Water Resources Research

RESEARCH ARTICLE
10.1029/2022WR033759

Controls of Variability in the Laurentian Great Lakes Terrestrial Water Budget

Samar Minallah[1,2], Allison L. Steiner[1], Valeriy Y. Ivanov[3], and Andrew W. Wood[2,4]

[1]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA, [2]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [3]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI, USA, [4]Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO, USA

Key Points:

- The Laurentian Great Lakes domain is not hydrologically uniform, with different regulators of the water budget in the five subbasins
- Dominant quantities characterizing the subregional terrestrial hydrology vary for daily, monthly, and annual timescales
- Climate change impact studies on regional hydrology need to account for basin-scale differences in the terrestrial hydroclimatic dynamics

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. Minallah,
minallah@umich.edu

Citation:

Minallah, S., Steiner, A. L., Ivanov, V. Y., & Wood, A. W. (2023). Controls of variability in the Laurentian Great Lakes terrestrial water budget. *Water Resources Research*, *59*, e2022WR033759. https://doi.org/10.1029/2022WR033759

Received 26 SEP 2022
Accepted 21 SEP 2023

Author Contributions:

Conceptualization: Samar Minallah, Allison L. Steiner
Formal analysis: Samar Minallah
Investigation: Samar Minallah
Methodology: Samar Minallah, Allison L. Steiner, Valeriy Y. Ivanov, Andrew W. Wood
Validation: Samar Minallah
Visualization: Samar Minallah
Writing – original draft: Samar Minallah
Writing – review & editing: Samar Minallah, Allison L. Steiner, Valeriy Y. Ivanov, Andrew W. Wood

Abstract The land surface hydrology of the North American Great Lakes region regulates ecosystem water availability, lake levels, vegetation dynamics, and agricultural practices. In this study, we analyze the Great Lakes terrestrial water budget using the Noah-MP land surface model to characterize the catchment hydrological regimes and identify the dominant quantities contributing to the variability in the land surface hydrology. We show that the Great Lakes domain is not hydrologically uniform and strong spatiotemporal differences exist in the regulators of the hydrological budget at daily, monthly, and annual timescales. Subseasonally, precipitation and soil moisture explain nearly all the terrestrial water budget variability in the southern basins, while the northern latitudes are snow-dominated regimes. Seasonal assessments reveal greater differences among the basins. Precipitation, evaporation, and runoff are the dominant sources of variability at lower latitudes, while at higher latitudes, terrestrial water storage in the form of ground snowpack and soil moisture has the leading role. Differences in land cover categorizations, for example, croplands, forests, or urban zones, further induce spatial differences in the hydrological characteristics. This quantification of variability in the terrestrial water cycle embedded at different temporal scales is important to assess the impacts of changes in climate and land cover on catchment sensitivities across the diverse hydroclimate of the Great Lakes region.

1. Introduction

The Laurentian Great Lakes region is one of the largest freshwater ecosystems with a unique hydroclimate that extends from boreal to humid continental and subhumid climates (Beck et al., 2018). The regional hydroclimate is modulated by the atmospheric phase of the water cycle, terrestrial hydrological processes, and lake surface processes, which together regulate lake water levels, biodiversity, vegetation dynamics, and wetland systems. The terrestrial water budget is an essential component of the water cycle as it determines the ecosystem water availability (e.g., through snowpack storage dynamics and soil moisture fluctuations) and atmospheric humidity (through evapotranspiration) that can alter cropping patterns and vegetation dynamics (Huntington et al., 2018). Additionally, terrestrial processes alter lake water quantity and quality through surface runoff, which directly impact navigational, recreational, and water use needs. Thus, understanding the modalities of the terrestrial water cycle is important for improved predictive capabilities of climate and hydrometeorological processes, and assessing the related sensitivities of the regional ecosystems.

For the Great Lakes region, various studies have assessed the atmospheric water budget (Li et al., 2010; Minallah & Steiner, 2021b), the net basin supply which constitutes over lake precipitation, lake evaporation, and runoff into the lake (Do et al., 2020; Fortin et al., 2012; Mailhot et al., 2019; Music et al., 2015), and the connection between the regional climate and lake levels (Bennington et al., 2015; Durnford et al., 2018). Various studies have also evaluated the variability in the terrestrial water cycle over the continental US (CONUS) or regional domains such as the Midwest US, however, a detailed terrestrial water budget assessment for the Great Lakes watersheds is needed. For example, Syed et al. (2004) assessed the process controls in the land hydrological cycle over the CONUS to identify the dominant drivers of spatiotemporal variability, but without due consideration of processes that may have important regional controls (e.g., snowmelt). Huntington et al. (2018), in their broad CONUS assessment, found that northern (southern) parts of the Great Lake basin had reduced (increased) soil water storage during the 1985–2014 period, as compared to the preceding decades, highlighting the importance of spatial differences within this region that require further investigation. At local scales, Yeh and Wu (2018) conducted a

trend assessment in land hydrology variables for Illinois (a subregion within the Great Lakes domain) over the 1983–2013 period and found an intensification of the hydrological cycle, with positive trends in precipitation, evapotranspiration, and runoff, but a decrease in the terrestrial water storage. In contrast, J. Niu et al. (2014) assessed the water storage in two basins in Michigan (the Grand River and the Saginaw Bay watersheds) and revealed an increase in storage over the 2002–2012 period. These studies highlight the need for a fine-scale spatiotemporal assessment of the land surface hydrology and to quantify the variability of each terrestrial water budget quantity in the regional water cycle of the Great Lakes.

This work focuses on the terrestrial water budget of the five Laurentian Great Lakes watersheds: Superior, Michigan, Huron, Erie, and Ontario. The study objectives are to (a) characterize the different catchment hydrological regimes of the region, (b) distinguish the dominant drivers of variability among the water budget components at different timescales, (c) study the relationships between the budget quantities and the relative contributions of change in each variable from other components, and (d) assess the influence of soil and land cover categories on the variability in the terrestrial water budget.

Identifying the extent of hydrometeorological variability embedded at different temporal resolutions is especially important to improve regional water budget accounting for water resources planning purposes (e.g., land use and land cover policies [Levia et al., 2020]) and predictability of future evolution of the regional water cycle. This work aims to provide a new fine-resolution, region-specific assessment of the land hydrology of the Great Lakes domain and to establish a baseline for evaluating the variability in the terrestrial water budget that will help assess future changes in the water cycle for different climate scenarios and land cover developments.

2. Methods

2.1. Domain

The model domain is approximately 1,640 × 1,460 kms in area and encompasses the watersheds of the five Laurentian Great Lakes: Superior, Michigan, Huron, Erie, and Ontario (Figure 1a). We conduct an assessment over the land component of the basin only, where lakes with surface area larger than 100 km^2 are masked out.

2.2. Modeling Framework

We use the Noah-MP land surface model (LSM; G.-Y. Niu et al., 2011) to simulate the land surface hydrology and compute the terrestrial water budget within the domain. We run Noah-MP within the WRF-Hydro V5.2.0 modeling system and use the NOAA National Water Model (NWM) configuration for the physics parameterization of the land surface processes (excluding surface runoff; Gochis et al., 2020). The NWM configuration has been tested for various domains and conditions across CONUS (Lahmers et al., 2021), while Noah-MP provides multiple options for simulation of overland and subsurface runoff and groundwater transfer (Barlage et al., 2015).

We run the LSM at 9 km resolution and hourly time step, with the coupled WRF-Hydro terrain routing at 900 m with a 300 s model time step, and use a 90 m hydrologically conditioned DEM for surface elevation (Minallah & Steiner, 2022). The information passed from the high-resolution terrain routing module back to the LSM is the volumetric surface head (instantaneous depth of ponded water on surface) and soil moisture (Gochis et al., 2020). After a 10-year treadmill spin-up (i.e., running the model with 1-year forcing multiple times) for a stable initialization of soil moisture and other related fields, we simulate the Great Lakes basin over the 2017–2020 hydrological years (October 2016 to September 2020). For runoff simulations, we test two schemes: the TOPMODEL approach mimicking lateral flows and groundwater (G.-Y. Niu et al., 2007) and the default runoff scheme in WRF-Hydro that lacks lateral moisture redistribution, with free drainage at the bottom of soil column (Gochis et al., 2020). The atmospheric forcing is obtained from the North American Mesoscale (NAM) 12 km, 6-hourly analysis (NCEP, 2015) for the period under consideration.

2.3. Terrestrial Water Budget

We define the terrestrial water budget as

$$\Delta W = P_t - \mathrm{ET} - \mathrm{RO} + R \quad (1)$$

where P_t is the total precipitation (mm day^{-1}), ET is the land evapotranspiration (mm day^{-1}), and RO is the total surface and subsurface runoff (mm day^{-1}). The change in the modeled terrestrial water storage (ΔW; mm day^{-1})

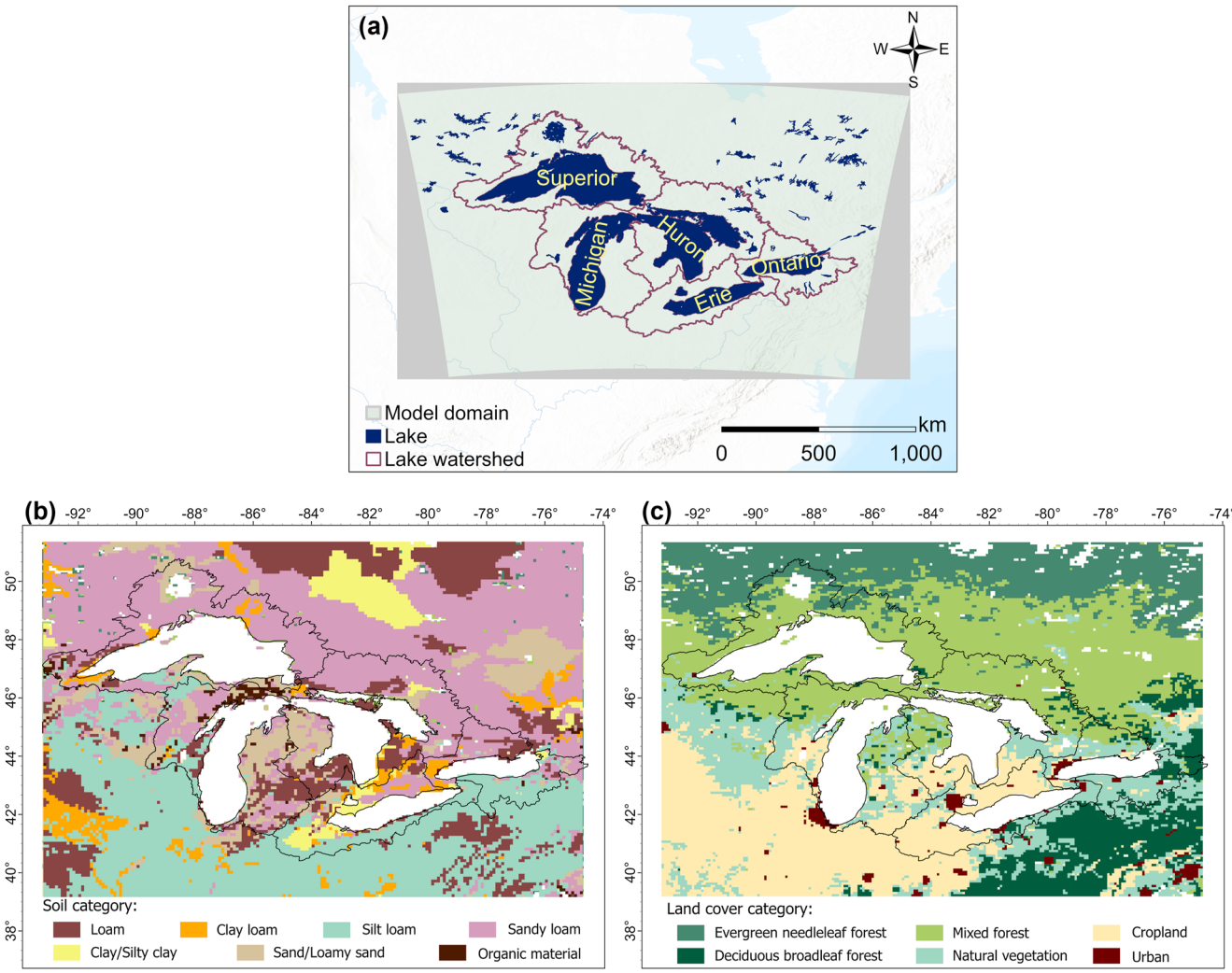


Figure 1. (a) The Noah-MP model domain (gray box) with the five Great Lakes basins. (b) Main soil types and (c) land cover categories of the Great Lakes domain. The complete categorizations of soil and land cover are provided in Figure S2 in Supporting Information S1.

