Development and application of a North American Great Lakes hydrometeorological database — Part I: Precipitation, evaporation, runoff, and air temperature

Timothy S. Hunter⁎, Anne H. Clites⁎, Kent B. Campbell⁎, Andrew D. Gronewold⁎

⁎ Great Lakes Environmental Research Laboratory, National Oceanic and Atmospheric Administration, Ann Arbor, MI 48108, USA

⁎⁎ Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 3 June 2014
Accepted 24 November 2014
Available online 7 February 2015

Communicated by Jay Austin

Index words:
Hydrological data
Precipitation
Temperature
Evaporation
Great Lakes
Database

ABSTRACT

Starting in 1983, the National Oceanic and Atmospheric Administration (NOAA) Great Lakes Environmental Research Laboratory (GLERL) has been developing and maintaining a historical time series of North American Great Lakes basin-scale monthly hydrometeorological data. This collection of data sets, which we hereafter refer to as the NOAA-GLERL monthly hydrometeorological database (GLM-HMD), is, to our knowledge, the first (and perhaps still the only) to assimilate hydrometeorological measurements into model simulations for each of the major components of the water budget across the entirety (i.e., both United States and Canadian portions) of the Great Lakes basin for a period of record dating back to the early and mid 1900s. Here, we describe the development of data sets in the first (GLM-HMD-I) of two subsets of the GLM-HMD including basin-scale estimates of over-lake and over-land precipitation and air temperature, runoff, and over-lake evaporation. Our synthesis of the GLM-HMD-I includes a summary of the monitoring network associated with each variable and an indication of how each monitoring network has changed over time. We conclude with two representative applications of the GLM-HMD aimed at advancing understanding of seasonal and long-term changes in Great Lakes regional meteorology and climatology. These two examples implicitly reflect the historical utility of the GLM-HMD in numerous previous studies, and explicitly demonstrate its potential utility in ongoing and future regional hydrological science and climate change research.

Published by Elsevier B.V. on behalf of International Association for Great Lakes Research.

Introduction

Starting in 1983, the National Oceanic and Atmospheric Administration (NOAA) Great Lakes Environmental Research Laboratory (GLERL) has been developing and maintaining a historical time series of North American Great Lakes basin-scale monthly hydrometeorological data. Most of this data, which we hereafter refer to collectively as the NOAA-GLERL monthly hydrometeorological database (GLM-HMD), either represents or is directly related to the major components of the Great Lakes water budget. The GLH-HMD can be accessed at the following site: www.glerl.noaa.gov/data/arc/hydro/mnth-hydro.html.

Here we describe the historical development and continuing evolution of the first (GLM-HMD-I) of two subsets of the GLM-HMD including over-lake evaporation, over-land and over-lake precipitation, runoff, and both over-lake and over-land air temperature. We intend to describe the variables from the second subset (GLM-HMD-II), including wind speed, cloud cover, and other hydrometeorological variables indirectly related to the major components of the Great Lakes water budget, in a separate study. We note that ice cover data, as well as hydrometeorological variables at higher spatio-temporal resolutions, are distributed separately through the NOAA Great Lakes Ice Atlas (Assel, 2003; Wang et al., 2012) and, respectively, the Great Lakes Coastal Forecasting System (Schwab and Bedford, 1994).

With the exception of historical over-lake evaporation estimates, which are currently derived from physically-based model simulations, data in the GLM-HMD-I are derived from measurements collected primarily at monitoring stations owned and operated by federal agencies in both the United States and Canada. Examples include tributary flow measurements from both the United States Geological Survey (USGS) and Water Survey of Canada (WSC), as well as meteorological station measurements (including precipitation and air temperature, among others) from both NOAA’s National Climatic Data Center (NCDC) and the Meteorological Service of Canada (MSC). For some of these variables, records extend back to the late 1880s; however, the spatiotemporal density of the respective monitoring networks has changed significantly over time.

The GLM-HMD (subsets I and II), to our knowledge, represents the only comprehensive database of Great Lakes basin-scale water budget variables that overcomes the challenges associated with assimilating measurements distributed across the international border from NCDC and MSC dating to the early and mid 1900s. For further discussion on internationally-coordinated Great Lakes hydrometeorological research, see Gronewold and Fortin (2012). Prior documentation of the GLM-
HMD is limited to two NOAA technical memoranda (Quinn and Kelley, 1983; Croley and Hunter, 1994) targeting an audience that, at the time of their publication, included a relatively small user group of federal agency representatives responsible for water level regulation on Lakes Superior and Ontario (Lee et al., 1997; Clites and Quinn, 2003). Since then, the database has been employed by a growing group of scientists, water resource managers, and planning agencies in applications ranging from historical analysis of regional ecosystem and climate trends to seasonal experimental and operational hydrological forecasting (e.g., Magnuson et al., 1997; Thorp and Casper, 2002; Lofgren et al., 2002; McBean and Motiee, 2008; Ghanbari and Bravo, 2008; Gronewold et al., 2011, 2013a; Deacu et al., 2012; Clites et al., 2014).

