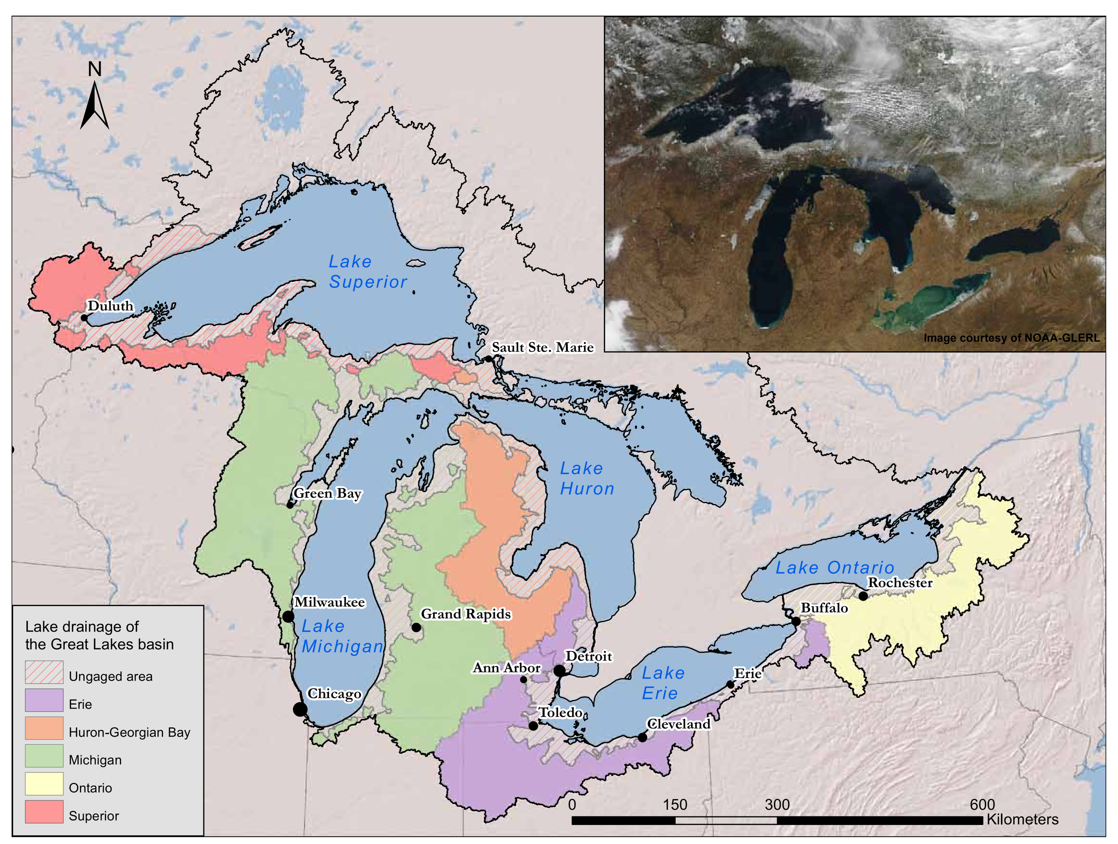