is balanced by the difference in the water influx (P_t) and the outgoing water fluxes (ET + RO) from each domain grid cell. The change in terrestrial water storage consists of four quantities:

$$\Delta W = \Delta\mathrm{SM} + \Delta\mathrm{SWE} + \Delta\mathrm{Aq} + \Delta\mathrm{Can} \tag{2}$$

where SM is the soil moisture in the top 2-m soil layer (mm), SWE is the ground snowpack (mm water equivalent), Aq is the aquifer recharge (mm), and Can is the canopy interception (mm). All ΔW quantities are individually computed at a daily time step (dt = 1 day) as

$$\Delta W = (W_t - W_{t-1}) / \mathrm{dt} \tag{3}$$

R in Equation 1 is the residual for cases where the water budget is not closed. In the Noah-MP model output, the seven budget quantities in Equations 1 and 2 complete the water balance and thus the residual is zero.

2.4. Model Output Evaluation

2.4.1. Runoff Evaluation

For the modeled runoff evaluation, we use streamflow data from over 230 USGS gauges based on the availability of their respective catchment outlines and the complete time series over the assessment period (Figure S1a in

Supporting Information S1), similar to the runoff verification approach used by Xu et al. (2021). We find high correlation between the modeled runoff estimates and the observed streamflow, with a coefficient of determination (R^2) of 0.71 for the TOPMODEL scheme and 0.69 for the free drainage scheme (Figure S1b in Supporting Information S1). The two schemes have small differences in the runoff magnitudes, with the TOPMODEL approach generally simulating marginally higher runoff (0.1–0.2 mm day^{-1}), especially over regions where the soils are characterized as loam or silty loam (Figures S2c and S2d in Supporting Information S1). However, the overall spatial patterns for the two parameterizations are similar (Figures S2a and S2b in Supporting Information S1). For the remainder of this study, we use the TOPMODEL approach that uses a groundwater scheme due to its slightly improved simulation of the terrestrial runoff, evaporation, and soil moisture (Wu et al., 2021).

For further evaluation of the TOPMODEL scheme, we use the Kling-Gupta Efficiency (KGE) criteria (Gupta et al., 2009), including the KGE score and the linear correlation between observations and simulations (r), using the monthly time series of all available gages within each of the five watersheds (Figure S3 in Supporting Information S1). A KGE score of 1 indicates a perfect agreement between simulations and observations, and KGE = −0.41 indicates agreement no better than the climatological mean (Knoben et al., 2019). In literature, various thresholds for these metrics are accepted to indicate "good" model simulations; for example, KGE score >0 (positive values only), KGE > 0.3, or KGE > −0.41 for mean flows (Knoben et al., 2019, and citations within). For our model output, the median KGE score across the gages for each of the five watersheds is larger than all these thresholds indicating strong model performance: 0.58 for the Lake Superior watershed, 0.45 for Michigan, 0.51 for Huron, 0.53 for Erie, and 0.63 for Ontario (Figure S3a in Supporting Information S1). The correlation coefficient (r) is greater than 0.75 for all basins (Figure S3b in Supporting Information S1), indicating that the model is capturing the runoff processes reasonably well.

2.4.2. Comparison With Observation-Based Gridded Products

One of the challenges of LSM evaluation is the limitations in spatiotemporal observations and differing scales of models and observations. Each terrestrial water budget variable has a different observational network, and many quantities are derived and not directly observed. Additionally, observations and reanalysis products do not close the water budget. Here, we use various gridded products to evaluate the modeled magnitudes and seasonal cycles of the individual terrestrial water budget variables.

The NAM forcing precipitation, which is regridded for the Noah-MP simulation to the model resolution, has similar magnitude and seasonality as compared to the observation-based 0.5° CRU data product (v4.05; Harris et al., 2020) for all five basins, with small differences (mostly <0.5 mm day^{-1}) in some months (Figure S4 in Supporting Information S1). We compare the simulated Noah-MP soil moisture with a gridded data set that uses the satellite-derived Soil Moisture Active Passive (SMAP) brightness temperatures combined with a hyperresolution (30-m) LSM to acquire surface soil moisture estimates (SMAP-HB; Vergopolan et al., 2020). While a one-to-one comparison in magnitude is difficult due to the differences in the surface soil layer depth in the available data products (10 cm for Noah-MP vs. 5 cm for SMAP-HB), we find that the seasonal cycle of the soil moisture content in the respective top layers is similar for the two data sets, except in the Superior basin, where SMAP-HB is out-of-phase in some months (Figure S5 in Supporting Information S1). For SMAP-HB, the magnitudes and seasonal cycle of the soil moisture are affected by canopy cover which makes satellite-based retrievals challenging in forested regions (Entekhabi et al., 2010). The higher forest cover in the Superior watershed (Figure S2e in Supporting Information S1) can be a reason for the differences in the seasonal cycles of the two data sets.

Evapotranspiration estimates are compared with those of the Global Land Evaporation Amsterdam Model (GLEAM v3.5b; Martens et al., 2017), which uses satellite retrieval of precipitation, soil moisture, vegetation optical depth, radiation, and temperature to estimate evaporation. The simulated Noah-MP evapotranspiration has smaller magnitudes, by approximately 70%–90% of the GLEAM model estimates (Figure S4 in Supporting Information S1). The parameterization of soil moisture, runoff, and evapotranspiration is closely coupled and controlled by multiple factors (e.g., soil characterization, land cover type, and model physics), where adjusting one parameter will affect all quantities in various ways. The modeled runoff matches well with the gauge-based streamflow observations and the change in soil moisture has similar magnitudes and cycle as compared to other data sets, however, we note that the Noah-MP evapotranspiration in the summer months is underestimated for all the basins. The lower ET could result from constraining the model to simulate streamflow if the precipitation inputs have a low bias, although it is difficult to quantify the degree to which these factors (i.e., relatively higher model calibration effort in Noah-MP and potential forcing bias) may influence ET without further study. Despite

the magnitude differences, the ET seasonality is simulated adequately, and the magnitude difference does not affect the variability analysis of this work.

2.4.3. Comparison With Reanalyses

We also compare the model output with the 9-km ERA5-Land reanalysis (Muñoz-Sabater et al., 2021) and the MERRA-2 land surface diagnostics at $0.5° \times 0.625°$ resolution (Global Modeling and Assimilation Office, 2015). In general, the Noah-MP magnitudes and seasonal cycle are similar to the ERA5-Land data for all basins and variables, except evapotranspiration, where Noah-MP estimates are 60%–80% of the ERA5-L magnitudes depending on the month (Figure S4 in Supporting Information S1). The MERRA-2 magnitudes and seasonal cycles are generally different from both Noah-MP and ERA5-L, with substantially larger evaporation, reduced runoff and SWE, and weaker amplitude of change in soil moisture.

2.5. Water Budget Assessment Approach

2.5.1. Principal Component Analysis

Multivariate empirical orthogonal function (EOF) analysis, also called principal component analysis (PCA), is intended to reduce interrelated multidimensional data (e.g., the seven budget quantities in Equations 1 and 2) to a set of arrays of lower dimension, while preserving maximum information on variability in the original data. This is done by linearly transforming the original variables into a set of synthetic uncorrelated variables, called principal components (PCs; Hannachi et al., 2007; Jolliffe & Cadima, 2016). PCA has been used extensively to analyze various components and processes associated with the hydrological cycle. For example, Syed et al. (2004) used PCA to study the underlying processes controlling the variability of the hydrologic cycle over CONUS, specifically focusing on four variables (precipitation, soil moisture, runoff, and potential evaporation) extracted from North American Land Data Assimilation System (NLDAS). Other hydrology-specific applications of PCA include assessment of hyperresolution hydrologic modeling outputs (Mascaro et al., 2015), analysis of soil moisture spatiotemporal patterns (Fang et al., 2015; Perry & Niemann, 2008), assessment of water quality (Zeinalzadeh & Rezaei, 2017), study of drought characteristics (Qiao et al., 2019; Santos et al., 2010) and spatial distribution of floods (Chang et al., 2022), groundwater quality assessments (Taşan et al., 2022), assessing predictive accuracy of rainfall-runoff models (Zhang et al., 2007), and synthesizing snowmelt patterns (Woodruff & Qualls, 2019).

In this study, PCs are used to determine the dominant quantities that account for most variability among the components of the terrestrial water budget, similar to the PCA application in Syed et al. (2004). We conduct PCA at two temporal resolutions for each basin: (a) spatially averaged, daily time series concatenated for each month over the 4-year period (113–124 days) to attribute controls of the subseasonal and intramonthly variation in the hydrological budget and (b) spatially averaged, monthly time series over the 4 years (48 months) to identify the components controlling seasonal variability of the budget components (Table S1 in Supporting Information S1).

2.5.2. Partial Least Squares Regression

Partial least squares regression (PLSR) is another multivariate statistical approach which projects data to a new component space. A key distinction from PCA is that it establishes a relationship between a dependent variable (Y, predictand or response variable) and a set of independent variables (X, explanatory variables) when they exhibit collinearity (Wold et al., 2001). The PLS components (Z, predictors) are linear combinations of X which "maximize the variance explained in Y and the correlation between X and Y" (Smoliak et al., 2010).

Here, we use the PLSR approach to establish the most important predictor variables that drive change in each budget quantity. We set four variables (evapotranspiration, runoff, change in soil moisture, and change in aquifer recharge) as the response variable individually, with the remaining six budget quantities as the explanatory variables for each case. Canopy interception and ground snowpack are excluded as the response variables because they are primarily precipitation- and temperature-dependent quantities. We then compute the variable importance in projection (VIP, also variable influence on projection) score (Chong & Jun, 2005) to measure the importance of each explanatory variable for the response variable, where a VIP score >1 is generally considered important for the projection. Further details on PLSR and VIP are available in literature (Chong & Jun, 2005; Wold et al., 2001), and more recently, this approach has also been adopted in the atmospheric sciences and hydrology fields (Baker et al., 2020; Black et al., 2017; Fu et al., 2015; Mendoza et al., 2017; Smoliak et al., 2010).

