NOAA Technical Memorandum GLERL-167

Reaffirmation of large biases in a long-used method for projecting changes in Great Lakes water levels in climate change scenarios

Brent M. Lofgren
Jonathan Rouhana*

NOAA Great Lakes Environmental Research Laboratory, 4840 S. State Rd., Ann Arbor, MI 48108
*current affiliation: University of Notre Dame, South Bend, Indiana

September 2015, revised May 2016

UNITED STATES
DEPARTMENT OF COMMERCE

Penny Pritzker
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION

Kathy Sullivan
Acting, Under Secretary for Oceans & Atmosphere
NOAA Administrator

NOTICE

This publication is available as a PDF file and can be downloaded from GLERL's web site: www.glerl.noaa.gov. Hard copies can be requested from GLERL Information Services,
4840 S. State Rd., Ann Arbor, MI 48108.
pubs.glerl@noaa.gov.

NOAA's Mission – To understand and predict changes in Earth's environment and conserve and manage coastal and marine resources to meet our nation's economic, social, and environmental needs

NOAA's Mission Goals:

- Protect, restore and manage the use of coastal and ocean resources through an ecosystem approach to management
- Understand climate variability and change to enhance society's ability to plan and respond
- Serve society's needs for weather and water information
- Support the Nation's commerce with information for safe, efficient, and environmentally sound transportation
- Provide critical support for NOAA's Mission

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Reaffirmation of large biases in a long-used method for projecting changes in Great Lakes water levels in climate change scenarios

Brent M. Lofgren
Jonathan Rouhana

ABSTRACT. A method for projecting the water levels of the Laurentian Great Lakes under scenarios of human-caused climate change, used almost to the exclusion of other methods in the past, relies very heavily on the Large Basin Runoff Model (LBRM) as a component for determining the water budget for the lake system. This model uses near-surface air temperature as a primary predictor of evapotranspiration (ET); in addition to previous published work, we show here again that its very high sensitivity to temperature makes it overestimate ET in a way that is highly inconsistent with the fundamental principle of conservation of energy at the land surface. Under the traditional formulation, the quantity that has been called "energy available for evapotranspiration," which is proportional to what we call "potential evapotranspiration" (PET), is increased by large factors in future scenarios—by a factor of nearly 600 in the Lake Superior basin under forcing by one GCM case, but more typically by factors between 3 and 10. Because of the way that LBRM is formulated and calibrated, these factors can be thought of as corresponding to the factor of increase in solar radiation incident on the Earth, or, more vividly, as the number of Suns present in the sky of the virtual world simulated by LBRM. Therefore, we have created alternative formulations, which we regard as more reflective of what is being simulated by the driving GCMs, for the way that climate change is ingested into the modeling system that includes LBRM. In addition to the energy adjustment method in which PET is increased by an amount proportional to the change in net radiative energy at the surface, we add the Priestley-Taylor method, which augments the energy adjustment method by inclusion of a temperature-dependent factor with rigorous theoretical underpinnings that is a much weaker function of temperature than in the LBRM's basic formulation, as well as the Clausius-Clapeyron method, in which the PET is increased by an amount proportional to the increase in water vapor capacity of the atmosphere, again a weaker function of temperature than in the LBRM. We establish that all three of these alternative methods show, relative to the traditional method, often astoundingly less PET and less ET, more runoff from the land and net basin supply for the lake basins, and higher lake water levels in the future. The magnitude of these discrepancies is highly correlated with the air temperature change in the driving GCM (larger temperature changes lead to larger discrepancies). Using various methods of estimating the statistical significance, we find that, at minimum, these discrepancies in results are significant at the 99.998% level.

1. INTRODUCTION

The Laurentian Great Lakes of North America contain approximately 20% of the world's surface fresh water. They support intrinsic and commercial interests for shoreline property owners, for ecosystems in the lakes themselves and coastal wetlands, and for commercial and recreational fishing, boating, shipping, and tourism. They are a binational resource of the United States and Canada, and over 30 million people live within their drainage basin (Government of Canada and U.S. Environmental Protection Agency 1995).

The issue of the potential influence of future climate change on the water levels of the Great Lakes has been a topic of a significant body of literature. Many examples have used the method pioneered by Croley (1990). A partial list of this class includes Hartmann (1990), Chao (1999), Lofgren et al. (2002), and Angel and Kunkel (2010). Some variants on the more common practice of using general circulation model (GCM) output to impose climate change on the regional hydrologic system were studies that used transposition of climate from other warmer, and sometimes drier, climates, to the Great Lakes region (Croley et al. 1998), and imposition of a somewhat arbitrary range of possible climates with the intention of illuminating the possibility of terminal lakes (i.e. having no outflow) approximately 8000 years ago (Croley and Lewis 2006).

All of these studies feature lake level prominently as an end result of the modeling activity, and generally showed that future climate change results in dropping lake levels, with the possibility that such drops will be of large magnitude. One of the possible projections, a 1.67-m drop during the 21st century in Lake Michigan-Huron, reported in Lofgren et al. (2002), was much quoted in the media long after its publication, regularly without mention that it was published alongside an alternative scenario that projected a rise in lake levels. Croley et al. (1998) showed that if the climate of the region of Oklahoma were to be transposed over the Great Lakes region, Lake Superior would become a terminal lake. The set of results using this method has also been featured in the assessment and synthesis literature (e.g. Great Lakes Regional Assessment Group 2000, Kling et al. 2003, International Joint Commission 2006, and Hayhoe et al. 2010).

Other studies have used different approaches to examine the hydrologic budget of the Great Lakes basin. Some have used direct analysis of the GCM-based regional water budget (precipitation minus evaporation) in the region (Manabe et al. 2004, Milly et al. 2005), while others have examined the convergence of water vapor flux in the atmosphere as indicating net water budget (Kutzbach et al. 2005). Still others have used regional climate modeling to downscale GCM results and project climate both in terms of the atmosphere and the surface (MacKay and Seglenieks 2013, Lofgren 2014, Notaro et al. 2015). Most of these studies did not include lake level explicitly as an output, although MacKay and Seglenieks (2013) and Notaro et al. (2015) do have such projections based on a limited range of GCM inputs. The group of studies that have not been based on the methodology of Croley (1990) have generally not shown large trends toward reduced water budgets in the Great Lakes, and even though not always explicitly shown, this necessarily implies small changes in lake levels. This group of studies has also been much less influential in the mass media and in assessment reports.

Lofgren et al. (2011, hereafter referred to as LHW) attempted to expose some serious problems with the methodology of Croley (1990) for projection of Great Lakes water budgets and levels under climate change scenarios. These specifically dealt with the Large Basin Runoff Model (LBRM), which is the component of the modeling system that deals with the land portion of the Great Lakes basin. The central argument was that the LBRM is overly reliant on near-surface air temperature as a predictor of evapotranspiration (ET) to the exclusion of the surface energy budget as a predictor; although the LBRM is calibrated such that potential evapotranspiration is constrained by incoming solar radiation during the calibration period, this constraint does not apply when considering periods with altered climate regimes. This was broken down into three main intersecting lines of argument:

1. The ET projected for future climate scenarios using input from GCMs was much greater than that directly simulated by the driving GCMs. The equivalent energy of the difference, in terms of latent heat flux, was shown to be around 30 W m^{-2}, a significant discrepancy in the surface energy budget.

2. During periods of observation at land-based flux measurement stations based on the eddy covariance principle, the near-surface air temperature and ET were decomposed into different temporal classes—the annual cycle, variations with a time scale greater than a month but not annual, and variations with a time scale less than a month. For the annual time scale, there was a very high correlation between air temperature and ET, while for the other time scales, there was almost no correlation. This seems to indicate that the seasonal cycle (i.e. solar energy input) is the major driver of both air temperature and ET at the annual scale, while air temperature is a poor predictor of ET variability on other time scales. Therefore, thinking of air temperature as either an actual driver of ET, or a proxy that is universally applicable, is highly problematic.

3. The potential ET (PET), proportional to what is called by Croley (1983) the "energy available for evapotranspiration," is greater by very large factors in the future scenarios than in the historical base case. The largest factor that was found in LHW for a full lake basin in the examples we ran was a 10-fold increase in the PET for all of the land in the Lake Superior Basin. For only the basin of the Montreal River, an eastern tributary of Lake Superior, the corresponding number was a 60-fold increase. These magnitudes of increase can be taken as changes that are equivalent to increases in the incoming sunlight by that factor. These very large increases were due to the formulation of PET as exponential functions of air temperature, with calibrated parameters indicating that the Montreal River Basin has PET increasing by 90% per degree C, and many of the other sub-basins of Lake Superior increasing in the range of 30-50% per

degree C. This contrasts to the point of comparison of the Clausius-Clapeyron relation, which has an increase in saturated water vapor pressure of 7% per degree C, and literature sources such as Held and Soden (2006) and Lorenz and DeWeaver (2007), which show smaller increases in large-scale ET under climate change.

These arguments were the central justifications for asserting that the methods that have been repeatedly used over more than two decades overestimate future ET. In addition to this, LHW gave a small-scale demonstration of what effect these problems might have on projected lake levels. They devised the "energy adjustment" (EA) method, in which the changes in PET from a future climate scenario relative to the historical period was given by the ratio of net surface radiation, rather than the exponential function of the air temperature change, dubbed the "temperature adjustment" (TA) method. This was carried out for two example GCMs taken from the Climate Model Intercomparison Project Phase 3 (CMIP3) set of models. One of these was shown in the main body of LHW, and showed a switch from falling lake levels under the TA method to rising lake levels under the EA method, a difference of about 90 cm on Lake Michigan-Huron. The other example GCM was shown in the supplementary material, and contrasted a large drop in lake levels under the TA method with a lesser drop under the EA method, reducing a drop of 2.1 m on Lake Michigan-Huron to only 1 m.

These results received very guarded recognition in the 2014 National Climate Assessment. One of the authors of the present paper (Lofgren) acquiesced under protest to the inclusion of the following language: "Recent studies have also indicated that earlier approaches to computing evapotranspiration estimates from temperature *may* have overestimated evaporation losses." [italics added] This statement appears both in the Midwest chapter (Pryor et al. 2014) and in the Climate Science Supplement (Walsh et al. 2014).

Some National Climate Assessment authors suggested that what was missing was an assessment of the statistical significance of the difference between the traditional TA method and the more physically-based EA method by using a wider selection of GCM data as input (D. Wuebbles and K. Hayhoe, personal communication). Therefore, that is a major goal of this study. We also introduce two alternative formulations for PET under climate change in addition to the already-used EA method, as additional points of comparison to the TA method. Arguments 1 and 2 from LHW, in the list above, stand as already stated there, but our further use of many more GCMs as input leads to further underline argument 3. Another major application of LBRM is in seasonal-scale projection of Great Lakes water levels within NOAA Great Lakes Environmental Research Laboratory's (GLERL) Advanced Hydrologic Prediction System (AHPS), which is distinct from a product with a similar name developed by the National Weather Service's Office of Hydrology. An assessment of the performance of that entire modeling system was given by Gronewold et al. (2011) and shows that the main disagreement between AHPS and observed water levels is that AHPS underestimates the amount of water level variability. LBRM was also compared to other models in terms of its stream runoff as part of the Great Lakes Runoff Intercomparison Project—Lake Michigan (GRIP-M, Fry et al. 2014). Its results were well in line with those of other models. However, these assessments did not include the systematic and sustained changes in air temperature associated with changes in the climate regime that were investigated in LHW and the present paper.

The rest of this paper is organized as follows. Section 2 will summarize the methodologies that are used here. Section 3 will present the results, including formal analysis of the statistical significance of the discrepancy among these alternative methods. Section 4 will discuss the results and the way that strikingly extreme intermediate results within a model can be disguised in the final results. Conclusions are presented in section 5.

2. METHODS AND EXPERIMENTAL DESIGN

2.1. The Large Basin Runoff Model (LBRM)

Key features of the formulation of the LBRM model for purposes of this study include the determination of "energy available for evapotranspiration" (for convenience referred to here as PET; see further discussion below) and the proportion of PET that is manifested as ET. As stated in Croley (1983), due to considerations of data availability, an *a priori* constraint in developing LBRM was that only daily air temperatures and precipitation be used as input data.

PET is assumed to take the form $Ae^{T/Tb}$, where T_b is a parameter with units of temperature fitted through calibration. During the calibration process, the annual total of PET is taken to be such that its latent heat equivalent is equal to the annual incoming solar radiative energy:

$$LA\Sigma^{year}e^{T/Tb} = \Sigma^{year}S, \qquad (1)$$

where L is the latent heat of evaporation per unit mass, A is a parameter with units of cm/day or other units of evapotranspiration, T is daily mean air temperature, and S is the top-of-atmosphere incident solar radiation. The summation in eq. (1) uses a daily timestep for the snow-free portion of the year. In the calibration process, for each test value of T_b, the value of A is uniquely determined using eq. (1). Thus, eq. (1) constrains the total PET over the year, while T_b controls its distribution throughout the year, and fluctuation with other modes of temperature variability. It should be noted, however, that outside of the calibration process, including in future projections using adjustments to the temperature based on GCM results, the constraint of eq. (1) is not enforced. That is, the PET increases by a factor of $e^{\Delta T/Tb}$ corresponding to the change in temperature between current and future conditions, equivalent to the solar radiation also increasing by the same factor.

To clarify the difference in nomenclature between this paper and Croley (1983), we here use the term potential evapotranspiration to refer to the quantity $Ae^{T/Tb}$, because it is the variable in LBRM that best corresponds to the intuitive and widely-used definition of PET as the amount of ET that would occur when unlimited soil moisture is available. Within the formulation of LBRM, this amount of ET can be approached, but is only reached when soil moisture is infinite. When multiplied by L, this quantity is referred to in Croley (1983) as "energy available for evapotranspiration." Croley uses the term PET to refer to the water equivalent of the portion of the energy available for ET that does *not* actually result in ET.

The full set of calibrated parameters for each sub-basin in LBRM is listed in Table 1. Although the actual calibration that led to this set of values is not well documented, the values are generally similar to those of Croley (1983).