Here, we provide a research-oriented synthesis of the GLM-HMD-I while also describing important modifications implemented over the past several decades following the previous updates of Quinn and Kelley (1983) and Croley and Hunter (1994). We begin with a brief synopsis of key attributes of the Great Lakes and the Great Lakes basin that serve as a basis for the spatial framework of the GLM-HMD-I hydrometeorological data sets. We then provide a detailed synthesis of the GLM-HMD-I components. We conclude with two representative examples of the GLM-HMD-I that improve understanding of seasonal and long-term changes in Great Lakes regional hydrometeorology and climatology.

Overview of the Great Lakes hydrologic system

The North American Laurentian Great Lakes (Fig. 1) contain roughly 20% of Earth’s fresh unfrozen surface water (close to 23,000 km²) and, with their surrounding watershed, cover an area of about 766,000 km² across the United States and Canada (Fig. 1 and Table 1). Changes in the rate of water flowing through the Great Lakes and changes in their surface water elevations are driven primarily by changes in the regional water budget, regional meteorology, and the hydraulics of the channels that connect the lakes. The flow of water through the Great Lakes and Great Lakes water levels are also affected (though to a lesser extent) by interbasin flow diversions, the regulation of outflows from Lake Superior and Lake Ontario (Derecki, 1985; Changnon, 2004), and glacial isostatic rebound (Mainville and Craymer, 2005).

Unlike most other freshwater basins of comparable magnitude, quantifying the water budget of the Great Lakes basin requires an explicit understanding of over-lake precipitation and over-lake evaporation. This challenge arises directly from the fact that the Great Lakes collectively constitute the largest freshwater surface on the planet (Lake Superior alone has the largest surface area of any lake) and that few over-lake hydrometeorological measurements are available, particularly when compared to the spatial density and length of record of corresponding terrestrial measurements (for further discussion, see Spence et al., 2011, 2013; Blanken et al., 2011; Holman et al., 2012).

Methodology for developing database components

In this section, we describe our approach to developing each of the components of the GLM-HMD-I including over-lake and over-land precipitation and surface air temperature (Precipitation and air temperature section), over-lake evaporation (Over-lake...
evaporation section), and basin-scale runoff (Runoff section). We then (in Net basin supply and Spatial aggregation of database components sections) describe our approach to calculating the net basin supply (NBS) for each lake basin and to aggregating the data in the GLM-HMD-I over different combinations of individual lake surfaces and their watersheds. The final section (Note on quality assurance section) provides an overview of quality control and quality assurance protocols for the measurements that we assimilate in the GLM-HMD-I.

Precipitation and air temperature

The NOAA-GLERL GLM-HMD-I includes over-land, over-lake, and basin-wide estimates of total monthly precipitation, as well as minimum, maximum, and average monthly air temperature for each of the Great Lakes and their watersheds (and certain combinations thereof). As described in greater detail in the following subsection, minimum and maximum monthly air temperatures are calculated as the average of the minimum daily temperatures over the course of a month, and the average of the maximum daily temperatures over the course of a month, respectively. Precipitation estimates are calculated in millimeters (mm) and, to facilitate comparison with different components of the Great Lakes water cycle (including, for example, interconnecting channel flow rates), converted to cubic meters per second (cms). Conversions between mm and cms for these and other variables in the GLM-HMD-I are based on internationally-coordinated land and lake surface area estimates (Table 1) and the total number of seconds in each month.

The GLM-HMD-I historical precipitation estimates (which range from the mid 1800s to present) are divided into three time periods. Air temperature estimates, as described in greater detail in the following section, are only developed for the most recent of these three periods. The beginning and end of each time period correspond with significant changes in the number of stations in the monitoring network (Figs. 2 and 3) and in the algorithms used to interpolate measurements from that network. For some time periods, monthly estimates are based on aggregating daily measurements. For other (generally older) periods, estimates are based on stations that recorded only monthly summaries.