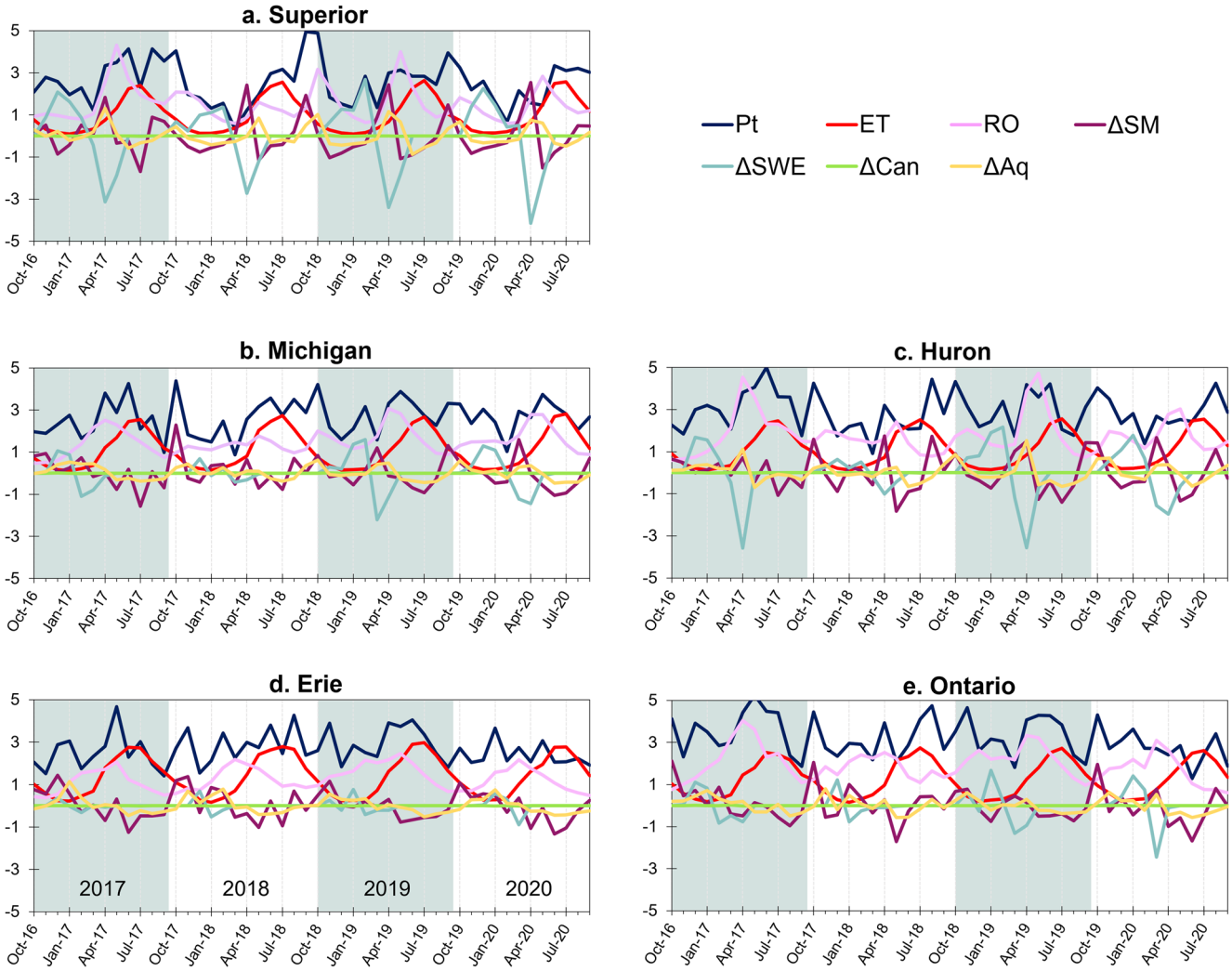


Figure 2. Monthly time series of the terrestrial water budget quantities over four hydrological years (October 2016 to September 2020) for the watersheds of lakes (a) Superior, (b) Michigan, (c) Huron, (d) Erie, and (e) Ontario. P_t is total precipitation, ET is evapotranspiration, RO is total runoff, SM is soil moisture, SWE is snow water equivalent, Can is canopy interception, and Aq is aquifer recharge. All variables are in mm day^{-1}.

3. Results and Discussion

3.1. Interbasin Differences in Budget Seasonality

All budget quantities (Equations 1 and 2) exhibit intra- and interannual and basin-wide differences (Figure 2). ET has a distinct seasonal cycle with the maximum in July for all basins and years, but the remaining quantities show interannual differences in their seasonal cycle. Runoff generally peaks at around 2.5–4.7 mm day^{-1} in spring (March–May), with the southern basins peaking earlier (March), while the runoff maxima in the Superior and Huron watersheds occurring in May. The rate of change of snow water equivalent ΔSWE in the Superior watershed is positive from November to February (max of 2.7 mm day^{-1}) and negative from March to May (min of −4.1 mm day^{-1}), with snowpack accumulation starting in the fall and steadily declining in the spring. The absolute magnitudes of ΔSWE for the Michigan, Ontario, and Erie watersheds are relatively smaller, ranging between −2.2 and 1.6, −2.5 and 1.7, and −0.9 and 0.8 mm day^{-1}, respectively, while for the Huron basin, ΔSWE ranges between −3.6 and 2.2 mm day^{-1}. The change in the canopy interception storage is orders of magnitude smaller than the other budget quantities (between −0.01 and +0.01 mm day^{-1}; Figure S6e in Supporting Information S1) and does not play a significant role in the water budget seasonal cycle.

For the annual budget, the change in the terrestrial water storage (ΔW) is negligible, as expected on longer timescales, and the water budget is predominantly controlled by the interplay of the terms in the P_t – ET – RO

10.1029/2022WR033759

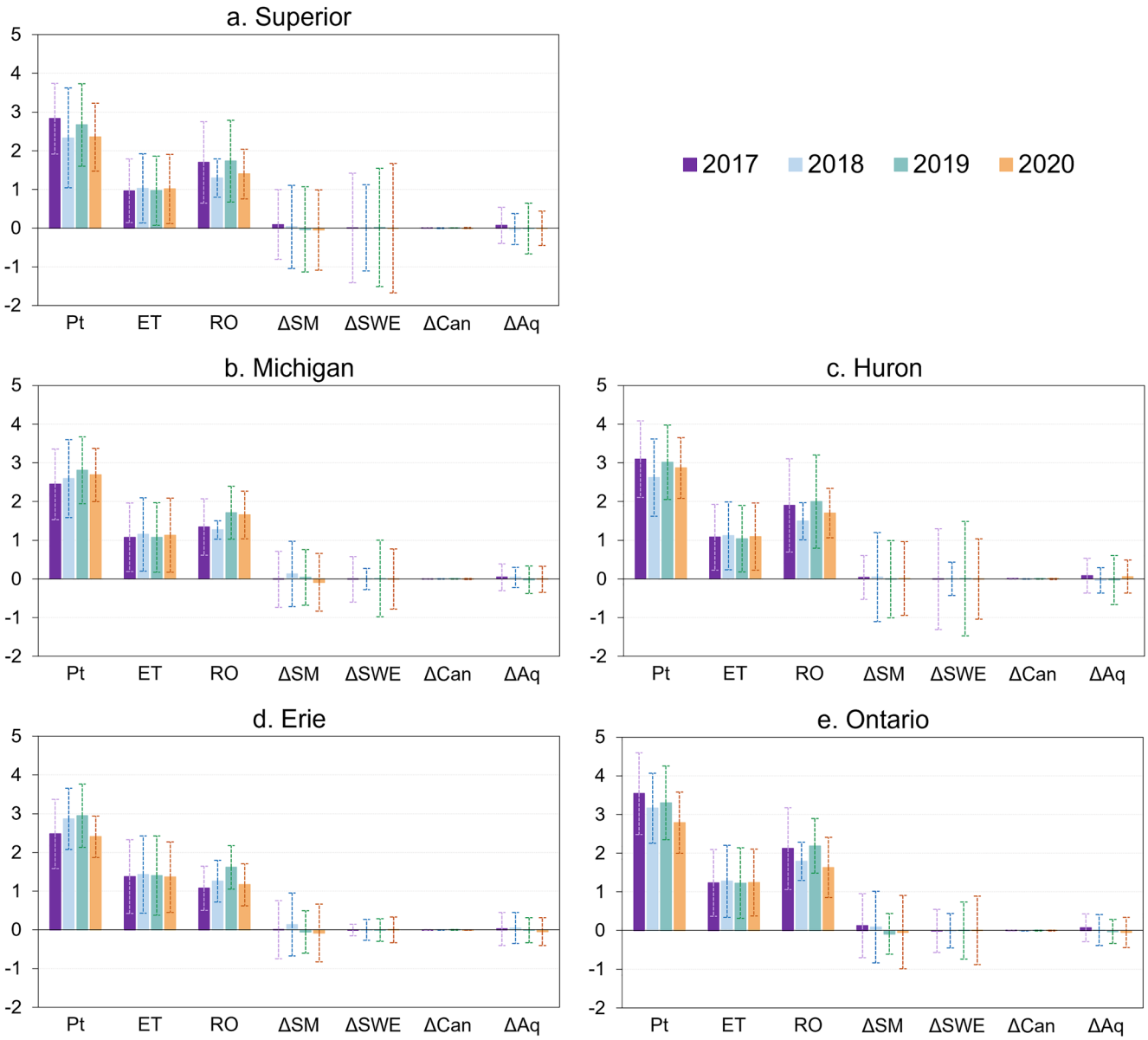


Figure 3. Annual-mean terrestrial water budget (mm day^{-1}) for the four hydrological years in the five watersheds of the Great Lakes region. The solid bars represent the annual mean, and the error bars represent one standard deviation, quantifying the month-to-month variability within each year.

difference (Figure 3). While these three variables dominate the water budget in their absolute magnitudes, the standard deviation is large in the rates of change in soil moisture (ΔSM), snowpack (ΔSWE), and, to a lesser extent, aquifer recharge (ΔAq), highlighting their roles at the subannual timescales. In the Lake Superior watershed, ΔSWE shows a large standard deviation (σ) of 1.1–1.7 mm day^{-1} (Figure 3a), while ΔSWE σ is the smallest in the Erie basin at 0.15–0.33 mm day^{-1} (Figure 3d). For the remaining basins, σ values are within the ranges estimated for these two basins. The standard deviation of ΔSM ranges between 0.5 and 1.1 mm day^{-1}, depending on the basin. In terms of magnitudes, ΔSM is negative for the Superior basin from November to February and from May to July (Table S2 in Supporting Information S1), while for the Michigan and Huron watersheds, ΔSM is negative for only two winter months (December–January) and the three summer months (May–July). Erie and Ontario watersheds have a different pattern of the seasonal cycle, with ΔSM being negative during the April–September period and mostly positive in the remaining months of the year. The seasonal differences in ΔSM dynamics among the basins are controlled by the dynamics of snowpack during colder months, snowmelt in spring, and ET in summer. Winter periods in higher latitudes (November–February in Superior and December–

January in Michigan/Huron watersheds) exhibit higher snow cover, preventing water infiltration and recharge of soil water that slowly drains, thus leading to negative ΔSM. In spring (March–April in Superior and February–April in Michigan/Huron), higher snowmelt and thaw in the top frozen soil layer promote infiltration and contribute to positive ΔSM. In summer (May–July), an increase in ET due to higher amount of available energy results in decreased soil moisture content. The ET dynamics reverse in late summer–autumn period (August–October in Superior and August–November in Michigan/Huron), and the resultant decline in ET drives positive changes in soil water content. Meanwhile, for the southern Erie and Ontario watersheds, the seasonality of ΔSM is primarily driven by ET and less so by ground snowpack, as increasing ET magnitude causes negative ΔSM over the summer months (Figure S6d in Supporting Information S1).