2.2 Experimental design

The community's general understanding of how evapotranspiration occurs concentrates on the surface as its locus of evapotranspiration, as well as the surface's energetics and gradients of temperature and specific humidity across the planetary boundary layer as drivers—liquid water in water bodies or soil or water within plant leaves acquires enough kinetic energy to change to the gaseous state, a process promoted by absorption of energy (heat) at the surface and a gradient of water vapor concentration into the planetary boundary layer. These principles form the paradigm that stands behind diagnostic equations of potential evapotranspiration, such as those of Penman and Monteith (Monteith 1973) and Priestley and Taylor (1972). They are also the basis for the surface transfer schemes embedded within climate models, from the early generation of bucket models (e.g. Manabe and Holloway 1975) to more complex soil-vegetation-atmosphere transfer schemes (Dickinson et al. 1986, Sellers et al. 1986) and their newer descendants.

Even most sources of information directed at the general public refrain from directly saying that the increased capacity for water vapor in warmer air *causes* increased evapotranspiration, but often do mention these concepts in close proximity to each other, e.g. Brean (2015). Therefore, there can tend to be a simplifying assumption that potential ET will increase with temperature in proportion to the Clausius-Clapeyron relationship. An analysis related to this concept is found in d'Orgeville et al. (2014), in which they compare regional precipitation changes associated with climate change simulations to the "expected" increases, assuming proportionality to the Clausius-Clapeyron relation. They did, in fact, find that in the Great Lakes Basin, their regional model showed an increase in precipitation that agreed well with this assumption, approximating 7% per degree C. The explanation in http://water.epa.gov/scitech/climatechange/Water-Impacts-of-Climate-Change.cfm (accessed May 12, 2015) actually seems to reverse the more popularly assumed cause-effect relationship: "Warmer temperatures increase the rate of evaporation of water into the atmosphere, in effect increasing the atmosphere's capacity to 'hold' water." So some confusion occurs surrounding the relationship between ET and water capacity of the atmosphere.

Table 1. LBRM parameter values for each sub-basin of all lakes. For purposes of these parameters, Georgian Bay is treated as separate from the rest of Lake Huron. T_b is the parameter used to calculate PET from temperature as in eq. (1). Parameter 2 is a coefficient determining snowmelt rate. Parameter 3 gives the rate at which water from the upper soil zone percolates into the lower soil zone. Parameter 4 controls the proportion of PET that becomes actual ET from the upper soil zone. Parameter 5 is a coefficient for the rate of interflow from the lower soil zone to the surface storage. Parameter 6 is a coefficient for the rate of deep percolation from the lower soil zone to the groundwater reservoir. Parameter 7 controls the proportion of PET that becomes actual ET from the lower soil zone. Parameter 8 is a coefficient for the rate of transfer from the groundwater to the surface storage. Parameter is a coefficient for the rate of outflow from the surface storage to the lake. For details, see Croley (1983).

Lake	Basin	T_b	2	3	4	5	6	7	8	9
Superior	1	3.00E+00	2.27E-01	3.00E-01	5.90E+01	2.10E-10	6.00E-03	1.00E-10	3.50E-02	8.60E-02
Superior	2	7.20E+00	1.90E-01	1.20E+02	1.00E-06	1.00E-01	7.60E-02	7.90E-09	3.30E-02	2.30E+01
Superior	3	4.70E+00	3.21E-01	2.30E+00	1.90E-08	3.90E-02	3.00E-02	3.20E+03	3.00E-02	2.20E-01
Superior	4	6.00E+00	2.64E-01	7.10E-01	8.50E-08	1.10E-10	1.80E-02	3.00E-09	6.70E-01	2.20E-01
Superior	5	6.30E+00	2.57E-01	2.30E+01	1.20E-10	4.00E+00	6.30E+00	2.60E+00	1.40E-02	2.60E-01
Superior	6	4.50E+00	2.98E-01	5.40E-01	4.80E-08	2.10E-10	2.70E-01	9.50E-08	5.80E-02	4.30E-01
Superior	7	5.20E+00	2.13E-01	1.00E+00	7.00E-08	1.10E-02	2.40E-03	3.50E-09	8.50E-03	1.70E-01
Superior	8	9.20E+08	4.10E+01	1.40E+01	4.50E-08	1.00E-10	9.90E-03	7.80E-10	1.40E-02	1.30E-02
Superior	9	2.30E+00	2.25E-01	4.60E-01	3.50E+03	1.40E-03	8.20E-03	4.00E-06	8.90E-04	7.50E-02
Superior	10	2.30E+00	2.25E-01	4.60E-01	3.50E+03	1.40E-03	8.20E-03	4.00E-06	8.90E-04	7.50E-02
Superior	11	2.30E+00	2.20E-01	4.60E-01	3.50E+03	1.40E-03	8.20E-03	4.00E-06	8.90E-04	7.50E-02
Superior	12	5.10E+00	2.57E-01	6.80E-01	2.20E+03	1.10E+03	7.90E+03	9.00E-01	1.40E-02	1.90E-01
Superior	13	1.10E+00	1.44E-01	9.30E+00	9.10E+03	1.10E-03	9.20E-03	6.90E-05	1.00E-03	9.50E-02
Superior	14	3.90E+00	2.44E-01	4.70E+00	9.10E-09	2.70E-03	4.30E-03	2.00E-10	1.70E-03	8.40E-02
Superior	15	9.70E+00	2.53E-01	4.70E+01	5.30E-07	2.00E-02	7.90E-03	3.90E-10	5.30E-01	1.10E-01
Superior	16	5.80E+00	2.33E-01	5.40E-01	8.70E-09	1.00E-10	2.70E-02	3.30E-09	6.00E-02	1.00E-01
Superior	17	1.80E+00	2.10E-01	5.50E-01	4.40E-02	3.20E-03	2.00E-03	9.80E-03	5.90E-03	7.10E-02
Superior	18	1.40E+00	2.01E-01	3.40E-01	6.40E-04	9.90E-04	1.00E-03	9.60E-02	9.30E+00	8.10E-02
Superior	19	2.00E+00	7.13E-02	9.10E+08	2.00E-10	7.90E-03	8.10E-02	2.70E-09	5.30E-03	5.20E-03
Superior	20	2.10E+00	1.88E-01	6.10E-01	1.30E-07	2.20E-03	8.30E-04	5.70E-08	9.90E+08	4.20E-02
Superior	21	2.00E+00	3.06E-01	5.70E+00	7.60E+00	4.20E-03	6.80E-03	4.60E-02	6.70E-03	1.10E-01
Superior	22	3.10E+00	1.58E-01	5.40E-01	6.90E-02	1.10E-10	1.80E-01	6.30E-01	2.60E-02	1.80E-01
Michigan	1	4.50E+00	2.44E-01	2.50E-01	1.60E-07	1.10E-01	2.50E-01	1.00E-10	5.00E-03	5.80E-02
Michigan	2	4.50E+00	2.45E-01	2.50E-01	1.60E-07	1.10E-01	2.50E-01	1.00E-10	5.00E-03	5.80E-02
Michigan	3	7.10E+00	2.09E-01	1.00E+01	9.90E-07	4.70E-02	2.50E-02	8.60E-09	1.20E-02	6.90E-01
Michigan	4	7.10E+00	2.08E-01	1.00E+01	9.90E-07	4.70E-02	2.50E-02	8.60E-09	1.20E-02	6.90E-01
Michigan	5	5.20E+00	2.07E-01	7.20E-01	1.20E-07	9.20E-03	9.00E-08	2.60E-09	3.30E-02	1.10E-01
Michigan	6	5.10E+00	1.90E-01	2.90E-01	1.10E-07	1.00E-05	6.80E-02	8.30E-02	5.30E-02	9.60E-02
Michigan	7	3.90E+00	2.30E-01	3.20E+00	1.10E-08	7.40E-03	4.60E-03	1.50E-09	9.80E-04	9.60E-02
Michigan	8	4.70E+00	2.54E-01	1.70E+00	1.20E-07	8.00E-03	6.20E-03	6.30E-09	5.20E-03	1.20E-01
Michigan	9	5.70E+00	2.66E-01	3.50E+00	2.90E-07	8.90E-03	1.00E-02	5.10E-09	3.40E-03	1.20E-01
Michigan	10	6.00E+00	2.67E-01	1.10E-01	1.80E+01	3.80E-06	4.00E-06	1.00E+01	6.40E-02	2.20E-01
Michigan	11	9.40E+00	7.32E+00	2.70E+02	4.30E-07	1.00E-02	9.10E-08	2.50E-10	4.00E-01	2.10E-01
Michigan	12	5.60E+00	3.82E-01	1.30E+00	2.00E+01	1.80E-02	9.50E-07	7.50E-01	2.50E-06	4.50E-01
Michigan	13	5.90E+00	3.98E-01	1.60E+00	1.90E-01	1.90E-02	8.20E-06	9.90E-03	9.80E-06	1.50E-01
Michigan	14	5.70E+00	3.15E-01	4.00E+00	3.20E-06	2.80E-02	3.40E-02	2.10E-07	1.00E-02	3.80E-01
Michigan	15	8.10E+00	3.22E-01	4.80E-01	3.90E-07	8.00E-06	6.80E+05	3.80E-01	2.20E-02	3.00E-01

Table 1. (Continued)

Lake	Basin	T_b	2	3	4	5	6	7	8	9
Michigan	16	8.30E+00	4.38E-01	1.70E+01	4.70E-08	8.60E-02	2.40E-01	3.90E-08	1.40E-02	1.30E-01
Michigan	17	5.60E+00	1.08E+01	8.40E+05	1.70E-07	4.20E-02	5.00E-02	1.30E-07	1.70E-02	2.60E+01
Michigan	18	8.50E+00	4.96E-01	8.90E+00	3.40E-07	1.10E-02	1.10E-02	1.30E-09	4.30E-03	1.80E-01
Michigan	19	5.90E+00	3.54E-01	1.10E-05	5.70E-03	9.30E+00	6.60E-01	1.10E-04	1.20E-05	5.80E-01
Michigan	20	5.90E+00	3.02E-01	2.00E-01	2.50E-07	1.00E-03	1.90E-02	9.00E-09	2.10E-02	5.70E-02
Michigan	21	4.80E+00	2.92E-01	5.00E+00	4.90E-07	4.40E-03	6.40E-03	5.10E-10	1.70E-04	1.50E-01
Michigan	22	6.00E+00	3.33E-01	3.80E+00	1.60E-07	5.70E-03	4.10E-03	3.90E-10	1.20E-03	1.30E-01
Michigan	23	4.80E+00	2.94E-01	5.00E+00	4.90E-07	4.40E-03	6.40E-03	5.10E-10	1.70E-04	1.50E-01
Michigan	24	7.50E+00	3.94E+01	1.80E+03	6.20E-05	6.00E-03	7.50E-03	2.40E-10	2.60E-04	7.60E+00
Michigan	25	4.80E+00	2.93E-01	5.00E+00	4.90E-07	4.40E-03	6.40E-03	5.10E-10	1.70E-04	1.50E-01
Michigan	26	6.40E+00	1.49E-01	6.00E+01	2.60E-06	2.20E-02	1.20E-01	3.20E-09	2.30E-04	5.10E+01
Michigan	27	6.40E+00	1.47E-01	6.00E+01	2.60E-06	2.20E-02	1.20E-01	3.20E-09	2.30E-04	5.10E+01
Huron	1	4.90E+00	2.44E-01	1.60E-01	2.90E-07	3.10E-02	8.80E-06	5.50E-10	7.50E-02	2.30E-01
Huron	2	6.60E+00	3.29E-01	1.40E+01	2.30E-07	6.50E-03	8.20E-03	1.00E-09	7.40E-04	1.40E-01
Huron	3	6.60E+00	3.30E-01	1.40E+01	2.30E-07	6.50E-03	8.20E-03	1.00E-09	7.40E-04	1.40E-01
Huron	4	6.60E+00	3.28E-01	1.40E+01	2.30E-07	6.50E-03	8.20E-03	1.00E-09	7.40E-04	1.40E-01
Huron	5	5.30E+00	2.97E-01	4.70E+00	3.20E-07	8.90E-03	8.20E-03	9.10E-09	4.20E-03	2.30E-01
Huron	6	5.30E+00	3.03E-01	4.70E+00	3.20E-07	8.90E-03	8.20E-03	9.10E-09	4.20E-03	2.30E-01
Huron	7	6.20E+00	2.13E-01	1.00E+01	2.00E-07	4.60E-03	1.00E-02	4.90E-10	9.60E-04	2.70E-01
Huron	8	5.60E+00	2.97E-01	3.00E+00	4.10E-07	1.00E-02	5.50E-07	8.30E-10	5.20E-04	3.60E-01
Huron	9	5.70E+00	2.67E-01	9.10E-07	6.60E-04	7.10E-01	6.20E-03	7.00E-10	3.00E-02	9.70E-02
Huron	10	5.40E+00	2.86E-01	2.00E-01	3.20E-07	9.70E-04	2.00E-02	5.50E-08	2.60E-02	1.10E-01
Huron	11	5.70E+00	2.94E-01	8.00E-07	2.70E-05	3.00E-01	6.00E-03	7.00E-10	4.10E-02	2.10E-01
Huron	12	6.30E+00	5.63E-01	1.00E-01	4.70E-03	1.70E-02	9.90E-07	2.60E-09	2.50E-03	3.00E-01
Huron	13	6.30E+00	5.63E-01	1.00E-01	4.70E-03	1.70E-02	9.90E-07	2.60E-09	2.50E-03	3.00E-01
Huron	14	4.80E+00	7.40E-01	4.20E-01	9.30E-06	1.50E-02	8.60E-07	1.50E-06	3.50E-04	2.30E-01
Huron	15	6.60E+00	9.20E-01	1.60E+00	1.30E-02	1.90E-02	9.50E-07	7.60E-03	8.00E-02	7.30E-01
Huron	16	5.30E+00	7.51E-01	1.60E+00	2.20E-07	1.20E-02	5.30E-03	5.00E-09	1.10E-02	1.70E-01
Georgian	1	4.30E+00	7.66E-01	1.10E-01	1.80E-04	2.10E-02	8.30E-07	1.10E-05	8.70E-04	8.10E-02
Georgian	2	6.10E+00	7.64E-01	3.70E+00	2.00E-07	1.50E-02	8.90E-07	4.30E-09	2.00E-03	2.40E-01
Georgian	3	6.00E+00	4.78E-01	3.30E+00	4.10E-07	1.00E-02	7.70E-03	9.40E-09	7.00E-03	3.50E-01
Georgian	4	4.70E+00	3.78E-01	8.40E+00	4.60E-07	7.10E-03	5.70E-03	1.40E-08	1.30E-04	1.30E-01
Georgian	5	4.80E+00	7.30E-01	4.30E+03	8.00E-06	1.60E-02	7.40E-03	2.70E-09	3.10E-04	3.40E-01
Georgian	6	3.90E+00	4.25E-01	5.20E+01	5.90E-05	1.20E-01	5.30E-02	5.90E-07	3.80E-06	7.90E-02
Georgian	7	2.60E+00	1.02E+00	5.70E+05	8.20E-10	7.50E-03	5.80E-03	5.90E-10	5.00E-06	7.90E-02
Georgian	8	3.70E+00	1.24E+00	2.10E+01	2.20E-08	4.60E-03	1.80E-03	8.10E-10	8.50E-04	1.70E-01
Georgian	9	4.40E+00	2.40E-01	1.20E+00	7.50E-09	6.20E-03	3.50E-04	2.40E-10	1.90E-02	7.50E-02
Georgian	10	4.40E+00	3.68E-01	7.40E-01	3.40E-07	2.60E-02	2.70E-06	2.20E-09	4.60E-02	5.40E-02
Georgian	11	2.20E+00	2.32E-01	3.00E+00	7.00E-08	4.80E-03	2.20E-03	5.00E-10	2.60E-04	9.90E-02
Georgian	12	4.90E+00	2.46E-01	1.60E-01	2.90E-07	3.10E-02	8.80E-06	5.50E-10	7.50E-02	2.30E-01
Georgian	13	4.90E+00	2.46E-01	1.60E-01	2.90E-07	3.10E-02	8.80E-06	5.50E-10	7.50E-02	2.30E-01