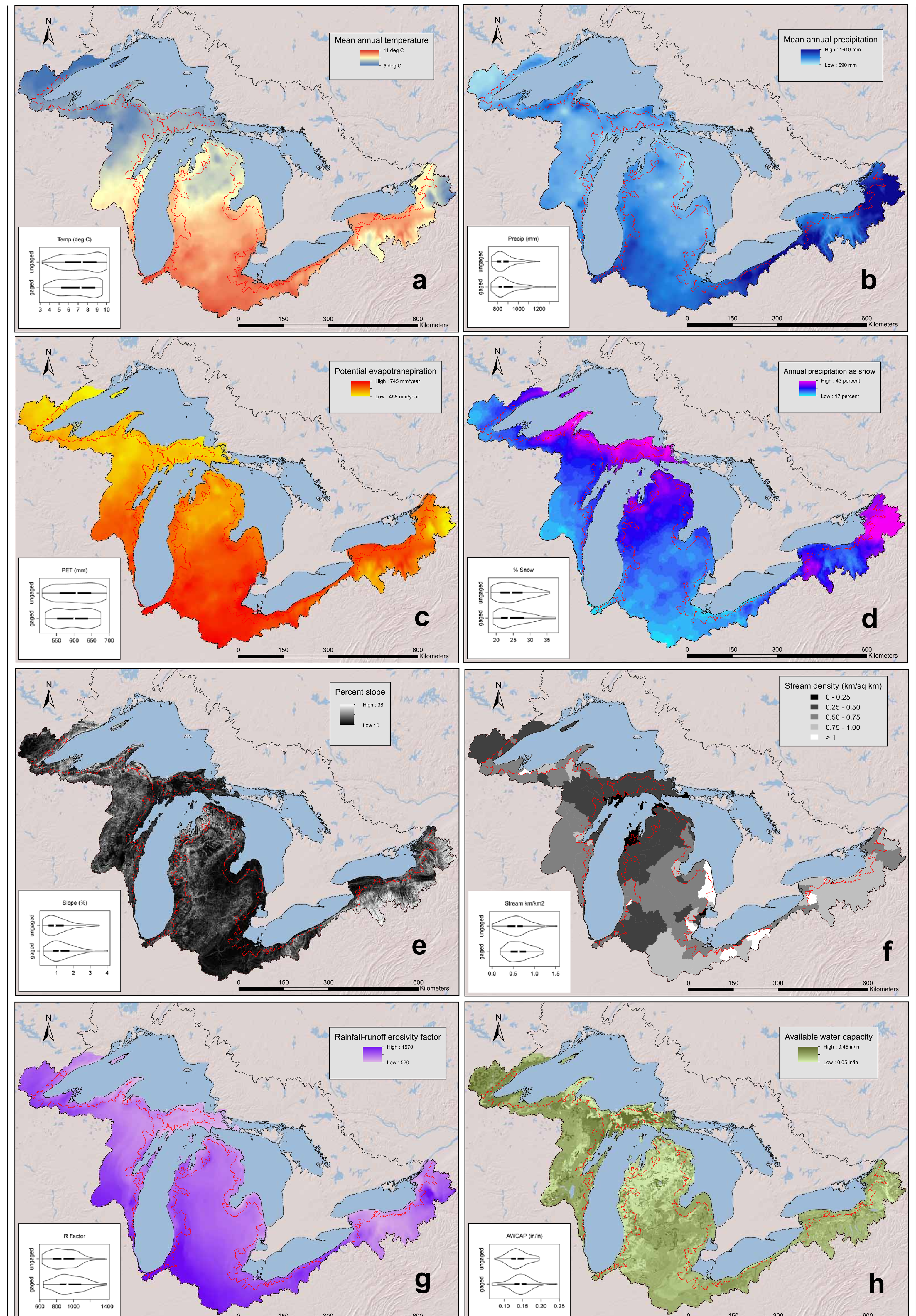