Change in the aquifer recharge also fluctuates from month-to-month, but without a distinct seasonal cycle (Figure 2). Within the modeling scheme, the aquifer recharge represents the residual amount after surface and subsurface runoff, evapotranspiration, and soil moisture are accounted for and is parameterized following the Darcy's law (G.-Y. Niu et al., 2011). It is therefore difficult to ascribe seasonality to this term of the water budget which is likely driven by intra- and interannual variability of soil water excess (i.e., above the soil field capacity).

We note two important implications of these interbasin differences in the seasonal cycles:

1. Generally, snow is considered to have a large significance in the Great Lakes region and consequently snowmelt is an important contributor for the lake levels. However, across the five basins, change in ground snowpack is trivial for Erie (maxima of $<|0.3|$ mm day^{-1}, averaged over the 4 years; Figure S6e in Supporting Information S1), and its significance is only prominent in the Superior and Huron basins, where the absolute maximum in the change in snowpack exceeds 3.3 and 2.5 mm day^{-1}, respectively, in April. The May runoff peak for the northern Superior and Huron basins follows this ΔSWE absolute maxima. The spring snowmelt increases lake levels, and the lag in the ΔSWE across the basins will alter the contribution of each watershed to lake levels.
2. Soil moisture increases substantially for Superior in April (by over 2 mm day^{-1}; Figure S6d in Supporting Information S1) and less so for Huron (∼1 mm day^{-1}), but ΔSM is smaller or even negative for other basins at this time of year. Similarly, aquifer recharge also peaks in April for these two basins (∼0.8 mm day^{-1} for Superior and Huron). This indicates that ground snowpack has a more prominent role in changes in groundwater storage and soil moisture for Superior and Huron, as compared to the other basins.

3.2. Dominant Quantities Inducing Temporal Variability

3.2.1. Subseasonal Timescale

The above seasonal cycle analysis for individual budget variables informs on fluctuations in magnitudes from month-to-month, but without quantifying their impact on the overall budget or on other quantities. PCA links these variables and their time response in the total water budget, highlighting only those quantities that explain the most temporal variation in the terrestrial water system.

In summer (June–September), the first PC accounts for nearly all the variance (>97%) in the five basins (Figure 4). The seasonal transition months of May and October in the southeastern basins (Erie and Ontario) and central basins (Michigan and Huron) are also controlled by the first PC. Generally, in the colder months (November–April), PC2 also explains a large percentage of the variability in the water budget (up to ∼40% for some months). However, there are differences among the basins. For the Lake Superior watershed, which is located at higher latitudes, PC2 contributes more to February–May (25%–42%), whereas in the southernmost Erie basin, PC2 effect is larger in December–February (explaining 27%–40% of the total variance).

Regardless of the monthly and basin-wide differences, only the first two PCs are important in explaining the characteristics of each basin's water budget at subseasonal timescales. The component loadings (correlations of the original seven budget quantities with PC1 and PC2) can provide physical insights into the drivers of variability. In summer and autumn (May–October), PC1 is highly correlated with only two variables in all basins: precipitation (correlation coefficient varying from 0.72 to 0.76, depending on basin) and the change in soil moisture (correlation of 0.63–0.69; Figure 5a). In the colder months (November–April), there are differences amongst the basins and the role of ground snow accumulation emerges. For Superior, the months from December to April have high correlation between PC1 and ΔSWE (0.60–0.82). This is also true for Michigan and Huron. But for Erie and

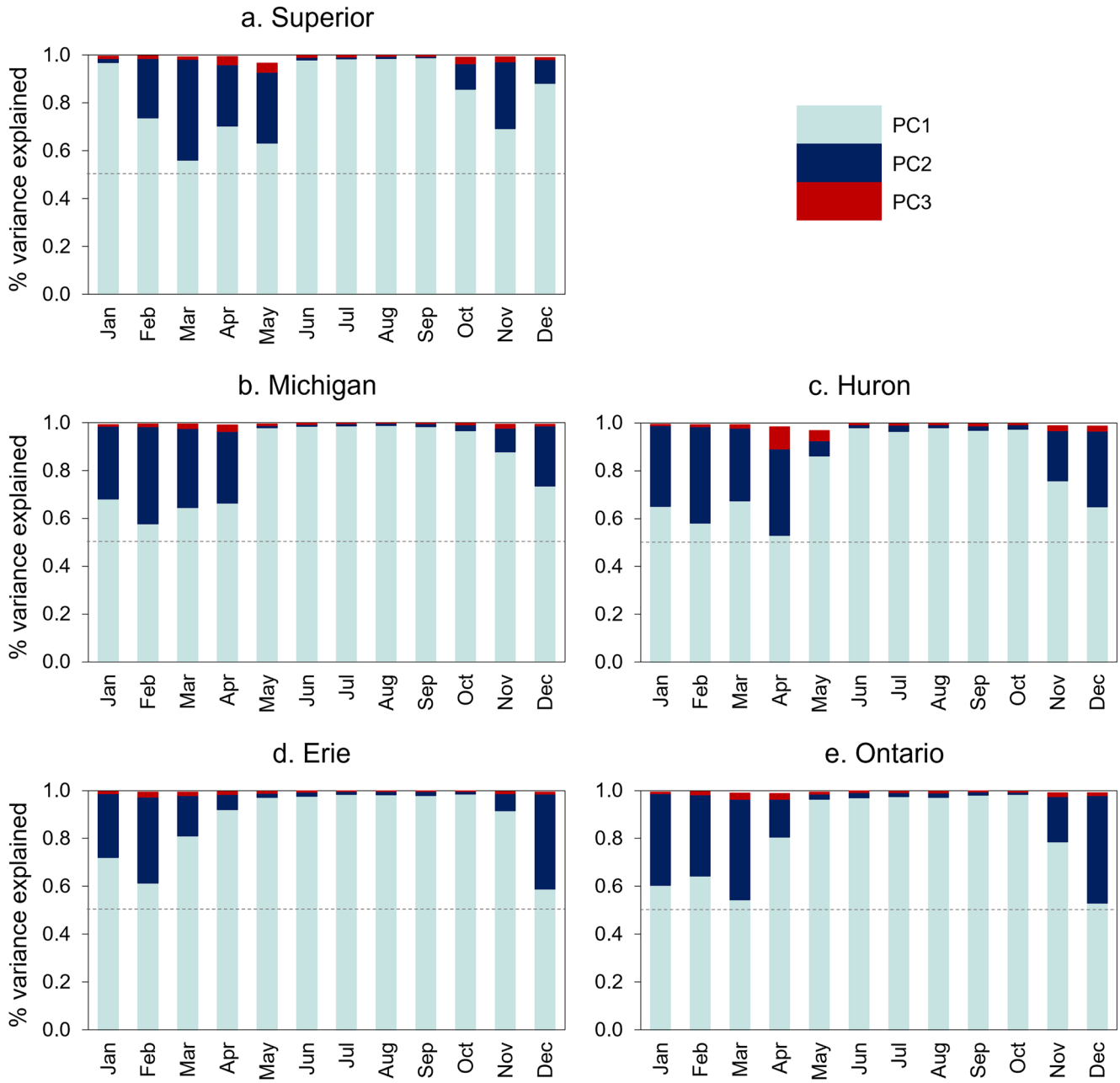


Figure 4. The proportions of variance explained by the first three principal components (PCs) for each month, computed using the daily time series.

Ontario, soil moisture continues to dominate in the winter months as well. In March for Ontario, PC1 is almost entirely correlated with ΔSWE (|0.72|) and ΔSM (0.68), while for other basins, precipitation is generally dominant throughout the year. Interestingly, the second PC, which explains a notable fraction of variance in colder months, is also entirely correlated with these three variables (P_t, ΔSWE, and ΔSM; Figure 5b). PCs by construct are uncorrelated; however, different PCs can be correlated with the same response variables, and this highlights the dominance of these three quantities in driving the submonthly variability within the budget.

The importance of this assessment lies in identifying the dominant budget components that control the variability in the system. In this regard, we highlight several points. (a) Precipitation is important in the budget of all basins year-round, not just in terms of magnitude but also in explaining the variability in the system. (b) The contribution of changes in soil moisture storage to monthly variability of water budget components is dominant for the entire domain in the summer months, and for the southernmost basins year-round. (c) Snowpack variations are