Table 1. (Continued)

Lake	Basin	T_b	2	3	4	5	6	7	8	9
St. Clair	1	6.40E+00	3.53E-01	9.70E-02	1.10E+00	9.00E-07	1.00E-02	9.80E-01	4.20E-02	2.50E-01
St. Clair	2	8.20E+00	1.66E-01	1.60E+01	4.20E-06	1.00E-01	1.60E-01	1.10E-07	1.50E-02	3.10E+00
St. Clair	3	8.20E+00	2.76E-01	1.60E+01	4.20E-06	1.00E-01	1.60E-01	1.10E-07	1.50E-02	3.10E+00
St. Clair	4	8.20E+00	1.66E-01	1.60E+01	4.20E-06	1.00E-01	1.60E-01	1.10E-07	1.50E-02	3.10E+00
St. Clair	5	8.50E+00	7.46E-01	1.10E-01	5.30E-06	5.10E-02	6.00E-06	8.60E-08	5.10E-02	6.50E-01
St. Clair	6	7.80E+00	4.17E-01	4.10E+04	7.20E+01	3.10E-01	1.10E-01	3.40E-08	2.10E-02	2.80E-01
St. Clair	7	6.40E+00	5.03E-01	4.50E-02	3.40E-05	9.70E-03	1.10E-06	5.10E-09	5.90E-02	1.80E-01
Erie	1	8.00E+00	2.76E-01	1.30E+00	6.70E-07	2.40E-02	5.30E-02	1.40E-07	1.40E-02	1.10E+00
Erie	2	1.00E+01	3.65E-01	3.00E+01	5.70E-06	4.50E-02	1.20E-01	5.90E-08	1.70E-02	4.00E-01
Erie	3	6.20E+00	3.18E-01	7.30E-01	3.70E-03	1.00E-05	1.70E-02	6.50E-04	2.70E-02	3.60E-01
Erie	4	7.90E+00	3.98E-01	8.80E-06	9.50E-03	9.00E-02	7.90E-02	8.80E-05	1.00E-01	1.00E-01
Erie	5	8.10E+00	4.14E-01	5.80E-05	1.10E-06	6.00E+04	9.90E-02	1.00E-04	2.00E-01	3.30E-01
Erie	6	6.60E+00	4.40E-01	3.90E-02	4.50E-08	9.70E-06	3.20E-06	6.40E-08	4.90E-02	1.90E-01
Erie	7	6.40E+00	3.63E-01	4.30E-02	4.70E-07	8.60E-06	6.10E-06	1.20E-06	6.40E-02	4.10E-01
Erie	8	5.90E+00	8.90E-01	4.90E-02	1.20E-06	6.00E-03	2.50E-06	1.60E-06	6.00E-02	2.80E-01
Erie	9	7.10E+00	6.17E-01	9.20E-07	4.90E-06	1.00E-05	3.00E-06	5.90E-08	5.00E-02	1.20E+00
Erie	10	5.20E+00	3.18E-01	1.10E-01	2.90E-07	1.20E-02	6.30E-06	2.30E-07	5.90E-02	6.60E-01
Erie	11	7.60E+00	2.75E-01	7.00E+00	8.40E-07	1.10E-01	5.70E-02	1.60E-08	2.20E-02	3.20E+00
Erie	12	5.80E+00	3.95E-01	1.10E+00	1.20E-06	6.50E-02	5.80E-06	1.90E-08	2.90E-02	3.40E+00
Erie	13	5.10E+00	3.29E-01	6.40E-07	1.30E-06	1.00E-05	3.00E-06	5.90E-08	5.00E-02	3.30E-01
Erie	14	4.50E+00	7.39E+00	6.40E-02	2.30E-06	8.50E-02	4.40E-06	1.90E-04	6.10E-02	6.80E-01
Erie	15	4.40E+00	3.49E-01	3.20E-02	4.90E-07	3.90E-02	5.40E-06	1.00E-10	5.80E-02	6.60E-01
Erie	16	4.30E+00	2.45E-01	1.70E+00	1.10E-06	1.50E-01	1.40E-01	5.30E-08	2.40E-02	6.80E+00
Erie	17	4.60E+00	4.27E-01	1.60E+00	1.40E-05	1.70E-01	2.00E-01	1.50E-06	2.70E-02	6.60E+00
Erie	18	4.60E+00	4.29E-01	1.60E+00	1.40E-05	1.70E-01	2.00E-01	1.50E-06	2.70E-02	6.60E+00
Erie	19	9.40E+00	4.63E-01	2.70E+01	8.00E-03	3.30E-01	4.00E-01	3.80E-08	1.80E-02	3.50E-01
Erie	20	6.90E+00	5.97E-01	2.60E+00	2.80E-06	5.70E-02	7.20E-02	1.60E-07	1.30E-02	5.10E-01
Erie	21	1.40E+01	6.96E-01	9.10E+01	1.10E-04	3.50E-01	4.40E-06	5.40E-08	7.50E-02	1.00E+00
Ontario	1	6.00E+00	3.68E-01	5.10E-02	4.40E-05	6.10E-03	3.30E-06	1.50E-04	2.40E-01	3.30E-01
Ontario	2	6.00E+00	3.68E-01	5.10E-02	4.40E-05	6.10E-03	3.30E-06	1.50E-04	2.40E-01	3.30E-01
Ontario	3	4.20E+00	2.49E-01	9.00E-05	5.60E-08	2.00E-02	7.30E-06	1.40E-10	6.70E-02	6.30E-02
Ontario	4	5.50E+00	2.47E-01	4.90E-01	3.30E-06	1.70E-06	1.90E-01	3.60E-07	2.20E-02	4.00E-01
Ontario	5	4.30E+00	3.81E-01	1.20E-01	1.90E-08	4.10E-02	7.40E-06	9.60E-08	2.90E-01	4.50E-02
Ontario	6	4.30E+00	2.81E-01	1.80E+00	3.30E-06	1.20E-01	8.70E-02	3.90E-08	4.20E-02	3.60E+00
Ontario	7	4.40E+00	3.81E-01	1.00E-06	1.10E-08	2.00E-05	3.00E-06	8.10E-08	5.00E-02	5.90E-02
Ontario	8	3.10E+00	3.39E-01	1.10E-01	1.00E-07	2.10E-02	1.40E-06	1.00E-10	3.60E-02	9.40E-02
Ontario	9	4.40E+00	5.94E-01	8.30E-01	5.50E-02	8.90E-07	1.50E-02	2.00E-03	7.90E+01	1.00E-01
Ontario	10	4.40E+00	5.75E-01	8.30E-01	5.50E-02	8.90E-07	1.50E-02	2.00E-03	7.90E+01	1.00E-01
Ontario	11	6.90E+00	4.88E-01	4.90E+00	5.90E-02	1.50E-01	1.20E-01	2.70E-07	2.20E-02	1.00E-01
Ontario	12	5.60E+00	6.09E-01	5.70E+05	1.10E-10	1.70E-02	4.90E-03	1.90E-09	9.50E-03	5.30E-01
Ontario	13	5.50E+00	3.47E-01	1.90E+00	2.30E-05	9.50E-03	2.40E-02	2.50E-08	7.10E-03	1.50E+00
Ontario	14	5.50E+00	3.19E-01	1.90E+00	3.10E-06	2.00E-02	2.80E-02	2.30E-07	1.20E-02	4.80E-01
Ontario	15	5.50E+00	3.17E-01	1.90E+00	3.10E-06	2.00E-02	2.80E-02	2.30E-07	1.20E-02	4.80E-01

The general framework of previous climate change sensitivity experiments using LBRM and the other associated models under the method of Croley (1990) was to have one run, the base case, which simply used station-based meteorological observations as drivers, for the period 1948-2005 in our case. Then to represent future time periods, the same historical meteorology was used as the foundation of the inputs to LBRM, but was adjusted to values intended to represent the climate change, based on results from GCMs. Croley (1990) notes that this "change factor" method was used because of instructions from the U.S. Environmental Protection Agency, the funding agency for that study. Along with the traditional method of treating climate change in the LBRM (here called the temperature adjustment method or "TA"), based on its embedded assumptions about the drivers of ET, we will be using three additional methods of applying GCM output to the LBRM for evaluation of response to climate scenarios (summarized in Table 2).

Table 2. Description of alternative methods of incorporating GCM-based climate projections into the LBRM.

Name	Abbreviation	Description	Temperature sensitivity of PET
Temperature adjustment	TA	Air temperatures adjusted by GCM climate projections are inserted directly into LBRM's equations for PET	7%-90% per °C depending on sub-basin, most between 15% and 50%
Energy adjustment	EA	PET values from LBRM in the base case climate are multiplied by the ratio of net radiation at the surface in the future case vs. the base case	No explicit dependence on temperature, only energy budget
Priestley-Taylor	PT	The energy adjustment method is applied first, then PET is further multiplied by the temperature-dependent factor from the Priestley-Taylor equation (eq. 5)	Explicit temperature dependence of ranging between 1.2% and 4.0% per °C, plus dependence on energy budget
Clausius-Clapeyron	CC	PET from the base case is multiplied by a factor representing the change in water capacity associated with the temperature change from the GCM	7% per °C

1. The energy adjustment method ("EA"), as in LHW, uses the sum of the sensible heat flux and latent heat flux given by the GCM as an indication of the net radiative energy flux at the surface. The ratio of the GCM's value in the future time period to the historical time period is used as a multiplier for the LBRM's PET from the base case.

2. The EA method of LHW got some legitimate criticism from Scheff and Frierson (2014) for ignoring the term that is directly driven by air temperature in formulations such as that of Priestley and Taylor (1972). The Priestley-Taylor formulation is similar to that of Monteith (1973), but uses an empirical constant multiplier in place of an additional term for water vapor deficit. We adopt the concept of Priestley-Taylor here and label these runs "PT", where in addition to multiplying PET by the ratio of surface radiative energy as in the EA method, we also multiply it by a factor directly driven by the air temperature change. This latter factor, however, ranges between 1.2% per °C (at 30°C) and 4.0% per °C (at 0° C), far less sensitive to temperature than the LBRM under the TA method.

3. The causal relationship between ET and atmospheric water holding capacity is not direct and has been a source of confusion, but the Clausius-Clapeyron relation seems to set a maximum bound on sensitivity of ET to climate change. So we use it as an additional point of comparison, multiplying the PET from the base case by the change in water vapor holding capacity of the atmosphere (expressed as a ratio) to project future PET. Even though this method also appears to overestimate the sensitivity of PET to climate change (based on e.g. Held and Soden 2006 and Lorenz and DeWeaver 2007), we regard it as "less wrong" than the TA method, and therefore find it useful as another point of comparison to demonstrate the problems that exist in the TA method.

We approximate the Clausius-Clapeyron relationship in an exponential form, in order to create an equation based on the Priestley-Taylor equation in closed form. The approximation that we use for saturation vapor pressure is

$$e_s = .632\ kPa\ e^{T/15.51°C} \tag{2}$$

This formula is fitted so that it gives the exact values from the table in Lide (2005) at 5°C and 25°C. In the range of 0-30°C, this simplified formula lies within 3.4% of the actual value. The Priestley-Taylor equation is

$$PET = \frac{\Delta(R_n - G)}{L(\Delta + \gamma)} \tag{3}$$

where Δ is the gradient of e_s with temperature (0.0407 kPa $°C^{-1}$ $e^{T/15.51°C}$), R_n is net radiative flux at the surface, *G* is conductive heat flux into the ground, *L* is latent heat of vaporization per unit mass, and γ is the psychrometic constant, approximately 0.067 kPa $°C^{-1}$. Filling in the exponential formulas (and dropping units of kPa $°C^{-1}$, since they cancel out), this transforms to

$$PET = \frac{0.0407\ e^{T/15.51°C}(R_n - G)}{L(0.0407\ e^{T/15.51°C} + 0.067)} \tag{4}$$

To simplify calculations when we are using two sets of numbers—the actual observations and the temperature offset from climate change based on GCM simulations—we will use the first-order Taylor expansion of this; i.e. we multiply the change in temperature by the derivative of eq, (4) with respect to *T* (assuming that *L* is constant with temperature), and then divide by eq. (4), in order to get a fractional change in PET due to temperature change. This factor is the multiplier that is applied to the PET from the EA method to get the PET from the PT method. Its value is

$$\frac{\partial PET/\partial T}{PET} = \frac{4.32 \times 10^{-3}}{0.0407\ e^{T/15.51°C} + 0.067} \tag{5}$$

2.3 Other Model Components

In addition to the runoff from the land in the lakes' basins, the other components of net basin supply are the over-lake precipitation and evaporation. As in Angel and Kunkel (2010) and preceding studies, over-lake precipitation was taken from Thiessen polygon weighting of data from land-based stations for the base case, and adjusted by ratios of GCM precipitation for the future cases. Also, as in previous studies, evaporation from the lakes was calculated using the Large Lake Thermodynamics Model (LLTM, Croley 1989), which projects water temperatures of the Great Lakes by treating them as lumped 1-dimensional lakes. It estimates solar and longwave radiation exchanges at the surface, and evaporation (latent heat flux) and sensible heat flux based on bulk aerodynamic formulas. Lake temperature profiles are then calculated using this energy budget and formulas for vertical mixing.