Figure 1. Drainage areas of lakes in the Great Lakes basin. Solid areas represent gaged portions, and hatched areas represent ungaged portions.

## Approach and Data

Differences in watershed physical characteristics between gaged and ungaged areas of the Great Lakes basin are examined at three spatial scales with tools from the Spatial Analyst and Spatial Statistics toolboxes in ArcGIS 10.0. Gaged areas are defined here as all catchments with: (1) 20+ years of continuous discharge records since 1950 or (2) currently active gages as of water year 2009. These criteria, plus the geospatial data for all watershed physical characteristics, were obtained from the source data of the GAGES-II dataset (USGS, 2011). Characteristics considered here (Figure 2a-i) are based on the findings of Kult et al. (2012) (see Poster H51I-1476).

Raster datasets from GAGES-II were aggregated to: 1) the 79 (U.S.) subbasin delineations used by the Large Basin Runoff Model to predict monthly runoff to each of the Great Lakes (Croley II & Hartmann 1986) and 2) 1415 USGS subwatershed (HUC12) delineations in the Lake Michigan basin. Contributing gaged and ungaged portions were determined for each delineation. The subbasin delineations assess gaged/ungaged differences in watershed physical characteristics over the entire basin and individual lake basins, while the subwatershed delineations explore variable gradients at a smaller scale.



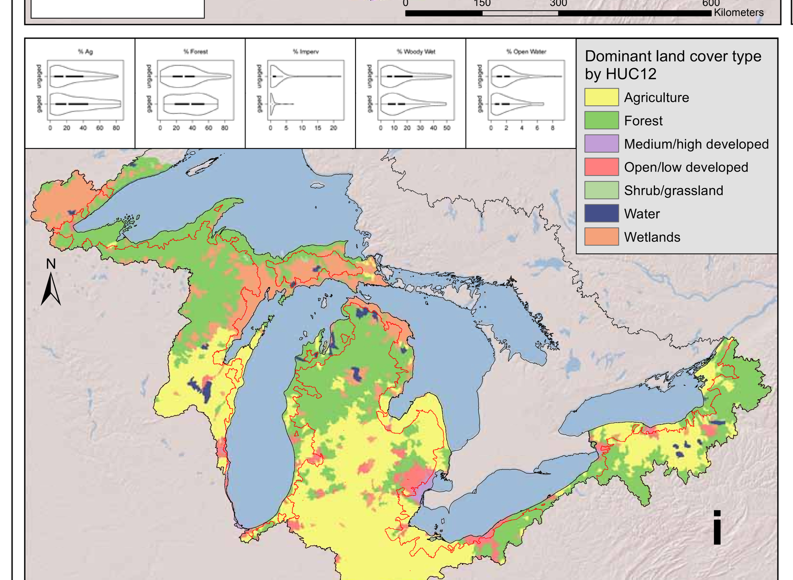




Figure 2. Watershed physical characteristics of the Great Lakes basin: (a) Mean annual temperature, (b) Mean annual precipitation, (c) Mean annual potential evapotranspiration, (d) Percent of precipitation as snow, (e) Slope, (f) Stream density, (g) Rainfall-runoff erosivity factor, (h) Available water capacity, and (i) Land cover.

Using the `vioplot` package in `R`, violin plots (Hintze & Nelson 1998) were created to display the range and distribution of subbasin values for each watershed physical characteristic. Violin plots present both the quartiles and density shape of a variable's distribution.

To assess gaged/ungaged differences at a smaller scale we compute the Gedis-Ord $G_i$ statistic for each of the HUC12 subwatersheds. $G_i$ is a frequently used spatial statistic for detecting local "hot spots" or "cold spots" in an attribute's value, that is, areas with concentrations of higher or lower values relative to a zonal mean (Getis and Ord 1992). The 1415 HUC12s were aggregated to thirteen zones for the analysis based on LBRM delineations. Each zone contains at least 50 HUC12 units spanning both gaged and ungaged areas and represents a typical spatial scale for regionalization applications. Statistically significant hot and cold spots (clusters of HUC12s with large absolute z-scores) are investigated based on their situation in the gaged and ungaged portions of each zone.

## Results

Violin plots of watershed physical characteristics for the gaged/ungaged portions of the entire Great Lakes basin are shown in Figure 2a-i. At this scale of analysis, only percent impervious surface exhibits notable differences between gaged and ungaged areas. Therefore, a model including impervious surface as a predictor variable, conditioned on characteristics of gaged catchments, may not adequately simulate flow in ungaged portions of the basin.