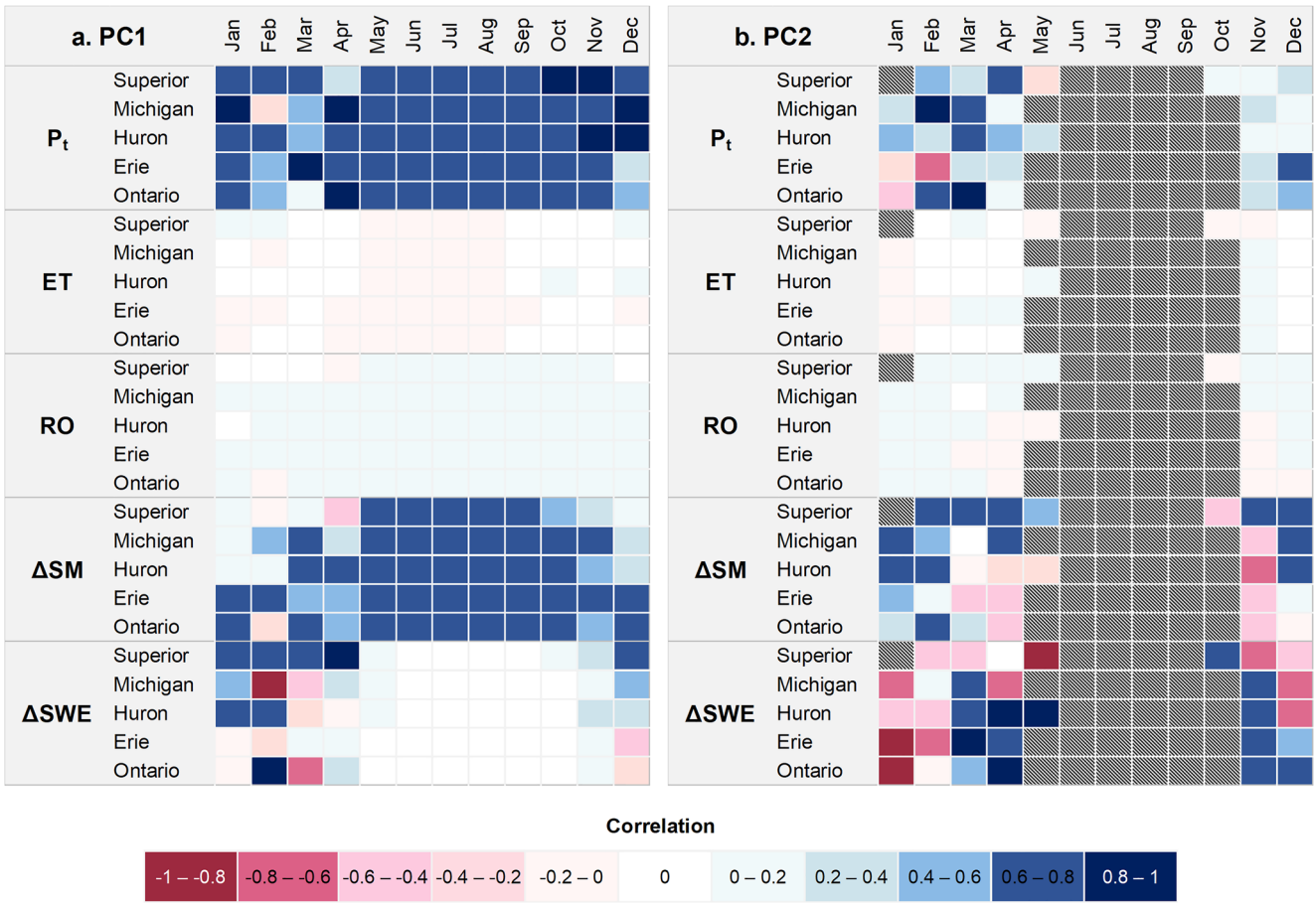


Figure 5. Correlation matrix plot for component loadings, showing the correlation of the five original budget variables with (a) PC1 and (b) PC2. Hatching in (b) shows months where proportion of variance explained by PC2 is negligible (<5%). The full matrix for all seven terrestrial water budget quantities is shown in Figure S7 in Supporting Information S1.

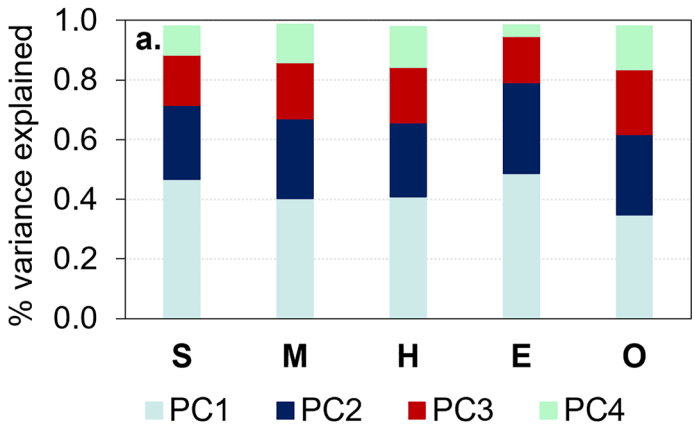
expectedly important for the colder season, especially for the northernmost basins, and less so in the southern watershed such as Erie. (d) ET and runoff, which have relatively larger magnitudes and are important for the annual terrestrial water budget, have little contribution in explaining the variability in the budget.

3.2.2. Seasonal Timescale

At seasonal timescales, the importance of other variables emerges, along with greater differences among the basins. First, a larger number of variables (PCs 1–4) are needed to explain the total variability in the water budget (Figure 6a). The percentage of the variance explained by PC1 ranges from 35% (Lake Ontario watershed) to 49% (Erie) and for PC2 between 25% (Superior and Huron) to 30% (Erie). PCs 3 and 4 also have relatively higher importance, explaining 16%–22% and 4%–15% of the variance, respectively, depending on the watershed.

For each watershed, these PCs are correlated with different variables. In Superior, PC1 is strongly correlated with ΔSWE (correlation coefficient |0.82|), PC2 primarily with P_t (|0.68|), and PC3 with ΔSM (|0.67|) and P_t (|0.55|). PC4, which explains 10% of the variance, is strongly correlated with ET (|0.74|) and moderately with runoff (0.57). While precipitation, snowpack, and soil moisture are still dominant at this timescale, ET and RO now emerge to explain part of the budget variability. For the water budget components of the Lake Michigan watershed, PC1 explains 40% of the variance and correlates with ET (0.74) and P_t (0.53), and PC2 (27%) correlates with ΔSM (0.68) and P_t (0.62). PC3 (18%) strongly correlates with ΔSWE (|0.85|) and PC4 (13%) with runoff (correlation 0.70). For Michigan, precipitation, evaporation, and soil moisture are the dominant variables in the budget, followed by snowpack and runoff.

For the Lake Huron basin, the first two PCs (representing 41%, and 25% of the variance) are correlated with ΔSWE and ΔSM, with correlation of |0.71| and |0.74|, respectively. We also see the roles of precipitation, evapotranspiration, and runoff in PC3 and PC4, which together explain 32% of the variance. Meanwhile, in the Erie



b. Superior	PC1	PC2	PC3	PC4
P_t	0.21	**-0.68**	**-0.55**	0.11
ET	0.18	-0.45	0.19	**-0.74**
RO	0.28	-0.38	0.32	**0.57**
ΔSM	0.37	0.35	**-0.67**	-0.10
ΔSWE	**-0.82**	-0.24	-0.32	0.08
ΔCan	0.00	0.00	0.00	0.00
ΔAq	0.20	0.04	-0.08	0.30

c. Michigan	PC1	PC2	PC3	PC4
P_t	**0.53**	**0.62**	-0.36	0.16
ET	**0.74**	-0.21	0.08	-0.48
RO	0.25	0.16	0.38	**0.70**
ΔSM	-0.28	**0.68**	-0.02	-0.43
ΔSWE	-0.04	-0.22	**-0.85**	0.22
ΔCan	0.00	0.00	0.00	0.00
ΔAq	-0.14	0.21	0.05	0.15

d. Huron	PC1	PC2	PC3	PC4
P_t	0.38	-0.31	**0.68**	0.33
ET	0.16	0.35	**0.59**	**-0.58**
RO	0.48	0.42	-0.06	**0.58**
ΔSM	0.28	**-0.74**	-0.07	-0.20
ΔSWE	**-0.71**	-0.14	0.39	0.40
ΔCan	0.00	0.00	0.00	0.00
ΔAq	0.17	-0.20	-0.17	0.13

e. Erie	PC1	PC2	PC3	PC4
P_t	0.15	**0.87**	-0.08	
ET	**0.82**	0.06	-0.34	
RO	0.07	0.33	**0.76**	
ΔSM	-0.48	0.34	**-0.53**	
ΔSWE	-0.02	-0.01	-0.06	
ΔCan	0.00	0.00	0.00	
ΔAq	-0.26	0.15	0.09	

f. Ontario	PC1	PC2	PC3	PC4
P_t	**-0.76**	0.40	0.15	-0.25
ET	0.19	**0.80**	0.31	0.24
RO	-0.31	0.19	**-0.79**	-0.02
ΔSM	**-0.51**	-0.35	0.43	0.38
ΔSWE	0.05	-0.06	0.26	**-0.86**
ΔCan	0.00	0.00	0.00	0.00
ΔAq	-0.18	-0.20	-0.06	0.01

Figure 6. (a) The proportion of variance explained by the first four principal components (PCs) computed using monthly data over the four hydrological years. (b–f) Correlation matrix plot for component loadings for each basin, where the color and hatching scheme are same as Figure 5. The bold and underlined values show correlation coefficients >0.5 for each PC. P_t is total precipitation, ET is evapotranspiration, RO is total runoff, SM is soil moisture, SWE is snow water equivalent, Can is canopy interception, and Aq is aquifer recharge.

watershed, P_t, ET, and RO quantities, and to a lesser degree ΔSM, dominate the budget variability, while for the basin of Lake Ontario, ΔSWE is highly correlated with PC4 (|0.86|), which explains 15% of the total budget variance.

We can infer from these results that in the northern regions, snowpack and precipitation are the most important variables determining the subannual/seasonal variability of the water budget, similar to subseasonal timescales.

In the lower latitudes, however, the role of evapotranspiration and soil moisture emerges, while for the southernmost basins, the precipitation–evapotranspiration–runoff combination represents the dominant source of variability at these timescales. Some of these results align with the findings of Syed et al. (2004), who determined the importance of precipitation and potential evapotranspiration across CONUS and of soil moisture in some parts of the Midwest. However, due to the broad spatial scale of their assessment, the role of snowpack was not discussed, which is an important component to consider for the Laurentian region. In our assessment, the differences across the five basins are prominent, where Lake Superior has a snow-dominated regime, while southern basins' water balance is controlled by the standard *P*–ET–RO relationship prevalent in most of CONUS. Thus, individual catchment characteristics are more important at this timescale to establish climate change effects on the water balance and terrestrial hydrology.

3.3. Relationship Among the Terrestrial Budget Quantities

To further understand the differences in the catchment characteristics at the seasonal timescale, we establish the influence of each quantity in inducing the variability within the remaining budget quantities using the PLSR approach (Figure 7).

The primary quantity which explains the most variance in evapotranspiration varies across the latitudes, from north to south: precipitation for Superior and Michigan, change in soil moisture for Erie and Ontario, while for the Huron basin, the results are less definitive. The VIP scores show a bidirectionality of the links, for example, the ΔAq – ET dependence in Huron (Figure 7a). In this case, the PLSR results reveal a correlation between the two, where ET would be responsible in drying out the aquifer, allowing less recharge. The change in aquifer recharge can be thus a predictor in ET fluctuations. This is also true for Erie and Ontario, where ΔAq is the second important quantity in explaining the ET variability. In Superior, expectedly, ΔSWE is also an important variable for ET, while for Michigan, ΔSM – ET is the second strongest dependence.

Runoff dependence is more definitive for the five basins, with ΔSWE as the primary variable for Superior, Michigan, and Huron, as inferred from the budget seasonality as well (Section 3.1). In the southern basins, precipitation is the main driver of variability in runoff. The RO – ΔSM interdependence is strong for all the basins as the quantity with second highest VIP score (Figure 7b) and we see this bidirectionality in the relationship between ΔSM – RO (Figure 7c). Change in soil moisture fluctuates with ET primarily within Michigan and Erie watersheds and precipitation within Huron and Ontario watersheds, while in Superior, the ΔSM variability is mostly driven by change in ground snowpack. In Ontario, ΔSM primarily drives variability in ET (Figure 7a), but the role of ET is weaker in explaining the ΔSM variability (Figure 7c).

The change in aquifer recharge is interesting and an important indicator for the fluctuations in groundwater reserves (Figure 7d). ΔAq changes in response to ET and ΔSM in all basins, except Superior where ΔSWE again has the dominant role, highlighting the significance of its snow-dominated hydrological cycle. We can assume that other watersheds in the northern regions of the North America will have similar snow-dominated characteristics. Warming scenarios will shift the regime to behave more like the southern Great Lakes watersheds, and conversely, long-term cooling trends in the region will alter characteristics to match that of Lake Superior's. This has strong implications for climate change assessments for the regional water resources.