Since the over-lake precipitation and evaporation are not affected by the method of adjusting the LBRM for climate change, they have only one value for each GCM realization. They are combined with the LBRM results for each of the methods (TA, EA, PT, and CC) to get net basin supply (NBS) values that correspond to each GCM realization and each method—runoff from LBRM, plus over-lake precipitation, minus over-lake evaporation.

These net basin supplies are then the input to the Coordinated Great Lakes Regulation and Routing Model (CGLRRM), developed by the U.S. Army Corps of Engineers and Environment Canada (T. Hunter, personal communication). Each set of net basin supply time series was input to the CGLRRM to derive lake levels for Lake Superior, the combined Lake Michigan-Huron, Lake St. Clair (not shown), and Lake Erie. Because of the frequent departure of actual operation of the dams regulating Lake Ontario from their official plans, CGLRRM does not calculate levels for Lake Ontario.
The values for all variables generated by the GCMs are spatially interpolated to the locations relevant for simulation by LBRM, LLTM, and the over-lake precipitation estimate by simple inverse-distance weighting from the nearest four

GCM grid points, without consideration for such factors as elevation or location relative to observing stations with special weather characteristics (e.g. lake breeze or lake-effect precipitation). We understand that this is not a state-of-the-art way of doing downscaling of GCM output, but we want to replicate the methodology of Angel and Kunkel (2010) and predecessor studies, in order to isolate the effects of the portions of the formulation that we expect to be even more problematic than the details of downscaling.

2.4 General Circulation Models (GCMs)

A summary of the GCM runs that were used is in Table 3. This group of model realizations was chosen based on the availability of the data needed for simulations using our modeling system. Our criterion for choosing the first five models listed was that we used the models from the Climate Model Intercomparison Study Phase 5 (CMIP5) for which the following variables were available on a monthly basis from the Earth System Grid Federation server at Lawrence Livermore National Laboratory for a historical period including 1986-2005 in a single file for each variable, and for future periods including 2056-2100 in a single file for each variable: air temperature at 2 m height, humidity at 2 m height, precipitation, cloud cover, sensible heat flux, latent heat flux, and wind at 10 m height (either as a scalar quantity or vector components). We also accessed the Earth System Grid Federation server at the Geophysical Fluid Dynamics Laboratory (GFDL) to access the models generated by GFDL; these output files were in 5-year increments, requiring additional data handling. No other criteria were used to select GCMs, such as their accuracy in historically simulating the climate of the region. The monthly data from these GCMs were spatially interpolated to the lakes and sub-basins as in previous studies using the method of Croley (1990).

For each of the GCM realizations, we calculated the climatological values for each month of the year from the period 2056-2075 (referred to as the mid-21st century) and 2081-2100 (referred to as the late 21st century) for comparison to the historical period 1986-2005. For the historical period, we used only the lowest-numbered model realization that was available on the data server. So with the 32 model realizations shown in Table 3 and the two time periods, we had 64 scenarios to use as input to the hydrologic model suite.

Table 3. Summary of GCM runs used. In the columns under the Representative Concentration Pathways (RCP), the number of model realizations using that RCP is indicated. All of these cases were used with the lowest-numbered model run from the historical time period.

Institution and country	Model name and reference	RCP 4.5	RCP 6.0	RCP 8.5
National Center for Atmospheric Research, USA	CESM1-CAM5, Hurrell et al. 2013	3	3	3
Centre National de Recherches Météorologiques, France	CM5, Voldoire et al. 2013	1	0	1
Commonwealth Scientific and Industrial Research Organization, Australia	Mk3-6-0, Rotstayn et al. 2012	2	2	2
Hadley Centre for Climate Science and Services, United Kingdom	HadGEM2-AO, Collins et al. 2011, Martin et al. 2011	1	1	0
Institut Pierre Simon Laplace, France	CM5A-LR, Dufresne et al. 2012	2	1	1
Geophysical Fluid Dynamics Laboratory, USA	CM3.0, Delworth et al. 2006, Donner et al. 2011	1	1	1
Geophysical Fluid Dynamics Laboratory, USA	ESM2G, Dunne et al. 2012, Dunne et al. 2013	1	1	1
Geophysical Fluid Dynamics Laboratory, USA	ESM2M, Dunne et al. 2012, Dunne et al. 2013	1	1	1

3. RESULTS

All box plots shown here use the default algorithm for creating box plots using the R software package. The heavy horizontal line inside the box indicates the median value of all of the data, the upper boundary of the box indicates the median of the subset of values at or above the overall median (approximating the 75% quantile), and the bottom of the box indicates the median of the values at or below the overall median (approximating the 25% quantile). The distance between the 25% and 75% quantile is known as the inter-quartile distance (IQD), and the maximum length of the whiskers is 1.5 times the IQD from the boundary of the box. The whiskers actually extend to the data value that is farthest from the box while still being within this maximum distance. Additional data points outside of this maximum length of the whiskers ("outliers") are indicated by individual circles. Each individual data point used in creating the box plots represents one GCM realization. While this algorithm for creating box and whiskers plots may not be appropriate for non-Gaussian datasets, one point of this study is to highlight the highly irregular and non-Gaussian nature of the data we are dealing with.

3.1. Lake Superior basin

Lake Superior is the lake whose outflow eventually proceeds to all of the other Great Lakes. Hence its basin's water budget can significantly influence the lake levels of all of the lakes. As an outcome of the vagaries of the calibration process of LBRM, the calibrated values of T_b for the sub-basins of Lake Superior are generally smaller than those for the other lakes, with many falling in the range of 2-5° C (contrasting with the Clausius-Clapeyron equivalent value of 15.51° C), and the most extreme low value being 1.1° C in the basin of the Montreal River at the eastern end of Lake Superior.

The first results that we show are the most extreme—the results for changes in PET in the Lake Superior basin. These are shown as a ratio of the value in the projected time period to the base case in Figure 1a. Because the figure's scale is so large to accommodate the changes that result from the TA method, the results for the CC, PM, and EA cases are almost indistinguishable; they cluster slightly above the value of unity, with just a few values below unity. The largest value is 1.72 in the case of the CC method for the RCP 8.5 run of the GFDL CM3.0 for the late 21^{st} century.

In sharp contrast to the CC, PM, and EA methods' values for PET are the values under the TA method in the Lake Superior basin. The median value is 5.05, with several values far more extreme than this. The highest value that is shown in the figure is 565.4, again for the RCP8.5 run of the GFDL CM3.0 for the late 21^{st} century. Although the median increase by a factor of 5.05 may seem small compared to the extreme value of 565.4, it is in fact shockingly large when one considers that it implies the equivalent of an increase in the incoming solar radiation by a factor of 5.05.

The actual ET is constrained not only by the PET, but also by the amount of moisture that is available—precipitation and its storage as soil moisture. Therefore the results for ET among the different methods scale reasonably to be displayed on the same graph, but still differ markedly. The differences in ET between the base case and climate change cases under the various methods are shown in Figure 1b. The median change in ET in the Lake Superior basin under the TA method is 0.024 cm/day, a 20.9% increase over the base case value of 0.115 cm/day. Under the PT method, ET is increased by only 0.004 cm/day. The fact that this increase is much less than the factor of increase in PET is reflective of the fact that both potential and actual ET are small during the winter, when ET is energy-limited, while the summer season's ET tends to be more moisture-limited, especially when PET is increased. Thus the proportional increase in actual ET is much less than in PET. Nevertheless, the median value of ET increase under the TA method is greater than the top extent of the whiskers for all three of the other methods. Also, the CC, PM, and EA methods' results are much more similar to each other than to the TA method. The difference of 0.020 cm/day between TA and PT is the equivalent of 5.8 W m^{-2}—not a really large difference in the surface energy budget, but we need to remember that this is averaged over the entire year and is limited by the availability of moisture, and that the total present direct forcing by human-made greenhouse gases is around 1 W m^{-2}.

In terms of runoff (Figure 1c), the results are influenced by those for actual ET, but also reflect changes in precipitation. Because there are some GCM runs that project large increases in precipitation, but few project significant decreases, the whiskers extending downward are relatively short, especially for the CC, PT, and EA cases, with no outliers on the low side. The entire box for the TA method is below the whiskers for the other three cases, although the whisker at the high end for the TA method extends into the range of the other boxes. This means that the GCM cases with larger increases in precipitation and smaller increases in temperature were able to have increased runoff, even under the TA method, although most cases have decreased runoff. Under the other methods, though, nearly all cases have increased runoff.

The changes in runoff propagate directly into changes in NBS (Figure 1d), because the same results for the other components of net basin supply (over-lake precipitation and evaporation) were used for all of the methods. Therefore, the discrepancies among methods are the same as for runoff.

The lake level of Lake Superior (Figure 1e) is less sensitive to net basin supply than those of Lakes Michigan-Huron and Erie. This is in part because there is no other Great Lake upstream to enhance its net basin supply with an altered inflow, but more importantly because of the regulation of the outflow into the St. Marys River. One major goal in managing this regulated outflow is to maintain a near-constant lake level, thus prohibiting large changes in level. Nevertheless, projected lake levels under the TA method are markedly lower than under the other three methods. Although some data points lie above zero (rising lake level), the 75% quantile is below zero, so the great majority of GCM cases have falling lake levels under the TA method; the median is a drop of 24.5 cm. The CC, PT, and EA methods also have median changes in lake level that are negative, but less decisively so (a drop of 3 cm under the PT method), and their 75% quantiles are all greater than zero, so they have significant chances of rises in lake level. Also, the results for the CC, PT, and EA methods all cluster much closer to one another than to the results of the TA method.

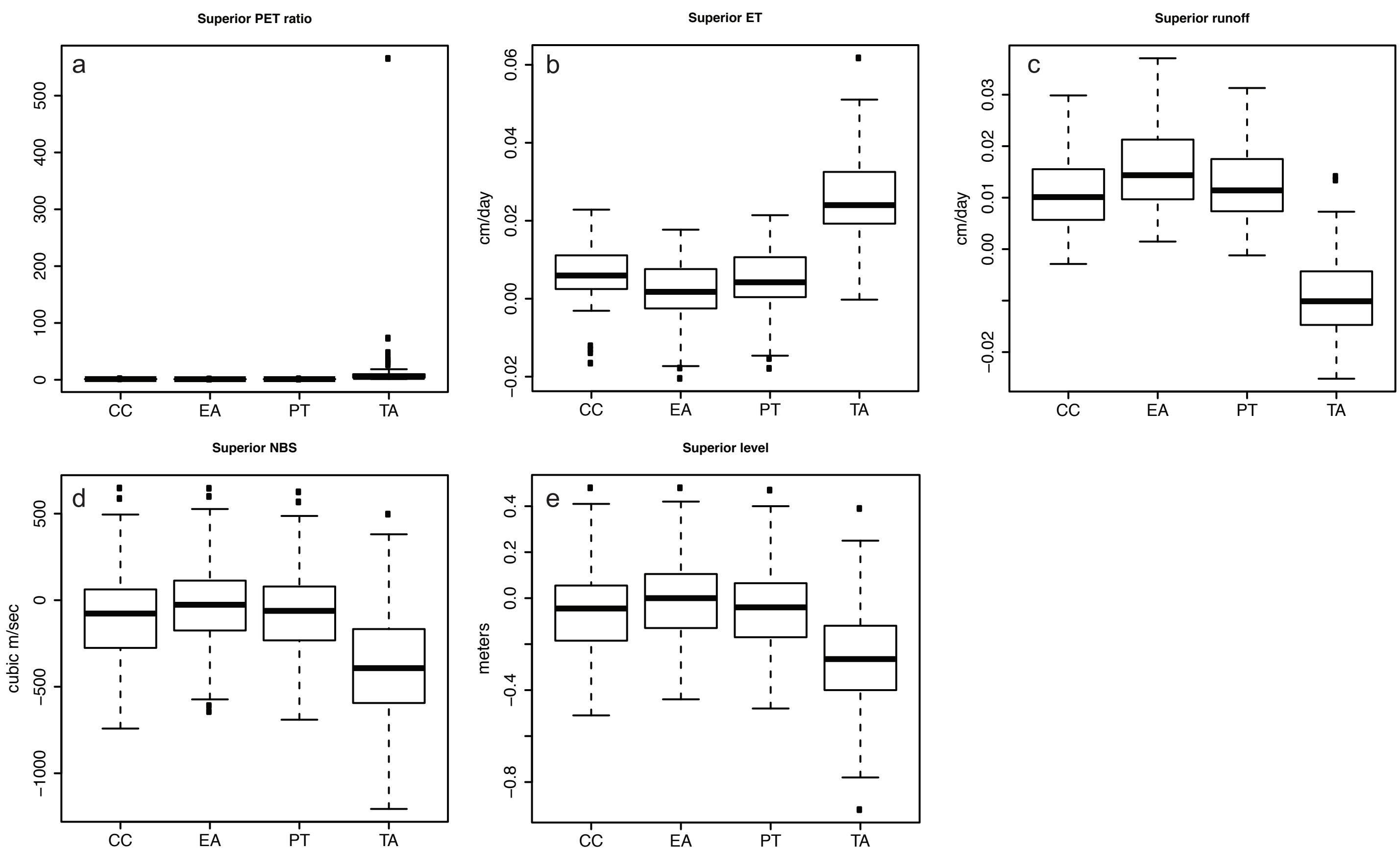


Figure 1. Box and whiskers plots of variables averaged over the Lake Superior basin and over the entire year, generated by GCM outputs in combination with LBRM using the temperature adjustment (TA), Clausius-Clapeyron (CC), Priestley-Tayor (PT), and energy adjustment (EA) methods. Each data point used in making the plots is a single GCM realization. (a) Ratio of PET for future and historical time periods, (b) ET for future minus historical time periods (cm/day), (c) runoff for future minus historical time periods (cm/day), (d) net basin supply for future minus historical time periods (m^3/s), and (e) lake level for future minus historical time periods (m).