Violin plots for the individual lake basins (Figure 3) indicate increased differentiation between gaged and ungaged areas for some variables. For example, it appears that precipitation is higher in gaged portions of the Lake Huron basin, but higher in ungaged portions of the Lake Superior basin. Additionally, compared to the gaged portions of the Lake Superior basin, percent woody wetlands is lower and percent forest is higher in ungaged areas. Models conditioned on watershed characteristic distributions in gaged areas at the lake basin scale may be not be robust when applied to ungaged areas exhibiting a substantially different distributions. For example, a regional regression model conditioned on slope may not be appropriate for regionalization across the Superior, Michigan, Erie, or Ontario lake basin.

Figure 4a depicts the thirteen zones used to aggregate HUC12s for computing the $G_i$ statistic, our smallest scale of analysis. Figure 4b illustrates hot and cold spots by zone for mean annual precipitation, based on z-scores of the $G_i$ statistics. Zone 10 in Figure 4c provides an exemplary case of a watershed physical characteristic gradient between gaged and ungaged areas: higher and lower precipitation relative to the zonal mean are observed in the ungaged and gaged areas, respectively. Interestingly, while the strong regional coastal-inland precipitation gradient is still present in the adjacent Zone 9, the overall distribution of highs and lows is far more homogeneous, as indicated by comparison of the boxplots of z-scores. In Figure 4d, there is a clear north-south precipitation gradient, this time largely perpendicular to the gaged/ungaged boundary. Consequently, hot and cold spots are observed in both gaged and ungaged portions of the zone.