3.4. Effects of Soil Types and Land Cover Categories

Besides the temperature-dependent basin characteristics, soil types and land cover categories can also influence the water budget variability. We again focus on the seasonal timescale and conduct PCA for the soil types in Figure 1b and land cover categories in Figure 1c, which are input in the Noah-MP model. In this case, to exclude the impact of the snow-heavy characteristics of the northern latitudes, we focus only on the midlatitudes of our modeling domain, approximately between 40° and 44°N.

PCA of soil categories shows nearly the same results for the five main soil categories in the region under consideration: (a) loam, (b) clay loam, (c) sandy loam, (d) silt loam, and (e) clay or silty clay, where the first three PCs explain nearly all the variability in the budget of these domains (Figure 8). The first PC proportion ranges between 37% for sandy loam, 46% and 47% for loam and clay loam, and 54% and 56% for silt loam and clay/silty clay. It is mainly correlated with evapotranspiration, with the correlation coefficients larger than 0.75 for all soil

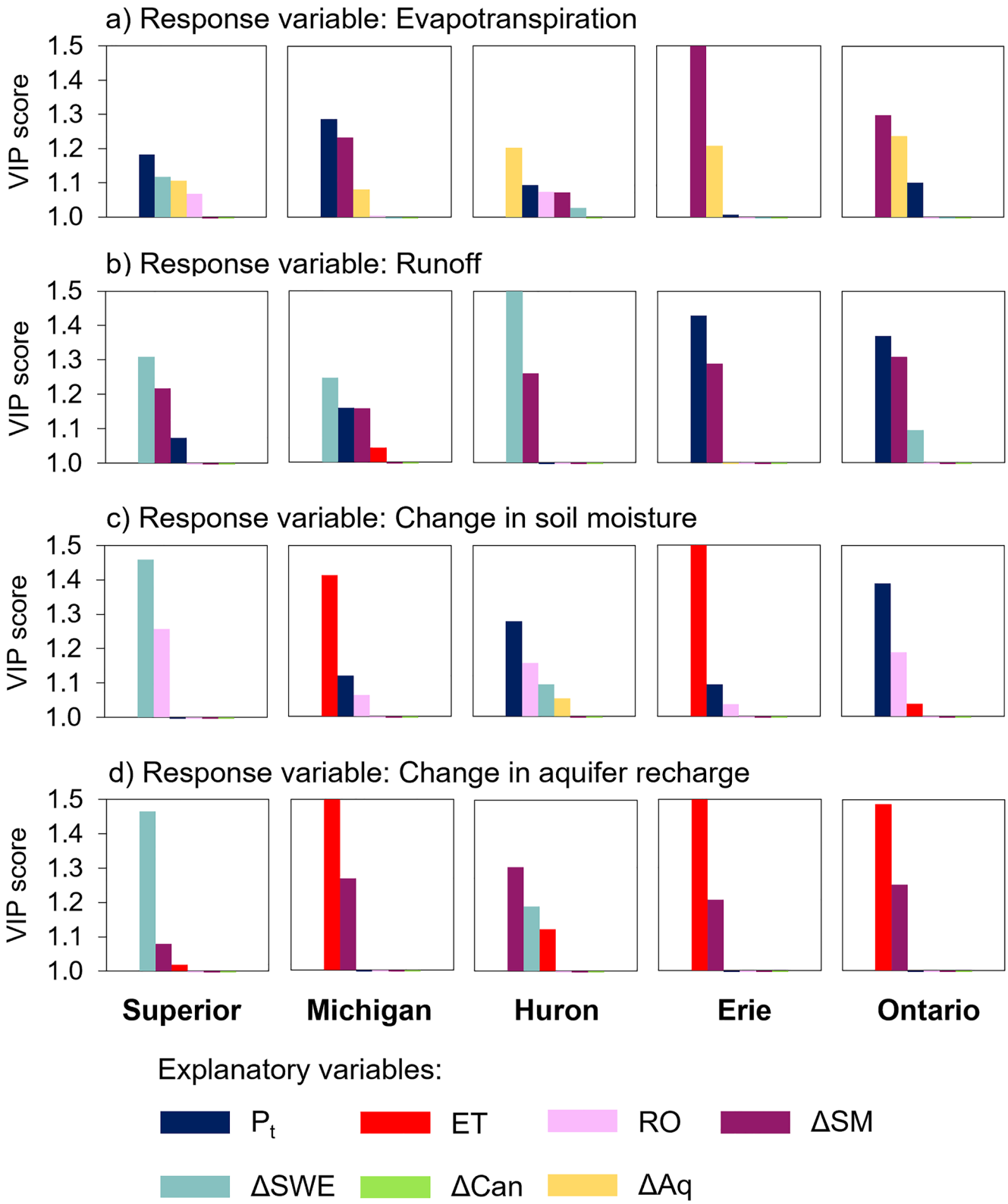


Figure 7. Variable importance in projection (VIP) scores from the partial least squares regression (PLSR) analysis for four response variables (a) ET, (b) RO, (c) ΔSM, and (d) ΔAq, while the seven explanatory variables are shown in various colors in the legend. PLSR is done using the monthly time series, similar to Figure 6, for each basin. Only VIP scores greater than 1 are shown on the *y*-axis.

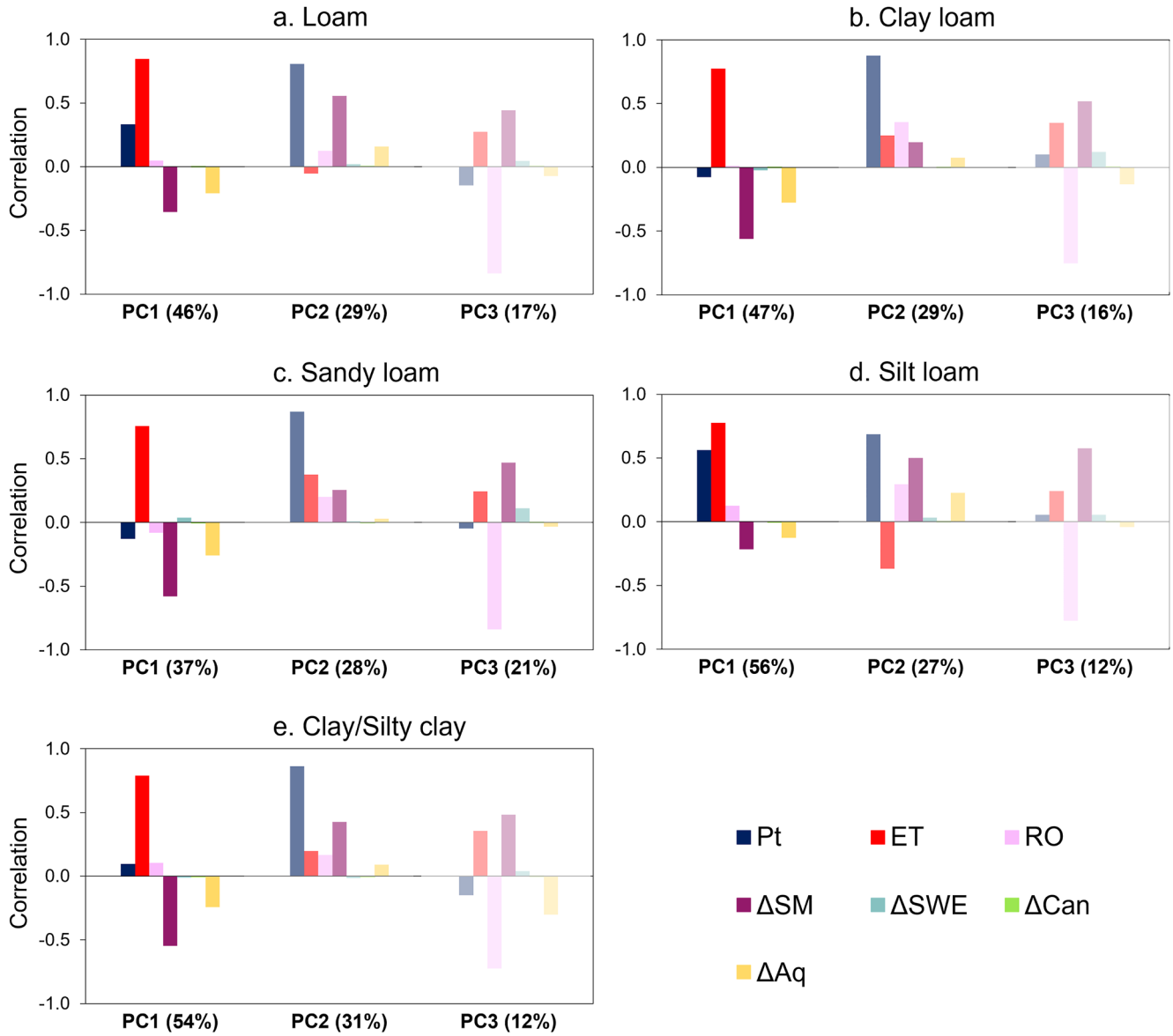


Figure 8. Principal component analysis (PCA) of the five main soil types for the domain between 40° and 44°N in Figure 1b: (a) loam, (b) clay loam, (c) sandy loam, (d) silt loam, and (e) clay and silty clay. PCA is conducted using the monthly time series, similar to Figure 6. The proportion of variance explained by the first three principal components (PCs) is shown on the *x*-axis, while the *y*-axis shows the correlation of each PC with the original budget quantities (colored bars).

categories. The second PC is mainly correlated with precipitation and PC3 with runoff for all soil types. This again highlights the *P*–ET–RO dominant regime of the southern basins, where the different soil categories have negligible impact on the budget variability.