3.2. Lake Michigan-Huron basin

Lakes Michigan and Huron are considered a single unit because their attachment at the Straits of Mackinac is wide and deep enough for them to always maintain the same hydraulic level. The results for Lakes Michigan-Huron, Erie, and Ontario largely reflect those of Lake Superior, although the results in terms of PET are less extremely dramatic, while the results in terms of lake levels are of greater magnitude (lake level results are not available for Lake Ontario).

The ratios of PET in the basin of Lake Michigan-Huron (Fig. 2a) are a muted version of the results for Lake Superior (Figure 1a), but again the changes in PET for the CC, PT, and EA methods cluster at values only moderately above unity, while under the TA method, the median value is just above 2, while several model runs project increases by factors greater than 4. These are not as extreme as the values for Lake Superior, mainly because the calibrated values of T_b are generally larger in the Michigan-Huron sub-basins (falling mostly in the range of 5 to 9° C) than those for Superior; the reason for this is not clear, and is part of the black-box nature of the calibration of LBRM.

The actual ET from the Michigan-Huron basin (Figure 2b) is again modulated by availability of moisture. The increases in ET under the TA method strongly outstrip those under the other three methods, but those under the other three methods lean more strongly toward increased ET than in the Lake Superior basin (compare Figure 1b).

In terms of runoff in the Lake Michigan-Huron basin (Figure 2c), as well as net basin supply (Figure 2d), much the same story is told again. The runoff and NBS under the TA method are considerably lower than under the other three methods. While the CC, PT, and EA methods generally show increased runoff, the balance of over-lake precipitation and evaporation place their median changes in NBS at negative values.

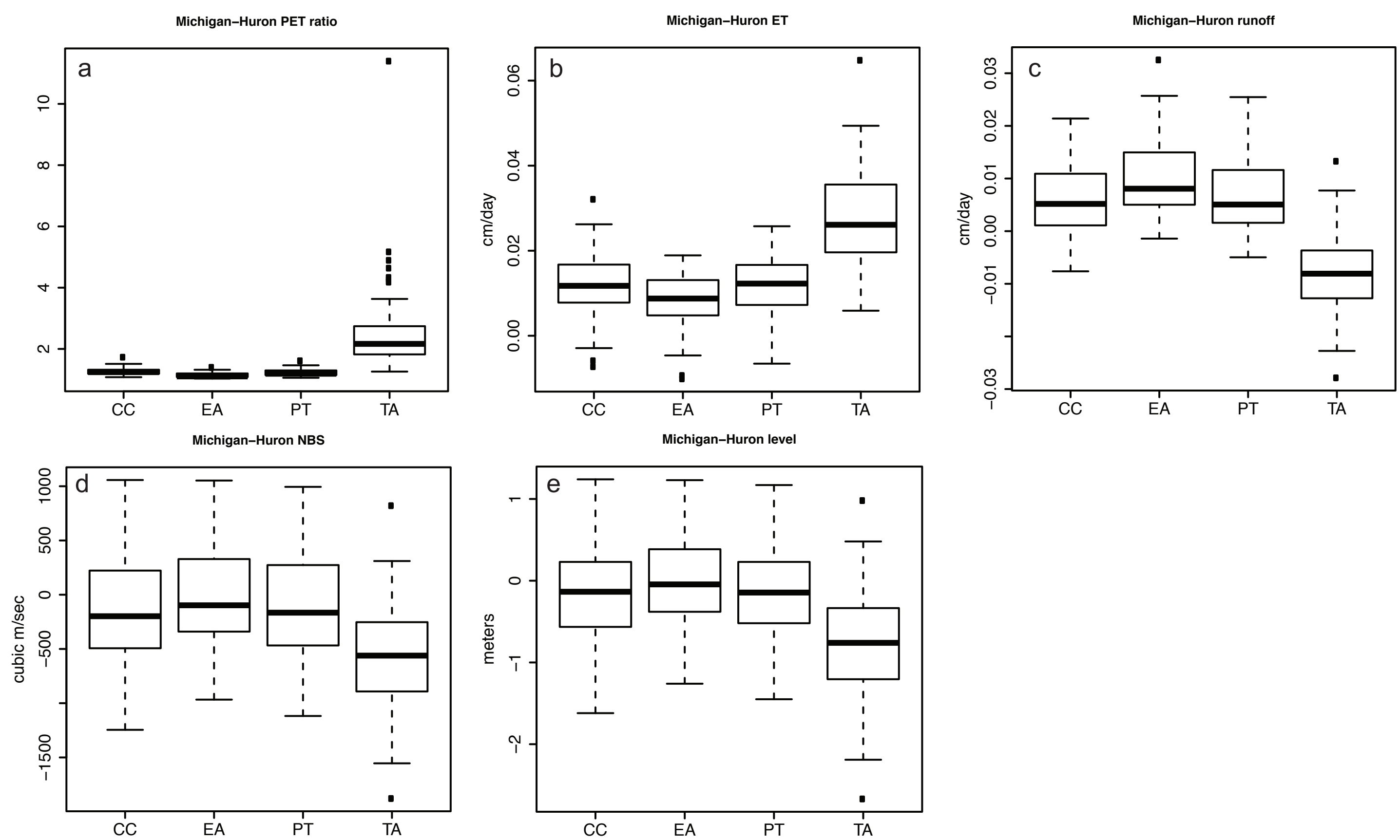


Figure 2. As in Fig. 1, but for the Lake Michigan-Huron basin.

Lake levels on Lake Michigan-Huron are more sensitive than Lake Superior levels because of a combination of factors. In the model results from all methods, the changes in NBS are generally of the same sign for all lakes. This leads to the effect that Lake Superior's outflow due to its NBS has a reinforcing effect on the change in NBS in Lake Michigan-Huron. Furthermore, Lake Michigan-Huron has no regulation at its outlet, in contrast to Lake Superior, so lake level variations cannot be artificially damped by direct human intervention; although this can be changed through construction of regulation works in the future, we assume the status quo in this respect in order to isolate the effect due to climate. Finally, with the accumulation of changes in NBS from the Superior, Michigan-Huron, and Erie basins, the level of Lake Erie also changes in the same sense as the change in NBS for those combined basins. Because the difference in level between Lake Michigan-Huron and Lake Erie is small (just over 2 m), a significant rise in Lake Erie's level can inhibit outflow from Lake Michigan-Huron, requiring a rise in Michigan-Huron's level to restore dynamic equilibrium, and vice-versa for a drop in lake levels. Thus the change in lake levels for Lake Michigan-Huron (Fig. 2e) is qualitatively similar to those for Lake Superior, but the magnitudes are distinctly larger. Under the TA method, the 75% quantile is less than zero, indicating a large majority of GCM cases leading to decreased lake levels. The median value is a drop of 67.5 cm, and the largest drop is 219 cm in the IPSL RCP8.5 run for the mid-21st century (greater by 6 cm than for the same model run in the late 21st century, because of a larger drop in precipitation). The TA results also have a large spread, with a few results showing rises in lake level greater than 1 m.

The CC, PT, and EA results again cluster closer to one another than to the TA results, with median values less than zero and 75% quantile values greater than zero. The median lake level change under the PT method is a drop of 14.5 cm, 53 cm higher than the TA method, and the extreme low, again from the IPSL RCP8.5 run for the mid-21st century, is a drop of 145 cm, 74 cm higher than the extreme under the TA method. The PT and EA methods show a narrower range of changes between the 25% and 75% quantiles than the TA method, but also have several outliers above the whiskers.

Because Lake Michigan-Huron has the largest signal in terms of lake level changes, we include a comparison of lake levels for subsets of the GCM realizations in Figure 3, broken down by time period and RCP scenario. The sample sizes for these subsets are smaller—there are 12 GCM realizations contributing to the RCP 4.5 sets for each time period, and 10 each for RCP 6.0 and RCP 8.5. Figure 3 includes only the TA and PT methods, and for all of the combinations of RCP and time period, PT results in higher lake levels in all of the following measures: upper and lower extremes, 25% quantile, median, and 75% quantile. In some of the panels, the 25% quantile for the PT method is greater than the 75% quantile for the TA method, i.e. the boxes do not overlap. The median projected lake levels under the RCP 4.5 scenarios are lower than those under the RCP 6.0 scenarios, in both the mid-and late century and using both the TA and PT methods, due to a different subset of GCMs included. All combinations of RCP scenario, time period, and method yielded median reduction in lake level. Only the late 21st century under RCP 6.0 and using the PT method has a median change in lake level greater than zero.

3.3. Lake Erie

The results for Lake Erie are again similar. Most values of T_b are even higher for Lake Erie's sub-basins than in the Michigan-Huron sub-basins, so the PET in the TA method (Figure 4a) does not reach such high extremes, but still is very distinct from the distributions for the other methods. Actual ET (Figure 4b) is also distinctly higher under the TA method than the others. Runoff (Figure 4c) is mostly decreased under the TA method, but increased under the other methods. For NBS (Figure 4d), the TA method has a strong decrease, while the other methods have strong chances of changes in NBS both above and below zero.

The lake level for Lake Erie is influenced by the sum of the NBS of Lakes Superior, Michigan-Huron, and Erie, because the others flow into Lake Erie. While the outflow from Lake Erie can be variably partitioned between routing to Niagara Falls for aesthetic purposes and water turbines for power generation purposes, the overall outflow is not controlled by human regulation. Because the head is very large between Lake Erie and Lake Ontario (just under 100 m), there is no backwater effect. The sensitivity of Lake Erie's levels is somewhat diminished relative to Lake Michigan-Huron's. Figure 4e shows that under the TA method, lake levels are shown as having well over 75% likelihood of going down, while under the other methods, zero change is within the range of the box.

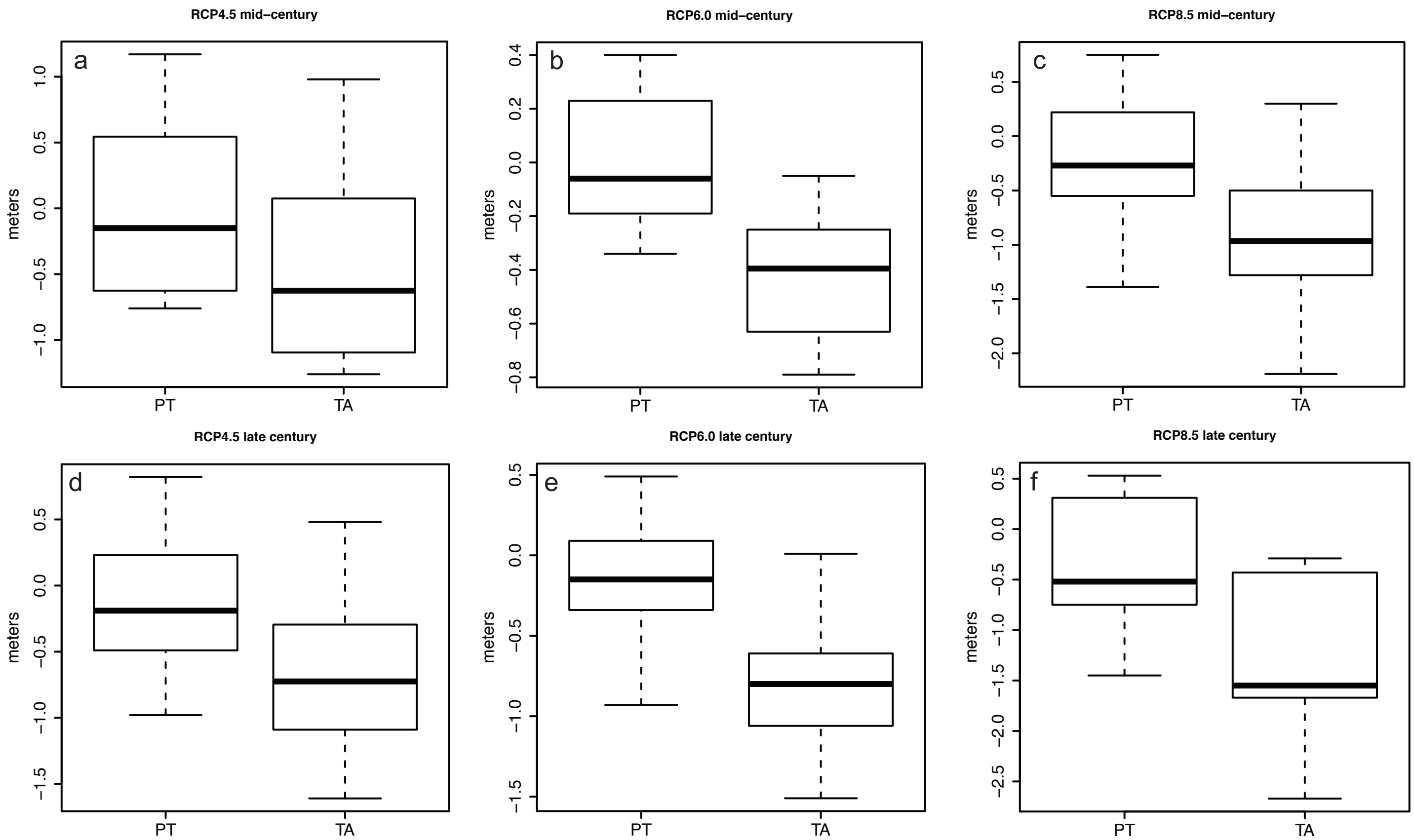
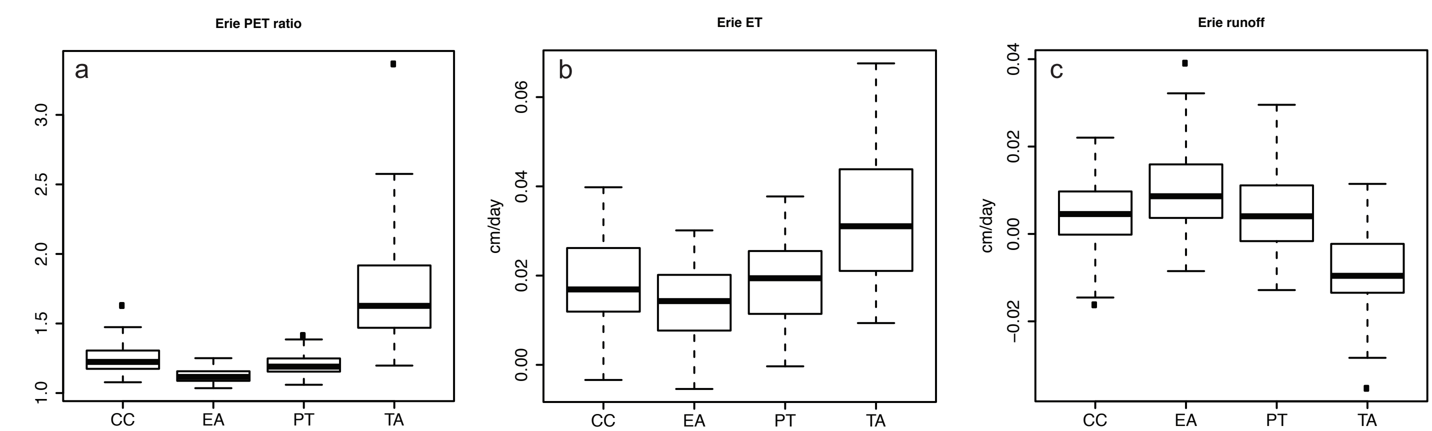


Figure 3. Lake Michigan-Huron water levels represented using the temperature adjustment and Priestley-Taylor methods for subsets of the GCM realizations. Note that the scale differs among the panels. (a) Mid-21st century for RCP 4.5, (b) mid-21st century for RCP 6.0, (c) mid-21st century for RCP 8.5, (d) late 21st century for RCP 4.5, (e) late 21st century for RCP 6.0, and (f) late 21st century for RCP 8.5.