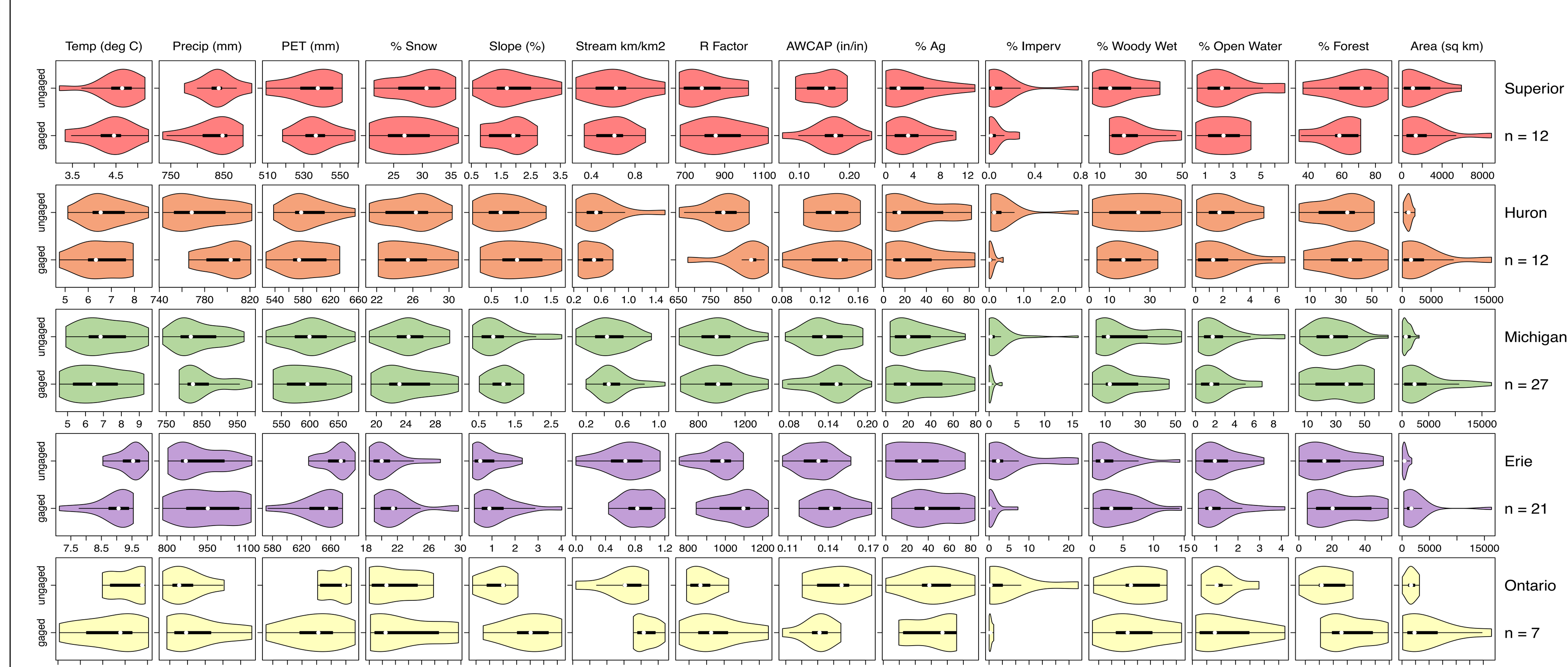