On the other hand, PCA of land cover has differences among the four main categories under consideration: (a) deciduous broadleaf forest, (b) cropland and natural vegetation mosaics, (c) croplands only, and (d) urban areas (Figure 9). Natural vegetation and croplands have similar results, where PC1 strongly correlates with ET (correlation coefficient >0.80), PC2 with precipitation (correlation >0.75), and PC3 with runoff (correlation >0.80). In the International Geosphere–Biosphere Programme (IGBP)-modified MODIS vegetation/land use categories input in the model, croplands include at least 60% of cultivated area, while the natural vegetation has overlap with croplands and is defined as mosaics of small-scale cultivation covering 40%–60% area with natural trees, shrubs, or herbaceous vegetation (Sulla-Menashe & Friedl, 2018). Therefore, the similarities between the two categories are expected. The region with deciduous broadleaf forests has precipitation as the main variable corresponding

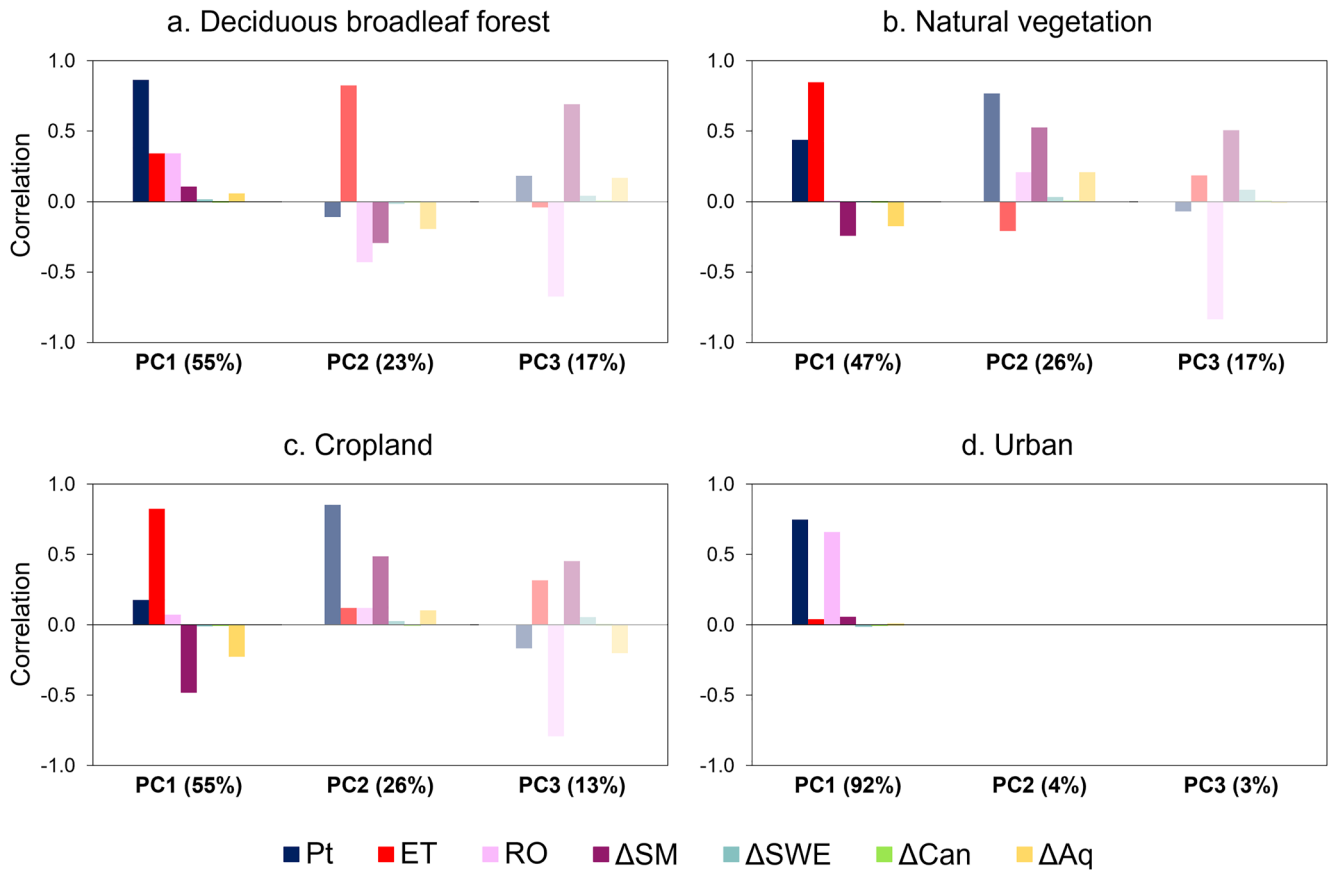


Figure 9. Principal component analysis (PCA), similar to Figure 8, of select land cover categories for the domain between 40° and 44°N in Figure 1c: (a) deciduous broadleaf forest, (b) natural vegetation, (c) cropland, and (d) urban areas.

with PC1 (explaining 55% of the variability), ET with PC2 (23% variability), and both RO and ΔSM nearly equally correlated with PC3, which explains only 17% of the variability. Urban areas, which are classified as having at least 30% impervious surface area, have definitive PCA results where precipitation–runoff are the only two dominant variables in the budget. This confirms a well-known hydrological regime of urban and built-up areas, where nearly all precipitation results in surface runoff.

Unlike the soil types, land cover has a more visible impact on the drivers of variability in the water budget and therefore we can anticipate a role of land cover/land use changes in altering hydrological regimes of regions.

4. Conclusions

Assessing the hydrometeorological variability at different temporal resolutions is important to improve water budget estimates for the hydroclimate of the North American Great Lakes basin. This study identified the terrestrial water budget quantities with the greatest variability at seasonal and subseasonal timescales and characterized the basin-specific hydrology of the region. We further established the relative importance of the essential variables in the regional terrestrial hydrological system and quantified the relationships among the water budget components in the Great Lakes subregions.

We found that the Great Lakes basins do not have hydrologically uniform regimes, and there is a substantial variability in the key drivers. Depending on the basin and period of the year, different variables play dominant roles in characterizing the hydrological cycle, while the drivers of dominant modes of hydrologic variability differ for daily, monthly, and annual budgets. The annual water budget is dominated by the contributions from the fluxes of precipitation, evapotranspiration, and runoff. PCA over the seasonal timescale shows that each basin is unique; for example, runoff and evapotranspiration are important contributors to monthly variability for

Lake Erie and Lake Ontario water budgets but play a smaller role for the Lake Superior basin. Change in the soil moisture storage is equally important for both the Lakes' Michigan and Huron basins, with the other dominant quantities being evapotranspiration and precipitation for Michigan, and snowpack for the Lake Huron watershed. At subseasonal timescales, we see a convergence in the behavior of all basins where precipitation and soil moisture are the main contributors to the variability, together with ground snowpack in the winter months, especially at higher latitudes.

The catchments at northern latitudes significantly differ from the standard rainfall-runoff relationships (Wang et al., 2016) prevalent in the Erie and Ontario basins and the rest of CONUS. In these snow-dominated regimes, variability in runoff, soil moisture, and aquifer recharge is a response to the variability in the ground snowpack. In southern regions of the Great Lakes basin, evapotranspiration and precipitation drive variability in the soil moisture, while aquifer recharge variability is explained by evapotranspiration and soil moisture. For surface runoff, which is a key component of the lake net basin supply, a clear distinction is present across the basins, where Superior and Huron are ground snowpack–soil moisture driven basins, while Erie and Ontario are rain–soil moisture modulated basins. These interbasin differences at seasonal timescales signify the different hydrological characteristics across the region and how that alters the water storage and availability for lake levels, groundwater reserves, and soil moisture. The later consequently impacts cropping patterns and vegetation dynamics of the basins. Climate change impact studies on the regional hydrology must account for these differences, as the subregional terrestrial hydrology is characterized by different variables, implying that each basin's vulnerability to climate change will differ as a result.

We also found that soil types have only a nominal role in explaining the spatial differences of the hydrological characteristics; however, land cover plays a more important role where the dominant quantities driving variability depends on how the land surface is categorized in the model. Future changes in land use/land cover will impact the terrestrial hydrology, and large-scale urbanization especially needs to be considered for local catchment hydrology.

This study has two implications for future research. First, it highlights the processes that require better consideration in modeling and assessing the hydrological cycle for the Great Lakes basin. This helps identify important variables, specifically changes in soil moisture and ground snowpack, which require an improved spatiotemporal observational network and component development in LSMs. Second, in a warming world, we can anticipate that the importance of these hydroclimatic drivers will change over time, including the emergence of processes that historically played inconsequential roles in the regional hydrological regimes. The terrestrial water cycle is intensifying in this region and globally (Huntington, 2006; Huntington et al., 2018), with the evapotranspiration rates projected to increase in the coming decades (Minallah & Steiner, 2021a) and soil moisture variability expected to change in response to warming (Green et al., 2019; Zhou et al., 2021). With these changes, we can expect a southward shift of hydroclimatic behavior, such that Lakes' Michigan and Huron water budgets have the potential to become similar to Lakes' Erie/Ontario, and the Lake Superior hydrological budget may transform to become similar to that of Lake Michigan watershed. Seasonal and spatial shifts in the hydroclimate will have implications for local ecosystems, biodiversity, soil aridification, agriculture, and water availability (Hayhoe et al., 2010). Soil moisture variability is a major control in land carbon uptake which in turns effects the gross primary production (Green et al., 2019). Similarly, seasonal shifts in different water "buckets" (e.g., snowpack and soil moisture) will alter the vegetation dynamics. Therefore, quantifying the sources of variability at different timescales in the current regime is crucial to assess catchment sensitivities because it highlights the region-specific mechanisms that are suppressed over annual averages and only become dominant at specific periods. This study provides a baseline understanding of the geographic variation in hydroclimate variable relationships, which informs strategies to determine shifts in signatures of the variability drivers, assess modifications in future hydrological budgets, and project the consequences of emergent processes.

Data Availability Statement

The WRF-Hydro modeling system is an open-source model, available at https://ral.ucar.edu/projects/wrf_hydro. The model output is openly accessible in the University of Michigan Deep Blue Repository as "Land surface hydrology data for the North American Great Lakes region" (Minallah & Steiner, 2022).

Acknowledgments
We thank Dr David Wright for guidance on the meteorological forcing data, Dr Yiwen Mei for WRF-Hydro model support, Dr Donghui Xu for providing the processed USGS gage data, and Dr Andrew Gronewold for sharing insights on the regional hydrology literature. This work was supported by the National Science Foundation (NSF) Grant OCE-1600012 to ALS, with additional support for SM by the University of Michigan Rackham Predoctoral Fellowship. VYI acknowledges the support of NSF CMMI Award 2053429. AWW was supported by research grants from the US Bureau of Reclamation Science and Technology Office, and the US Army Corps of Engineers Climate Preparedness and Resilience Program.

References

Baker, S. A., Wood, A. W., & Rajagopalan, B. (2020). Application of postprocessing to watershed-scale subseasonal climate forecasts over the contiguous United States. *Journal of Hydrometeorology*, *21*(5), 971–987. https://doi.org/10.1175/JHM-D-19-0155.1

Barlage, M., Tewari, M., Chen, F., Miguez-Macho, G., Yang, Z.-L., & Niu, G.-Y. (2015). The effect of groundwater interaction in North American regional climate simulations with WRF/Noah-MP. *Climatic Change*, *129*(3–4), 485–498. https://doi.org/10.1007/s10584-014-1308-8

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*, 180214. https://doi.org/10.1038/sdata.2018.214

Bennington, V., Notaro, M., & Lofgren, B. (2015). Dynamical downscaling–based projections of Great Lakes water levels. *Journal of Climate*, *28*(24), 9721–9745. https://doi.org/10.1175/JCLI-D-14-00847.1

Black, J., Johnson, N. C., Baxter, S., Feldstein, S. B., Harnos, D. S., & L'Heureux, M. L. (2017). The predictors and forecast skill of Northern Hemisphere teleconnection patterns for lead times of 3–4 weeks. *Monthly Weather Review*, *145*(7), 2855–2877. https://doi.org/10.1175/mwr-d-16-0394.1

Chang, L.-C., Liou, J.-Y., & Chang, F.-J. (2022). Spatial–temporal flood inundation nowcasts by fusing machine learning methods and principal component analysis. *Journal of Hydrology*, *612*, 128086. https://doi.org/10.1016/j.jhydrol.2022.128086

Chong, I.-G., & Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, *78*(1–2), 103–112. https://doi.org/10.1016/j.chemolab.2004.12.011

Do, H. X., Smith, J. P., Fry, L. M., & Gronewold, A. D. (2020). Seventy-year long record of monthly water balance estimates for Earth's largest lake system. *Scientific Data*, *7*(1), 276. https://doi.org/10.1038/s41597-020-00613-z