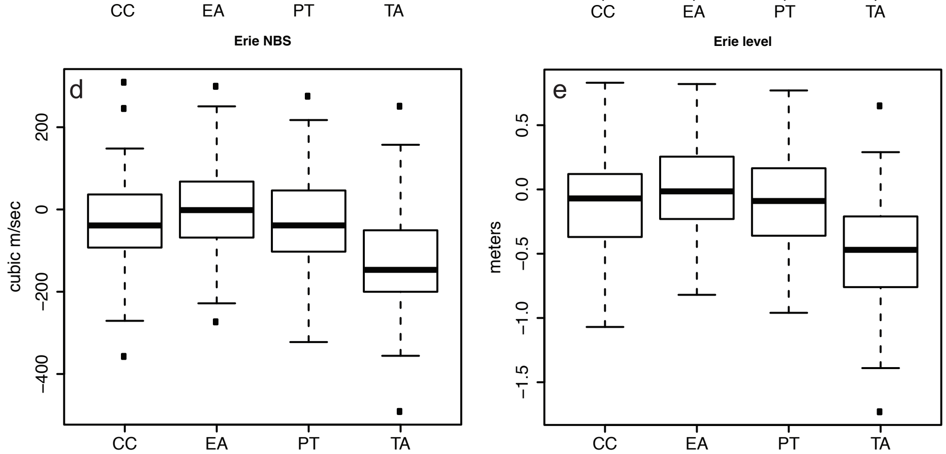


Figure 4. As in Fig. 1, but for the Lake Erie basin.

3.4 Lake Ontario basin

Additional similar results occur on Lake Ontario. The ratios of PET for Lake Ontario (Fig. 5a) occupy a similar range to those for Lake Michigan-Huron, due to the values of T_b there. Actual ET (Fig. 5b), runoff (Fig. 5c), and NBS (Fig. 5d) all show patterns similar to the other lakes, except that the median values of changes in runoff and NBS under the CC, PT, and EA methods are all greater than zero, with the 25% quantile near zero. Largely as a result of lake regulation in practice that frequently deviates from the official regulation plans for the outlet of Lake Ontario, there is not a reliable segment of the CGLRRM to calculate Lake Ontario outflows and levels, and therefore no lake levels are shown here.

3.5 Correlation of Discrepancies with Temperature Change

The amount by which the TA method overestimates the influence of temperature on the various outputs is linked to the magnitude of the temperature change that is a primary input. Figure 6 shows the discrepancy in Lake Superior lake level (result from the CC method minus the TA method and PT method minus TA method) as a function of the difference between the GCM-simulated historical and future air temperature. Lake Superior was chosen to eliminate the influence from other lakes. Figure 6 shows very high correlations—R^2=0.904 in the case of CC minus TA and R^2=0.869 in the case of PT minus TA. This accounts for much of the narrowing of the distributions of all of the variables in Figs. 1-5 under the CC, PA, and EA methods relative to the TA method; the largest increases in PET and ET under the TA method are associated with the largest overestimates and largest temperature increases, while the GCM cases with smaller temperature increases have smaller influences under the TA method, and thus lesser disagreement with the other methods. The narrowing of the distributions under the other methods relative to TA are actually more distinct in the runoff and, especially, ET plots rather than the NBS and lake level plots.

3.6 Statistical significance

The discrepancies between the TA method and all three alternative methods are highly systematic, stemming first from the very large differences in the projected changes in PET. This leads to the results being biased in the same direction regardless of the GCM that is used as input. Therefore, the TA method's projected results are greater than those for the CC, PT, and EA methods in terms of PET and ET, and less in terms of runoff, NBS, and lake level.

One simple way to think of the statistical significance of these comparisons is to think of each comparison between the TA method and each alternative method as having equal chances of either result being greater or less than the other, under the null hypothesis that both distributions are the same. However, this leaves us with the question of how many of these comparisons are independent of each other, for which there is not a clear answer. However, we will use a conservative guess at the number of independent samples that are present. Since PET controls ET, and in turn, runoff, NBS, and lake level, we will consider all of these variables to be dependent on each other, and therefore together constitute only one independent variable. Likewise, we consider the separate results from all of the lakes to constitute only one independent variable, as well as the various results from different RCP scenarios and realizations of the same GCM, and the two different time periods that were compared. Thus, the only dimensions of independence that we have left are the comparison between the CC and TA method on one hand, and between the PT and TA method on the other hand (we regard PT and EA to not be independent), as well as the eight different GCMs that were used. Thus, we regard the full set of experiments as resulting in sixteen independent comparisons between the TA method and alternative methods, all of which show discrepancies of the same sign. The probability of this happening under the null hypothesis is $(0.5)^{16}$ = 1.53×10^{-5} (single-tailed test), meaning that in terms of percentage significance, it is around 99.998%.

We also use the Wilcoxon rank sum method (Hollander and Wolfe 1973) on the lake level data. This is a non-parametric test for comparison of sets of data samples to find whether their distributions are distinct. It requires no assumptions about the shape of the distribution of each, although by combining multiple realizations from each GCM, we are admittedly violating the assumption of complete independence of the "observations" (in this case meaning the method to yield model results). Using the lake level results of the 64 GCM realizations under the TA method and the 64 results from each alternative method, we rank the results from 1 to 128, with lower rank numbers indicating higher lake level, and show the sum of the ranks for all of the realizations under the TA method in parentheses in Table 4. Under the theory of the

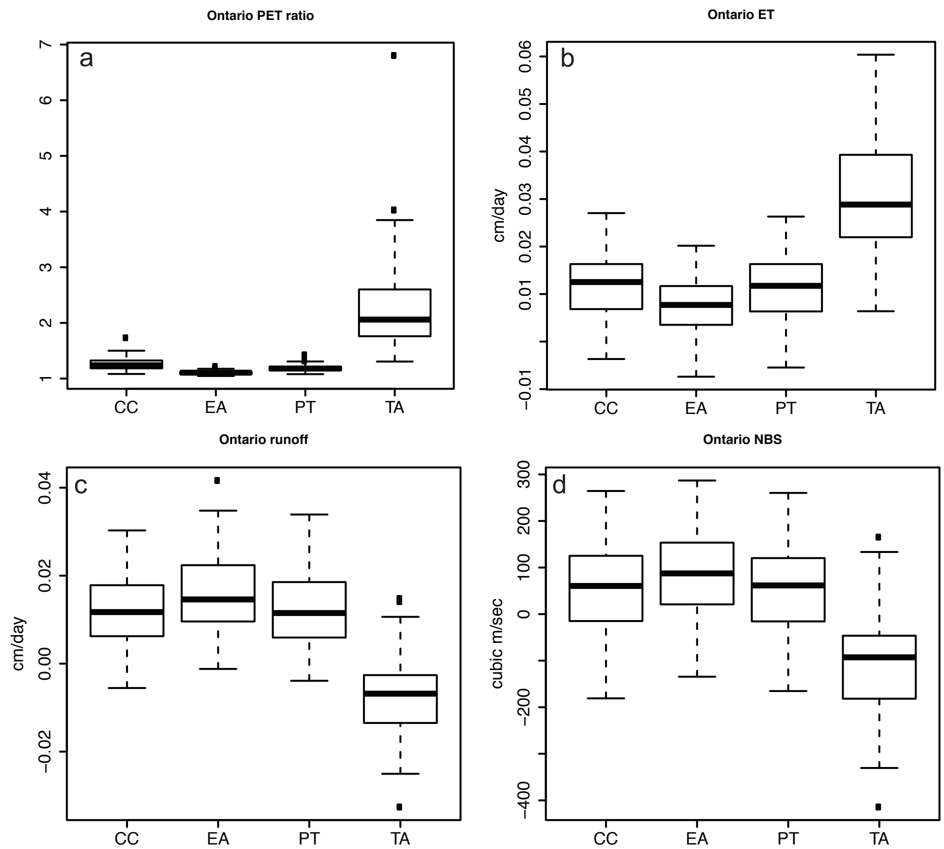


Figure 5. As in Figure 1, but for the Lake Ontario basin, and the lake levels are excluded, as lake levels were not calculated for Lake Ontario.

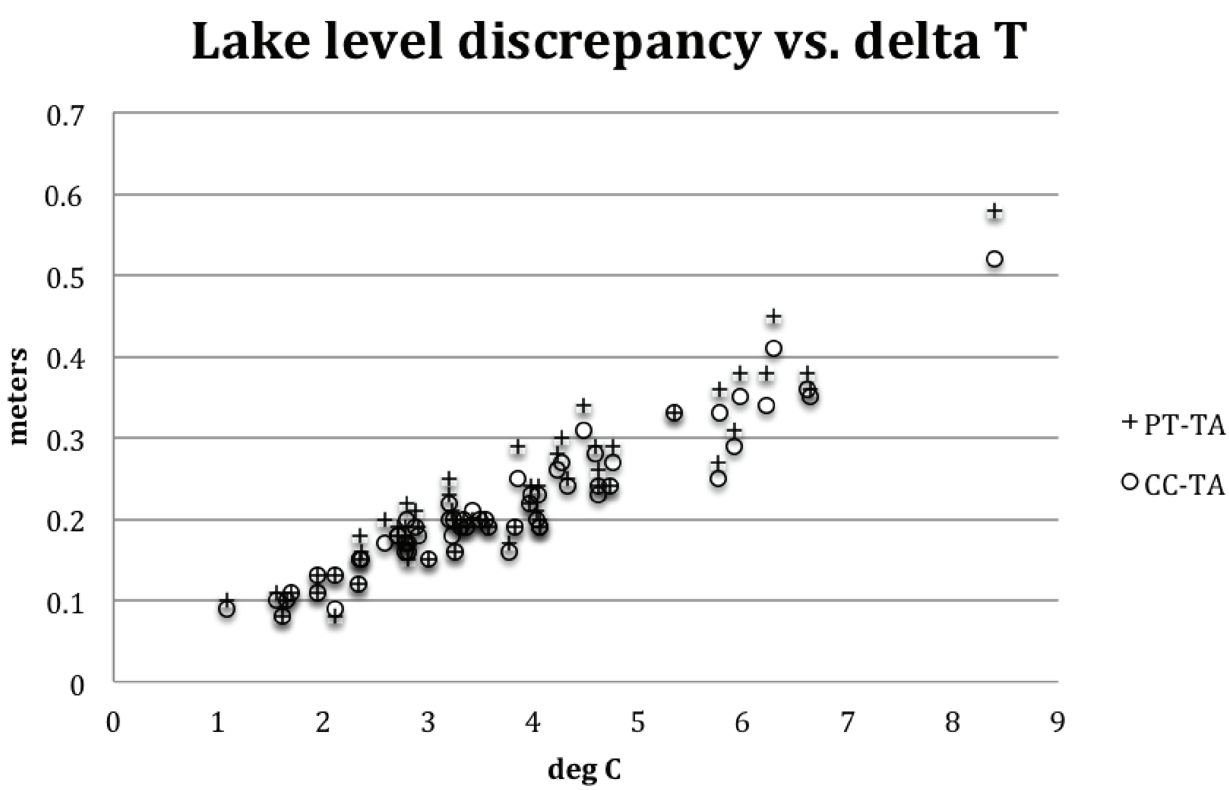


Figure 6. Scatterplot of the difference in lake level (units of m) for Lake Superior, TA minus PT (plus signs) and TA minus CC (circles) vs. change in temperature.

Table 4. Results of the Wilcoxon rank sum test (Hollander and Wolfe 1973) for the lake level of the various lakes, comparing the results under the Clausius-Clapeyron method, the Priestley-Taylor method, and the energy adjustment method to the temperature adjustment method. The statistical significance is stated in terms of the complementary error function (probability that the results under the temperature adjustment method and the alternative method come from the same distribution), while the rank sum for the temperature adjustment method is in parentheses.

	Clausius-Clapeyron	Priestley-Taylor	Energy adjustment
Lake Superior	6.00×10^{-16} (5328)	2.85×10^{-17} (5382)	1.76×10^{-20} (5504)
Lake Michigan-Huron	4.29×10^{-15} (5292)	1.11×10^{-15} (5317)	1.46×10^{-20} (5507)
Lake Erie	5.03×10^{-15} (5289)	9.87×10^{-16} (5319)	1.14×10^{-20} (5511)

Wilcoxon rank sum test, the rank sums for two groups of 64 members each, under the null hypothesis of samples being drawn from the same distribution, have a Gaussian distribution with a mean of 4128 and a standard deviation of 209.8. Thus the complementary error function, indicating the probability that the distributions are the same, can be derived from standard Gaussian probability functions, and are shown in Table 4. The smallest rank sum that we found is under the CC method for Lake Erie. The rank sum of 5289 is 5.53 standard deviations above the mean value, indicating a probability of about five in one quadrillion that the distributions are the same.