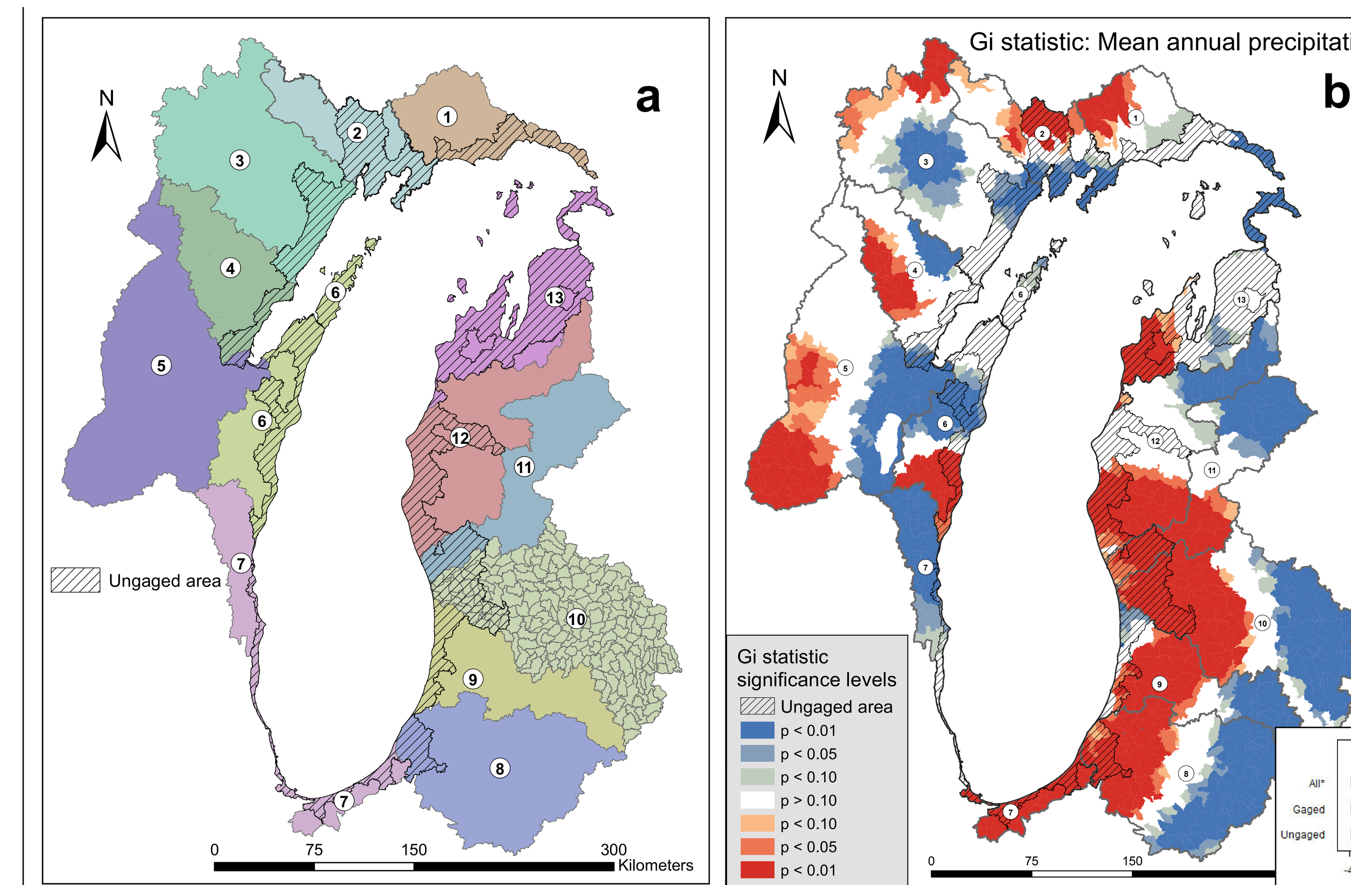