Durnford, D., Fortin, V., Smith, G. C., Archambault, B., Deacu, D., Dupont, F., et al. (2018). Toward an operational water cycle prediction system for the Great Lakes and St. Lawrence River. *Bulletin of the American Meteorological Society*, *99*(3), 521–546. https://doi.org/10.1175/BAMS-D-16-0155.1

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The Soil Moisture Active Passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/jproc.2010.2043918

Fang, Z., Bogena, H., Kollet, S., Koch, J., & Vereecken, H. (2015). Spatio-temporal validation of long-term 3D hydrological simulations of a forested catchment using empirical orthogonal functions and wavelet coherence analysis. *Journal of Hydrology*, *529*, 1754–1767. https://doi.org/10.1016/j.jhydrol.2015.08.011

Fortin, V., Deacu, D., Klyszejko, E., Spence, C., & Blanken, P. D. (2012). Predicting the net basin supply to the Great Lakes with a hydrometeorological model. *Journal of Hydrometeorology*, *13*(6), 1739–1759. https://doi.org/10.1175/JHM-D-11-0151.1

Fu, Q., Lin, P., Wallace, J. M., & Smoliak, B. V. (2015). Dynamical adjustment of the Northern Hemisphere surface air temperature field: Methodology and application to observations. *Journal of Climate*, *28*(4), 1613–1629. https://doi.org/10.1175/JCLI-D-14-00111.1

Global Modeling and Assimilation Office. (2015). MERRA-2 tavg1_2d_lnd_Nx: 2d,1-hourly,time-averaged,single-level,assimilation,land surface diagnostics V5.12.4. https://doi.org/10.5067/RKPHT8KC1Y1T

Gochis, D. J., Barlage, M., Cabell, R., Casali, M., Dugger, A., FitzGerald, K., et al. (2020). The NCAR WRF-Hydro modeling system technical description (Version 5.2.0).

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., Lawrence, D. M., & Gentine, P. (2019). Large influence of soil moisture on long-term terrestrial carbon uptake. *Nature*, *565*(7740), 476–479. https://doi.org/10.1038/s41586-018-0848-x

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, *27*(9), 1119–1152. https://doi.org/10.1002/joc.1499

Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, *7*(1), 109. https://doi.org/10.1038/s41597-020-0453-3

Hayhoe, K., VanDorn, J., Croley, T., Schlegal, N., & Wuebbles, D. (2010). Regional climate change projections for Chicago and the US Great Lakes. *Journal of Great Lakes Research*, *36*, 7–21. https://doi.org/10.1016/j.jglr.2010.03.012

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology*, *319*(1–4), 83–95. https://doi.org/10.1016/j.jhydrol.2005.07.003

Huntington, T. G., Weiskel, P. K., Wolock, D. M., & McCabe, G. J. (2018). A new indicator framework for quantifying the intensity of the terrestrial water cycle. *Journal of Hydrology*, *559*, 361–372. https://doi.org/10.1016/j.jhydrol.2018.02.048

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Lahmers, T. M., Hazenberg, P., Gupta, H., Castro, C., Gochis, D., Dugger, A., et al. (2021). Evaluation of NOAA National Water Model parameter calibration in semiarid environments prone to channel infiltration. *Journal of Hydrometeorology*, *22*(11), 2939–2969. https://doi.org/10.1175/JHM-D-20-0198.1

Levia, D. F., Creed, I. F., Hannah, D. M., Nanko, K., Boyer, E. W., Carlyle-Moses, D. E., et al. (2020). Homogenization of the terrestrial water cycle. *Nature Geoscience*, *13*(10), 656–658. https://doi.org/10.1038/s41561-020-0641-y

Li, X., Zhong, S., Bian, X., Heilman, W. E., Luo, Y., & Dong, W. (2010). Hydroclimate and variability in the Great Lakes region as derived from the North American regional reanalysis. *Journal of Geophysical Research*, *115*(D12), D12104. https://doi.org/10.1029/2009JD012756

Mailhot, E., Music, B., Nadeau, D. F., Frigon, A., & Turcotte, R. (2019). Assessment of the Laurentian Great Lakes' hydrological conditions in a changing climate. *Climatic Change*, *157*(2), 243–259. https://doi.org/10.1007/s10584-019-02530-6

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Mascaro, G., Vivoni, E. R., & Méndez-Barroso, L. A. (2015). Hyperresolution hydrologic modeling in a regional watershed and its interpretation using empirical orthogonal functions. *Advances in Water Resources*, *83*, 190–206. https://doi.org/10.1016/j.advwatres.2015.05.023

Mendoza, P. A., Wood, A. W., Clark, E., Rothwell, E., Clark, M. P., Nijssen, B., et al. (2017). An intercomparison of approaches for improving operational seasonal streamflow forecasts. *Hydrology and Earth System Sciences*, *21*(7), 3915–3935. https://doi.org/10.5194/hess-21-3915-2017

Minallah, S., & Steiner, A. L. (2021a). Analysis of the atmospheric water cycle for the Laurentian Great Lakes region using CMIP6 models. *Journal of Climate*, *34*(12), 4693–4710. https://doi.org/10.1175/JCLI-D-20-0751.1

Minallah, S., & Steiner, A. L. (2021b). Role of the atmospheric moisture budget in defining the precipitation seasonality of the Great Lakes region. *Journal of Climate*, *34*(2), 643–657. https://doi.org/10.1175/JCLI-D-19-0952.1

Minallah, S., & Steiner, A. L. (2022). Land surface hydrology data for the North American Great Lakes region [Dataset]. University of Michigan - Deep Blue Data. https://doi.org/10.7302/hry7-kv95

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Music, B., Frigon, A., Lofgren, B., Turcotte, R., & Cyr, J.-F. (2015). Present and future Laurentian Great Lakes hydroclimatic conditions as simulated by regional climate models with an emphasis on Lake Michigan–Huron. *Climatic Change*, *130*(4), 603–618. https://doi.org/10.1007/s10584-015-1348-8

NCEP. (2015). NCEP North American Mesoscale (NAM) 12 km analysis. https://doi.org/10.5065/G4RC-1N91

Niu, G.-Y., Yang, Z.-L., Dickinson, R. E., Gulden, L. E., & Su, H. (2007). Development of a simple groundwater model for use in climate models and evaluation with gravity recovery and climate experiment data. *Journal of Geophysical Research*, *112*, D07103. https://doi.org/10.1029/2006JD007522

Niu, G.-Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *Journal of Geophysical Research*, *116*, D12109. https://doi.org/10.1029/2010JD015139

Niu, J., Shen, C., Li, S.-G., & Phanikumar, M. S. (2014). Quantifying storage changes in regional Great Lakes watersheds using a coupled subsurface-land surface process model and GRACE, MODIS products. *Water Resources Research*, *50*(9), 7359–7377. https://doi.org/10.1002/2014WR015589

Perry, M. A., & Niemann, J. D. (2008). Generation of soil moisture patterns at the catchment scale by EOF interpolation. *Hydrology and Earth System Sciences*, *12*(1), 39–53. https://doi.org/10.5194/hess-12-39-2008

Qiao, Y., Zhan, C., Yang, P., & Xia, J. (2019). Analysis of changes in drought and terrestrial water storage in the Tarim River Basin based on principal component analysis. *Hydrology Research*, *50*(2), 761–777. https://doi.org/10.2166/nh.2019.033

Santos, J. F., Pulido-Calvo, I., & Portela, M. M. (2010). Spatial and temporal variability of droughts in Portugal. *Water Resources Research*, *46*(3), W03503. https://doi.org/10.1029/2009WR008071

Smoliak, B. V., Wallace, J. M., Stoelinga, M. T., & Mitchell, T. P. (2010). Application of partial least squares regression to the diagnosis of year-to-year variations in Pacific Northwest snowpack and Atlantic hurricanes. *Geophysical Research Letters*, *37*(3), L03801. https://doi.org/10.1029/2009GL041478

Sulla-Menashe, D., & Friedl, M. A. (2018). *User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product* (Vol. 1, p. 18). USGS.

Syed, T. H., Lakshmi, V., Paleologos, E., Lohmann, D., Mitchell, K., & Famiglietti, J. S. (2004). Analysis of process controls in land surface hydrological cycle over the continental United States. *Journal of Geophysical Research*, *109*(D22), D22105. https://doi.org/10.1029/2004JD004640

Taşan, M., Demir, Y., & Taşan, S. (2022). Groundwater quality assessment using principal component analysis and hierarchical cluster analysis in Alaçam, Turkey. *Water Supply*, *22*(3), 3431–3447. https://doi.org/10.2166/ws.2021.390

Vergopolan, N., Chaney, N. W., Beck, H. E., Pan, M., Sheffield, J., Chan, S., & Wood, E. F. (2020). Combining hyper-resolution land surface modeling with SMAP brightness temperatures to obtain 30-m soil moisture estimates. *Remote Sensing of Environment*, *242*, 111740. https://doi.org/10.1016/j.rse.2020.111740

Wang, C., Wang, S., Fu, B., & Zhang, L. (2016). Advances in hydrological modelling with the Budyko framework. *Progress in Physical Geography: Earth and Environment*, *40*(3), 409–430. https://doi.org/10.1177/0309133315620997

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 109–130. https://doi.org/10.1016/s0169-7439(01)00155-1

Woodruff, C. D., & Qualls, R. J. (2019). Recurrent snowmelt pattern synthesis using principal component analysis of multiyear remotely sensed snow cover. *Water Resources Research*, *55*(8), 6869–6885. https://doi.org/10.1029/2018WR024546

Wu, W.-Y., Yang, Z.-L., & Barlage, M. (2021). The impact of Noah-MP physical parameterizations on modeling water availability during droughts in the Texas–Gulf region. *Journal of Hydrometeorology*, *22*, 1221–1233. https://doi.org/10.1175/JHM-D-20-0189.1

Xu, D., Ivanov, V. Y., Li, X., & Troy, T. J. (2021). Peak runoff timing is linked to global warming trajectories. *Earth's Future*, *9*(8), e2021EF002083. https://doi.org/10.1029/2021EF002083

Yeh, P. J. F., & Wu, C. (2018). Recent acceleration of the terrestrial hydrologic cycle in the U.S. Midwest. *Journal of Geophysical Research: Atmospheres*, *123*(6), 2993–3008. https://doi.org/10.1002/2017JD027706

Zeinalzadeh, K., & Rezaei, E. (2017). Determining spatial and temporal changes of surface water quality using principal component analysis. *Journal of Hydrology: Regional Studies*, *13*, 1–10. https://doi.org/10.1016/j.ejrh.2017.07.002

Zhang, X., Wu, F., & Hu, T. (2007). Rainfall–runoff modeling using principal component analysis and neural network. *Hydrology Research*, *38*(3), 235–248. https://doi.org/10.2166/nh.2007.010

Zhou, S., Williams, A. P., Lintner, B. R., Berg, A. M., Zhang, Y., Keenan, T. F., et al. (2021). Soil moisture–atmosphere feedbacks mitigate declining water availability in drylands. *Nature Climate Change*, *11*(1), 38–44. https://doi.org/10.1038/s41558-020-00945-z