These estimates of statistical significance are not perfectly rigorous due to it being unclear how many of the realizations should be regarded as independent, but they indicate levels of statistical significance far beyond the often-used standard threshold of 95% or even 99%.

4. DISCUSSION

The LBRM was carefully crafted to have features such as a mathematical solution that is entirely analytic within each time step (see Croley 1983) and therefore numerically stable under all conditions, as well as a calibration process that thoroughly searches the parameter space. It also conditions the energy available for evapotranspiration (which, divided by the factor of latent heat of evaporation per unit mass, we refer to as potential ET) on total insolation at the surface *during the calibration runs*. However, it seems that the question was not asked whether that energy constraint would carry over to other climatic regimes. One signal that gives a hint at the answer to this is how sensitive the purported energy available for evapotranspiration is to air temperature, with the Clausius-Clapeyron relationship of approximately a 7% increase per degree C being a significant benchmark as a value that should not be greatly exceeded.

In LHW and here, caution has been used to say that in the TA method, air temperature is taken to be a proxy for PET that is equally valid across all time scales and climate regimes (the "universal proxy" assumption), and not to say that the TA method assumes that air temperature causes PET (the causation assumption). Now we will examine how these two assertions relate to one another. The causation assumption is a specific case of the universal proxy assumption. Functionally, in terms of setting up model simulations, these two types of assumption are entirely equivalent; the calculations that are used in the modeling are the same whether one regards air temperature as a universal proxy for PET or as an actual cause. The assumption that air temperature provides a universally valid proxy across time scales and regimes could conceivably also be valid if the same set of mechanisms applied to the relationship between air temperature and PET, regardless of time of year, GHG concentration, and other factors. Could this be the case?

To examine some of the mechanisms by which air temperature varies and how this relates to ET, let us first consider the annual cycle. It is driven by different amounts of solar radiation being incident to the surface at a given location in different seasons. As incident solar radiation is a key constituent of the surface net radiation, which is an explicit factor in the process-based Priestley-Taylor PET, the annual cycle yields a very direct forcing on ET. Air temperature variability on scales of several days is instead driven most strongly by variability in the direction of atmospheric advection, associated in the mid-latitudes with the frequent passage of fronts and air masses. This does not directly force ET in the way that solar radiation does, and in fact warmer air is regularly moister air, thus inhibiting ET. These contrasts between the annual and daily scale were demonstrated in LHW's Fig. 2, which shows high correlation between air temperature and ET for the annual cycle, but almost none at other time scales. This invalidates the universal proxy assumption, as in assertion 2 listed in the present paper's Introduction.

Some other methods are popular for assessment of climate change's impact on water resources and ET, using a one-way coupling of data feeding from a GCM to a hydrologic model. One of the most popular uses the PET equation of Thornthwaite (1948), often within the context of the Palmer Drought Severity Index. Also sometimes used is the formulation of Hamon (1961). These formulations use aspects of the seasonal cycle, often in terms of length of daylight, as explicit predictors within their formulation. This contrasts with LBRM, in which all of the variability in PET is attributed to air temperature as the predictor (regardless of whether this is framed as the universal proxy assumption or the causation assumption). By attributing much of the variability within historical times directly to the seasonal cycle, instead of to air temperature, the sensitivity to air temperature under these other formulations is much less than in LBRM.

Warming due to greenhouse gases has yet another mechanism (aside from solar input and advection), with a seminal explanation demonstrated in Manabe and Wetherald (1967). The warming can be thought of as being "seeded" by a small increase in downward longwave radiation due to the presence of enhanced CO_2 in the atmosphere. The resultant warming of the surface then mixes through the entire depth of the troposphere via convection and circulations driven by baroclinic instability. This warmer troposphere further enhances downward longwave radiation, so the surface and the troposphere mutually warm each other. Furthermore, a warmer atmosphere has more water vapor in it, which is also a greenhouse gas, and this further enhances downward longwave radiation, and further magnifies the warming. However, as the surface warms, its upward longwave radiation also increases, so that the change in the surface net radiation is minimal, despite substantial changes in temperature. The way that this is manifested in the GCMs that we used as forcing can be seen in the relatively small increases in PET that resulted under the EA and PT methods.

Another problematic assumption underlying the LBRM and the TA method is that the constraint of eq. (1), enforced in the calibration runs of LBRM as a means of keeping PET in line with solar radiative heating, universally ensures that the LBRM conserves energy. Eq. (1) does assure is that the total PET over the year is entirely unaffected by the value of T_b, and that actual ET is only minimally affected. But outside of the calibration runs and particularly in altered climate regimes, PET is crucially affected by T_b, as demonstrated by the contrast in the results of the TA and CC runs.

The extreme changes in PET demonstrated under the TA method are easily verified by simple calculations. The lowest calibrated value of T_b is 1.1° C for Lake Superior sub-basin 13 (the Montreal River near eastern Lake Superior). With a high-end increase in temperature of 7°, the factor of increase in PET is $e^{7.0/1.1}$=580.4. Even with a lesser increase, such as 3°, the factor is 15.3. Most of the other Lake Superior sub-basins have T_b values between about 2° and 5° C, yielding lesser but still large increases in PET. Because the base case's PET is constrained to be equivalent to the incident sunlight by eq. (1), these factors of increase in PET can be considered as factors of increase in equivalent incident sunlight, or, more vividly, as the equivalent number of suns in the sky of the virtual world created in the LBRM.

The assessments by Gronewold et al. (2011) and Fry et al. (2014) of LBRM and its application to seasonal to multi-seasonal projection, as well as hindcasting, do not involve large and systematic shifts of air temperature as in this paper. Therefore, the results found there and here do not impact each other very strongly. Additional assessment of LBRM in the context of the connection between air temperature anomalies and errors in runoff in historical periods is called for.

Other strange features of LBRM become evident by examining the set of calibrated parameters in Table 1. For example, parameters 4 and 7 from that table are used in the formula:

$$ET_{USZ} = Ae^{T/Tb} P_4 USZM/(1 + P_4 USZM + P_7 LSZM), \qquad (6)$$

where ET_{USZ} is ET from the upper soil zone, $Ae^{T/Tb}$ is as in eq. (1), P_4 and P_7 are calibrated parameters 4 and 7, and *USZM* and *LSZM* are soil moistures values of the upper soil zone and lower soil zone. There is a similar formula for ET from the lower soil zone. The values of parameters 4 and 7 are shown to vary by more than ten orders of magnitude among the various sub-basins. The other parameters also vary widely, as do the resultant values of storage in the upper soil zone, lower soil zone, groundwater, and surface storage reservoirs. Physical interpretation of this amount of variance is impossible, and is evidence that there is a high amount of compensation among the values of the calibrated parameters in optimizing the runoff against observations. However, as already shown, the sensitivity to climate change under the TA method is crucially dependent on the calibrated value of T_b.

It has been convenient to use the term "paradigm" to describe the repeated use of the LBRM under the TA method to represent response to climate change. However, comparison with Kuhn's (2012) definition of "paradigm" reveals some telling contrasts with his understanding of how a paradigm is created and used. The basic element of Kuhn's definition is that a paradigm is an agreed and shared set of principles that are understood among the practitioners of a branch of science, and which enables further research to refine that understanding. Once the elements of the paradigm are deemed valid in the minds of nearly all practitioners and as long as they remain that way, they become material for textbooks rather than research articles, in order to initiate new practitioners into the paradigm. Because of this need for initiation, reports of the research done by those in the field are largely unintelligible to learned outsiders. Godfrey-Smith (2003) reinforces the idea of paradigms supporting progress, speaking of the need for scientists to rely on the knowledge amassed by the larger scientific community: "If each individual insisted on testing everything himself, science would never advance beyond the most rudimentary ideas."

In contrast to this, use of the LBRM under the TA method does not generate new understanding, but only yields numbers for runoff, which contributes to calculations of NBS and lake level. Instead of requiring hard-earned background knowledge, it has been designed to have inputs and outputs that are easily understood, and internal workings that are invisible to its users, and, to a large extent, to its developers; it is a black box. Its reasonable outcomes for replicating the past were taken as sufficient evidence that its underlying assumptions would carry it to success under contrasting future conditions. The system has two key pieces whose conceptual development and subjection to peer review were carried out separately: LBRM itself and the change factor method that lies behind the TA method. Unlike a paradigm as seen by Kuhn, this combined system's use has not been contingent on critical evaluation by a community, but rather has been uncritically accepted by a collection of users who find an opportunity in extracting the final level outputs that the model suite including LBRM puts forward, correct or not, verifiable or not. We have chosen instead to lift the hood of this black box and evaluate more of the intermediate results that lead to the final results, and whose connection between past and future scenarios can be more easily traced.

The paradigms underlying all fields of science combine objects, their characteristics, and the processes that govern those objects and the interactions among them. It may be that cross-disciplinary research tends to lack true paradigms, and that the substitutes are paradigms that exist within only one field (in this case hydrology), and that recognize only some objects and their properties from a neighboring field (in this case meteorology), but not the relevant processes in that neighboring field's paradigm. In the case of LBRM, even this doesn't account for the problems. The conservation of energy at the surface is a part of the paradigm of hydrology, and is given cursory recognition in the calibration stage of LBRM via eq. (1), but full conservation is not enforced throughout, and the consequences of this were apparently not investigated until LHW.

The climate modeling community does have a strong paradigm for interaction between the atmosphere and surface, whose principles are directly derived from hydrology (it has the same roots as the Priestley-Taylor equation). But the application of LBRM under the TA method has summarily rejected the results of the climate models' calculation of surface fluxes, rather than engaging them and attempting to justify corrections to them.

5. CONCLUSIONS AND FUTURE PLANS

A method for projecting the water levels of the Laurentian Great Lakes under scenarios of human-caused climate change, used almost to the exclusion of other methods in the past, relies very heavily on the Large Basin Runoff Model (LBRM) as a component for determining the water budget for the lake system. LBRM's exclusive reliance on air temperature as a predictor of PET and its very high sensitivity to air temperature cause it to overestimate future ET and cause the entire modeling system to underestimate future lake levels. This excessive sensitivity is at least partially due to LBRM's failure to distinguish between the ET-air temperature correlation associated with the annual cycle (therefore driven largely by changes in incoming solar radiation and net surface radiation) and variability in ET and air temperature associated with other time scales and other causes.

We created three alternative methods for transferring climate outputs from GCMs into the LBRM: the energy adjustment method (EA), the Priestley-Taylor method (PT), and the Clausius-Clapeyron method (CC). We deem all of these alternative methods, especially the PT method, to be more reflective of what is actually happening within the GCM than the TA method. All three of these alternative methods show, relative to the temperature adjustment method (TA), less PET and ET, more runoff from the land and net basin supply for the lake basins, and higher lake water levels in the future. The results from the EA, PT, and CC methods cluster much closer to one another than to the TA method results. The median change in lake level of Lake Michigan-Huron among all the GCM runs shifts from a drop of 67.5 cm under the TA method to a drop of only 14.5 cm under the PT method. This contrast makes the better guess at the median value within a range that is quite small compared to the natural variability that has been observed historically, although the possibility of other outcomes for future multi-decadal mean values has also been shown.

The magnitude of the discrepancies between the TA and other methods is highly correlated with the air temperature change in the driving GCM (larger temperature changes lead to larger discrepancies). This narrows the distributions among the driving GCM runs in changes of all of the variables, although the spread still remains sizable.

Using various methods of estimating the statistical significance, we find that, at minimum, these discrepancies in results are significant at the 99.998% level. Because of the highly systematic sources of error within the formulation of the LBRM under the TA method, we regard these extremely high values of statistical significance as entirely inevitable and largely beside the point—simply the result of a model formulation that, intentionally or not, yields excessive sensitivity of ET to air temperature.

The alternative methods of transferring GCM-predicted variables into the LBRM for future projections are relatively quick and easy modifications to the existing modeling system, for purposes of demonstration of the problems with the traditional method, and are not final answers. As a next step, therefore, we would like to put forward a larger effort to model the Great Lakes basin's water budget using a hydrologic model with a strong basis in the surface energy budget, and driven by a variety of GCM data using a more modern climate downscaling technique.

6. ACKNOWLEDGEMENTS

This paper benefited from discussions with A. D. Gronewold, D. H. Lee, and R. Bolinger, and corrections from anonymous reviewers of its companion paper, to appear in the *Journal of Hydrometeorology*.

7. REFERENCES

Angel, J.R., and K.E. Kunkel. The response of Great Lakes water levels to future climate scenarios with an emphasis on Lake Michigan-Huron. *Journal of Great Lakes Research* 36, Supplement 2, 51–58 (2010).

Brean, H. Climate change will increase evaporation of Colorado River. Las Vegas Review-Journal, February 6, 2015. http://www.reviewjournal.com/news/water-environment/climate-change-will-increase-evaporation-colorado-river, accessed May 12, 2015 (2015).

Chao, P. Great Lakes water resources: Climate change impact analysis with transient GCM scenarios. *Journal of the American Water Resources Association* 35:1499-1507 (1999).

Collins, W.J., N. Bellouin, M. Doutriaux-Boucher, N. Gedney, P. Halloran, T. Hinton, J. Hughes, C.D. Jones, M. Joshi, S. Liddicoat, G. Martin, F. O'Connor, J. Rae, C. Senior, S. Sitch, I. Totterdell, A. Wiltshire, and S. Woodward. Development and evaluation of an Earth-System model-HadGEM2. *Geoscientific Model Development* 4(4):1051-1075 (DOI:10.5194/gmd-4-1051-2011) (2011).

Croley, T.E., II. Great Lakes basins (U.S.A.-Canada) runoff modeling. *Journal of Hydrology* 64:135-158 (1983).

Croley, T.E., II. Verifiable evaporation modeling on the Laurentian Great Lakes. *Water Resources Research* 25:781-792 (1989).

Croley, T.E., II. Laurentian Great Lakes double-CO_2 climate change hydrological impacts. *Climatic Change* 17:27-47 (1990).

Croley, T.E., II, and C.F.M. Lewis. Warmer and drier climates that make terminal Great Lakes. *Journal of Great Lakes Research* 32:852-869 (2006).

Croley, T.E., II, F.H. Quinn, K.E. Kunkel, and S.A. Changnon. Great Lakes hydrology under transposed climates. *Climatic Change* 38:405-433 (1998).