Figure 3. Violin plots of watershed physical characteristics for gaged and ungaged subbasins. Subbasins are color-coded for reference to Figure 1. In these plots, the white circle represents the median, the bar extends from the first to the third quartile, and the line represents the spread of the values.



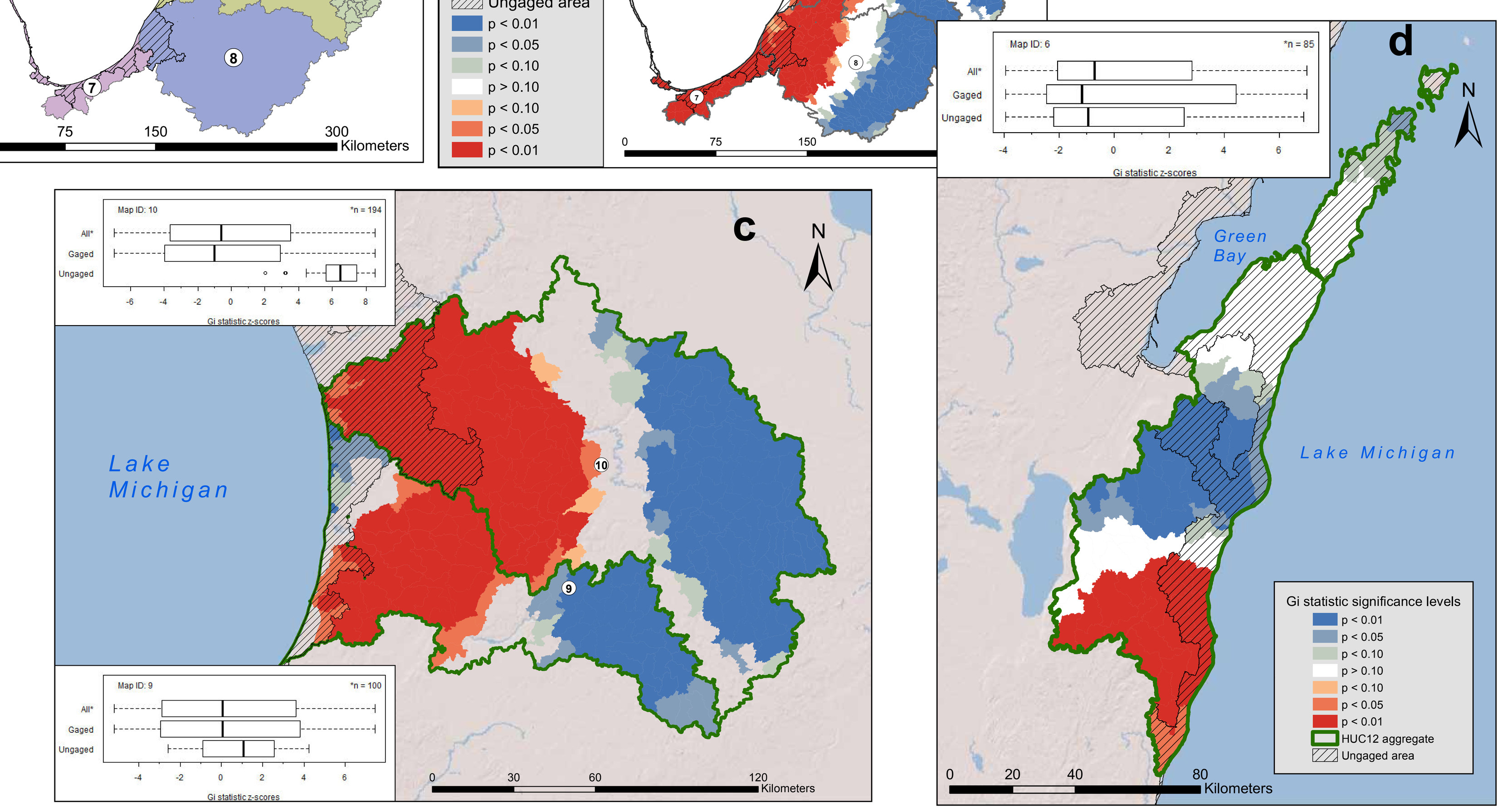




Figure 4. (a) Zones used to aggregate Lake Michigan basin HUC12s for computing the $G_i$ statistic. (b) Precipitation hot spots and cold spots based on $G_i$ statistics. Cold spots (blue) represent clusters of HUC12s with precipitation lower than the zonal mean. Hot spots (red) represent clusters of HUC12s with precipitation higher than the zonal mean. (c) Precipitation hot and cold spots for Zones 9 and 10 in southwestern Michigan, alongside boxplots of z-scores of the $G_i$ statistics. Zone 10 exhibits a strong west-east precipitation gradient, parallel to the gaged/ungaged boundary. (d) A north-south precipitation gradient, perpendicular to the gaged/ungaged boundary is evident in eastern Wisconsin, resulting in both hot and cold spots in gaged and ungaged portions of the zone.

## Conclusions

Understanding the degree of physical similarity between gaged and ungaged catchments is an important step toward predicting flow in ungaged basins. In this study, differences in watershed physical characteristics between gaged and ungaged portions of the Great Lakes basin were analyzed at multiple spatial scales. These differences begin to appear at the individual lake basin scale and become more pronounced (and more complicated) at smaller scales. Our results illustrate how localized spatial heterogeneity of watershed physical characteristics may result in significant differences between gaged and ungaged portions of a subbasin, thereby complicating or potentially invalidating regionalization approaches based on spatial proximity. While this study focused on in the Great Lakes basin, the spatial analyses and statistics conducted here are readily applicable to other coastal environments, regions with gage siting biases, or multi-scalar considerations for large basins.

## Acknowledgements

This work was partially completed under a post-doctoral fellowship with the Cooperative Institute for Limnology and Ecosystems Research (CILER), awarded under a Cooperative Agreement between the University of Michigan and the NOAA Great Lakes Environmental Research Laboratory. Additional support was provided through CILER's summer fellowship program.

## References

Croley II, T. E., & Hartmann, H. C. (1986). NOAA Technical Memorandum ERL GLERL-61: Near-Real-Time Forecasting of Large-Lake Water Supplies; A User's Manual. Ann Arbor, MI.

Hintze, J. L., & Nelson, R. D. (1998). Violin plots: A Box plot-density trace synergism. *The American Statistician* 52(2), 181–184.

Kult, J. M., Fry, L. M., & Gronewold, A. D. (2012). Accounting for inter-annual and seasonal variability in regionalization of hydrologic response in the Great Lakes basin. AGU Fall Meeting Abstracts. San Francisco, CA.

Vogel, R. M. (2006). Regional calibration of watershed models. In V. P. Singh & D. F. Frevert (Eds.), *Watershed Models* (pp. 549–567). Boca Raton, FL: CRC Press.

Wagener, T., & Montanari, A. (2011). Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resources Research* 47(6), 1–8. doi:10.1029/2010WR009469