Delworth, T.L., A.J. Broccoli, A. Rosati, R.J. Stouffer, V. Balaji, J.A. Beesley, W.F. Cooke, K.W. Dixon, J. Dunne, K.A. Dunne, J.W. Durachta, K.L. Findell, P. Ginoux, A. Gnanadesikan, C.T. Gordon, S.M. Griffies, R. Gudgel, M.J. Harrison, I.M. Held, R.S. Hemler, L.W. Horowitz, S.A. Klein, T.R. Knutson, P.J. Kushner, A.R. Langenhorst, H.C. Lee, S.J. Lin, J. Lu, S.L. Malyshev, P.C.D. Milly, V. Ramaswamy, J. Russell, M.D. Schwarzkopf, E. Shevliakova, J.J. Sirutis, M.J. Spelman, W.F. Stern, M. Winton, A.T. Wittenberg, B. Wyman, F. Zeng, and R. Zhang. GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *Journal of Climate* 19(5):643–674 (DOI:10.1175/jcli3629.1) (2006).

Donner, L.J., B.L. Wyman, R.S. Hemler, L.W. Horowitz, Y. Ming, M. Zhao, J.-C. Golaz, P. Ginoux, S.J. Lin, M.D. Schwarzkopf, J. Austin, G. Alaka, W.F. Cooke, T.L. Delworth, S.M. Freidenreich, C.T. Gordon, S.M. Griffies, I.M. Held, W.J. Hurlin, S.A. Klein, T.R. Knutson, A.R. Langenhorst, H.-C. Lee, Y. Lin, B.I. Magi, S.L. Malyshev, P.C.D. Milly, V. Naik, M.J. Nath, R. Pincus, J.J. Ploshay, V. Ramaswamy, C.J. Seman, E. Shevliakova, J.J. Sirutis, W.F. Stern, R.J. Stouffer, R.J. Wilson, M. Winton, A.T. Wittenberg, and F. Zeng. The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *Journal of Climate* 24(13):3484-3519 (DOI:10.1175/2011JCLI3955.1) (2011).

d'Orgeville, M., W.R. Peltier, A.R. Erler, and J. Gula. Climate change impacts on Great Lakes Basin precipitation extremes. *Journal of Geophysical Research: Atmospheres* 119, 10799-10812, doi:10.1002/2014JD021855 (2014).

Dufresne, J.-L., M.-A. Foujols, S. Denvil, A. Caubel, O. Marti, O. Aumont, Y. Balkanski, S. Bekki, H. Bellenger, R. Benshila, S. Bony, L. Bopp, P. Braconnot, P. Brockmann, P. Cadule, F. Cheruy, F. Codron, A. Cozic, D. Cugnet, N. de Noblet, J.-P. Duvel, C. Ethé, L. Fairhead, T. Fichefet, S. Flavoni, P. Friedlingstein, J.-Y. Grandpeix, L. Guez, E. Guilyardi, D. Hauglustaine, F. Hourdin, A. Idelkadi, J. Ghattas, S. Joussaume, M. Kageyama, G. Krinner, S. Labetoulle, A. Lahellec, M.-P. Lefebvre, F. Lefevre, C. Levy, Z.X. Li, J. Lloyd, F. Lott, G. Madec, M. Mancip, M. Marchand, S. Masson, Y. Meurdesoif, J. Mignot, I. Musat, S. Parouty, J. Polcher, C. Rio, M. Schulz, D. Swingedouw, S. Szopa, C. Talandier, P. Terray, N. Viovy, and N. Vuichard. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics* (DOI:10.1007/s00382-012-1636-1) (2012).

Dunne, J.P., J. John, E. Shevliakova, R.J. Stouffer, J.P. Krasting, S. Malyshev, P.C.D. Milly, L.T. Sentman, A. Adcroft, W.F. Cooke, K.A. Dunne, S.M. Griffies, R.W. Hallberg, M.J. Harrison, H. Levy II, A.T. Wittenberg, P. Phillipps, and N. Zadeh. GFDL's ESM2 global coupled climate-carbon Earth System Models Part II: Carbon system formulation and baseline simulation characteristics. *Journal of Climate* (DOI:10.1175/JCLI-D-12-00150.1) (2013).

Dunne, J.P., J.G. John, A.J. Adcroft, S.M. Griffies, R.W. Hallberg, E. Shevliakova, R.J. Stouffer, W. Cooke, K.A. Dunne, M.J. Harrison, J.P. Krasting, S.L. Malyshev, P.C.D. Milly, P.J. Phillipps, L.T. Sentman, B.L. Samuels, M.J. Spelman, M. Winton, A.T. Wittenberg, and N. Zadeh. GFDL's ESM2 Global Coupled Climate-Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics. *Journal of Climate* 25(19):6646-6665 (DOI:10.1175/jcli-d-11-00560.1) (2012).

Fry, L. M., A.D. Gronewold, V. Fortin, S. Buan, A.H. Clites, C. Luukkonen, D. Holtschlag, L. Diamond, T. Hunter, F. Seglenieks, D. Durnford, M. Dimitrijevic, C. Subich, E. Klyszejko, K. Kea, and P. Restrepo. The Great Lakes Runoff Intercomparison Project, Phase 1: Lake Michigan (GRIP-M). *Journal of Hydrology* 519:3448-3465, doi:10.1016/j.hyrol.2014.07.021 (2014).

Godfrey-Smith, P. *Theory and Reality: An Introduction to the Philosophy of Science*. University of Chicago Press, 272 pp. (2003).

Government of Canada and U.S. Environmental Protection Agency. *The Great Lakes: An Environmental Atlas and Resource Book*. Third Edition, 46 pp. (1995).

Great Lakes Regional Assessment Group. Preparing for a Changing Climate: The Potential Consequences of Climate Variability and Change, Great Lakes Overview. A Report of the Great Lakes Regional Assessment Group for the U.S. Global Change Research Program, Sousounis, P., and J. Bisanz (Eds.), 106 pp. (2000).

Gronewold, A.D., A.H. Clites, T.S. Hunter, and C.A. Stow. An appraisal of the Great Lakes Advanced Hydrologic Prediction System. *Journal of Great Lakes Research* 37:577-583, doi:10.1016/j.jglr.2011.06.010 (2011).

Hamon, W.R. Estimating potential evapotranspiration. J. Hydraul. Div. American Society of Civil Engineers 87:107-120 (1961).

Hartmann, H.C. Climate change impacts on Laurentian Great Lakes levels. *Climatic Change* 17:49-67 (1990).

Hayhoe, K., J. VanDorn, T.E. Croley II, N. Schlegal, and D. Wuebbles. 2010. Regional climate change projections for Chicago and the Great Lakes. *Journal of Great Lakes Research* 36 (Supplement 2):7-21.

Held, I.M., and B.J. Soden. Robust responses of the hydrological cycle to global warming. *Journal of Climate* 19:5686-5699 (2006).

Hollander, M., and D.A. Wolfe. *Nonparametric Statistical Methods*. Wiley, 503 pp. (1973).

International Joint Commission. Options for Managing Lake Ontario and St. Lawrence River Water Levels and Flows. Final report by the International Lake Ontario-St. Lawrence River Study Board to the International Joint Commission, 262 pp. (2006).

Hurrell, J., M.M. Holland, S. Ghan, J.-F. Lamarque, D. Lawrence, W.H. Lipscomb, N. Mahowald, D. Marsh, P. Rasch, D. Bader, W.D. Collins, P.R. Gent, J.J. Hack, J. Kiehl, P. Kushner, W.G. Large, S. Marshall, S. Vavrus, and M. Vertenstein. The Community Earth System Model: A framework for collaborative research. *Bulletin of the American Meteorological Society* (DOI:10.1175/BAMS-D-12-00121) (2013).

Kling, G.W. et al. *Confronting Climate Change in the Great Lakes Region*. Union of Concerned Scientists, 92 pp. (2003).

Kuhn, T.S. The Structure of Scientific Revolutions: 50th Anniversary Edition. U. of Chicago Press, 264 pp. (2012).

Kutzbach, J.E., J.W. Williams, S.J. Vavrus. Simulated 21st century changes in regional water balance of the Great Lakes region and links to changes in global temperature and poleward moisture transport, *Geophysical Research Letters* 32: L17707, doi:10.1029/2005GL023506 (2005).

Lide, D. R. CRC Handbook of Chemistry and Physics: 86th Edition. CRC Press, 2544 pp. (2005).

Lofgren, B.M. Simulation of atmospheric and lake conditions of the Laurentian Great Lakes region using the Coupled Hydrosphere-Atmosphere Research Model (CHARM). NOAA Technical Memorandum GLERL-165. NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, MI, 23 pp. (2014).

Lofgren, B.M., T.A. Hunter, and J. Wilbarger. Effects of using air temperature as a proxy for potential evapotranspiration in climate change scenarios of Great Lakes basin hydrology. *Journal of Great Lakes Research* 37:744-752, doi: 10.1016/j.jglr.2011.09.006 (2011).

Lofgren, B.M., F.H. Quinn, A.H. Clites, R.A. Assel, A.J. Eberhardt, and C.L. Luukkonen. Evaluation of potential impacts on Great Lakes water resources based on climate scenarios of two GCMs. *Journal of Great Lakes Research* 28:537-554 (2002).

Lorenz, D.J., and E.T. DeWeaver. The response of the extratropical hydrological cycle to global warming. *Journal of Climate* 20:3470-3484, doi: 10.1175/JCLI4192.1 (2007).

MacKay, M., and F. Seglenieks. On the simulation of Laurentian Great Lakes water levels under projections of global climate change. *Climatic Change* 117:55-67, doi:10.1007/s10584-012-0560-z (2013).

Manabe, S., and J.L. Holloway. The seasonal variation of the hydrologic cycle as simulated by a global model of the atmosphere. *Journal of Geophysical Research* 80, 1617-1649 (1975).

Manabe, S., and R.T. Wetherald. Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *Journal of Atmospheric Science* 24:241-258 (1967).

Manabe, S., R.T. Wetherald, P.C.D. Milly, T.L. Delworth, R.J. Stouffer. 2004. Century-scale change in water availability: CO2-quadrupling experiment. *Climatic Change* 64:59-76.

Martin, G.M., N. Bellouin, W.J. Collins, I.D. Culverwell, P.R. Halloran, S.C. Hardiman, T.J. Hinton, C.D. Jones, R.E. McDonald, A.J. McLaren, F.M. O'Connor, M.J. Roberts, J.M. Rodriguez, S. Woodward, M.J. Best, M.E. Brooks, A.R. Brown, N. Butchart, C. Dearden, S.H. Derbyshire, I. Dharssi, M. Doutriaux-Boucher, J.M. Edwards, P.D. Falloon, N. Gedney, L.J. Gray, H.T. Hewitt, M. Hobson, M.R. Huddleston, J. Hughes, S. Ineson, W.J. Ingram, P.M. James, T.C. Johns, C.E. Johnson, A. Jones, C.P. Jones, M.M. Joshi, A.B. Keen, S. Liddicoat, A.P. Lock, A.V. Maidens, J.C. Manners, S.F. Milton, J.G.L. Rae, J.K. Ridley, A. Sellar, C.A. Senior, I.J. Totterdell, A. Verhoef, P.L. Vidale, A. Wiltshire, and G.E.M.D.T. Had. The HadGEM2 family of Met Office Unified Model climate configurations. *Geoscientific Model Development* 4(3):723-757 (DOI:10.5194/gmd-4-723-2011) (2011).

Milly, P.C.D., K.A. Dunne, and A.V. Vecchia. Global pattern of trends in streamflow and water availability in a changing climate. *Nature* 438:347-350 (2005).

Monteith, J.L. P*rinciples of Environmental Physics*. Edward Arnold, 241 pp. (1973).

Notaro, M., V. Bennington, and B. Lofgren. Dynamical downscaling-based projections of Great Lakes water levels. *Journal of Climate,* accepted (2015).

Priestley, C.H.B., and R.J. Taylor. On the assessment of surface heat flux and evaporation using large scale parameters. *Monthly Weather Review* 100:81-92 (1972).

Pryor, S.C., D. Scavia, C. Downer, M. Gaden, L. Iverson, R. Nordstrom, J. Patz, and G. P. Robertson. Chapter 18: Midwest. In *Climate Change Impacts in the United States: The Third National Climate Assessment*, J.M. Melillo, Terese (T.C.) Richmond, and G.W. Yohe (Eds.). U.S. Global Change Research Program, 418-440, doi:10.7930/J0J10112N (2014).

Rotstayn, L.D., S.J. Jeffrey, M.A. Collier, S.M. Dravitzki, A.C. Hirst, J.I. Syktus, and K.K. Wong. Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a study using single-forcing climate simulations. *Atmospheric Chemicsty and Physics* 12:6377-6404 (DOI:10.5194/acp-12-6377-2012) (2012).

Scheff, J., and D.M.W. Frierson. Scaling potential evapotranspiration with greenhouse warming. *Journal of Climate* 27:1539-1558. Doi:10.1175/JCLI-D-13-00233.1 (2014).

Sellers, P.J., Y. Mintz, Y.C. Sud, and A. Dalcher. A simple biosphere model (SiB) for use within general circulation models. *Journal of Atmospheric Science* 43:505-531 (1986).

Thornthwaite, C.W. An approach toward a rational classification of climate. *Geographical Review* 38:55-94 (1948).

Voldoire, A., E. Sanchez-Gomez, C. Decharme, S. Cassou, I. Valcke, A. Beau, M. Alias, M. Chevallier, J. Deshayes, L. Coquart, and F. Chauvin The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dynamics* 40(9-10):2091-2121 (DOI:10.1007/s00382-011-1259-y) (2013).

Walsh, J., D. Wuebbles, K. Hayhoe, J. Kossin, K. Kunkel, G. Stephens, P. Thorne, R. Vose, M. Wehner, J. Willis, D. Anderston, V. Kharin, T. Knutson, F. Landerer, T. Lenton, J. Kennedy, and R. Somerville. Appendix 3: Climate Science Supplement. In *Climate Change Impacts in the United States: The Third National Climate Assessment,* J.M. Melillo, Terese (T.C.) Richmond, and G.W. Yohe (Eds.). U.S. Global Change Research Program, 735-389, doi:10.7930/J0KS6PHH (2014).