Period I (mid 20th century to present)

We estimate monthly total over-lake and over-land precipitation, as well as minimum, maximum, and average monthly (dry bulb) air temperature from 1948 to present using daily monitoring station data from NOAA’s NCDC daily data sets including the Global Historical Climate Network — Daily (or GHCN-D, as described in Menne et al., 2012), as well as MSC’s DLY04 and DLY02 data sets. We also develop estimates of daily precipitation and temperature values at individual stations from NOAA’s NCDC hourly Integrated Surface Data (Smith et al., 2011).

For both over-land and over-lake estimates, we utilize data from all stations within each lake basin, and from those stations that are outside but within a limited distance of each lake basin. For time periods when the gage network is sparse, we may utilize stations up to 50 km outside a lake’s basin; however when the gage network is dense, we may only utilize stations that are very close to the basin boundary.

While there are other sources of meteorological station data, the NCDC and MSC measurement summaries are well-suited for the GLM-HMD-I because they are derived from a relatively dense network of terrestrial monitoring stations (Figs. 2 and 3), and because they are

Table 1
Lake and land surface area estimates for each lake basin from the Coordinating Committee on Great Lakes Basic Hydraulic and Hydrologic Data (1977).

<table>
<thead>
<tr>
<th>Basin</th>
<th>Lake surface area (km²)</th>
<th>Land surface area (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior</td>
<td>82,100</td>
<td>128,000</td>
</tr>
<tr>
<td>Michigan</td>
<td>57,800</td>
<td>118,000</td>
</tr>
<tr>
<td>Huron (excluding Georgian Bay)</td>
<td>40,600</td>
<td>51,200</td>
</tr>
<tr>
<td>Georgian Bay</td>
<td>19,000</td>
<td>82,800</td>
</tr>
<tr>
<td>St. Clair</td>
<td>1,110</td>
<td>15,700</td>
</tr>
<tr>
<td>Erie</td>
<td>25,700</td>
<td>61,000</td>
</tr>
<tr>
<td>Ontario</td>
<td>19,000</td>
<td>64,000</td>
</tr>
<tr>
<td>Total</td>
<td>245,310</td>
<td>520,700</td>
</tr>
</tbody>
</table>

Fig. 2. Spatial distribution of meteorological stations within the Great Lakes basin (boundary represented by gray line), and outside but within 50 km of the Great Lakes basin, reporting daily precipitation totals in 1880, 1910, 1940, 1970, and 2000. Bottom right-hand panel indicates total number of corresponding meteorological stations from 1840 to present.
accompanied by metadata and clearly documented data quality assurance procedures. We note, however, that these monitoring networks include very few stations either over the lake surface or on islands distant from the lake shoreline. The beginning of this time period (i.e., 1948) in the GLM-HMD-I is based primarily on a time when many states began collecting data that contributed to the NCDC and MSC daily data sets. The beginning of the period of record for these summaries, in turn, is determined by a significant increase in the number of stations in the monitoring network (Figs. 2 and 3).

We develop precipitation and temperature basin-scale estimates by first interpolating daily precipitation and daily minimum and maximum air temperature measurements across a NOAA-GLERL defined set of subbasins and lake surface areas (Fig. 4) using a modified version (Croley and Hartmann, 1985) of conventional Thiessen weighting (Thiessen, 1911), hereafter referred to as the GLERL-DTP (or daily Thiessen polygon) method.

We then calculate cumulative monthly precipitation over each subbasin and lake surface as the sum of the daily precipitation estimates. Similarly, we calculate monthly minimum and maximum air temperatures for each subbasin and lake surface as the average of the daily minimum and maximum (respectively) air temperatures for each month. Finally, we estimate total monthly precipitation and average air temperature across each lake basin as areally-weighted averages (based on the areas in Table 1) of the corresponding monthly estimates for each subbasin.

Subbasin delineations employed in this procedure (Fig. 4) were developed by NOAA-GLERL in the early 1980s using topographic maps. It is informative to note that, as part of this conventional procedure, the total area of each subbasin includes the surface areas of off-shore islands as well (the surface areas of Isle Royale, Manitoulin Island, and Drummond Island, for example, are included as part of the surface area of the nearest mainland subbasin). The subbasin boundaries were subsequently transcribed onto a 1 km × 1 km grid by assigning each cell to either the lake surface, the land surface within a basin, or the land surface outside of a basin. This process, though it represented the state-of-the-art when it was initiated, leads to a set of subbasin delineations that are outdated relative to many contemporary geospatial frameworks. However, we know of no readily-available alternative representation of the land surface on a comparable spatial scale that is applied consistently across both sides of the US–Canadian border. We therefore present the subbasins in Fig. 4 with the explicit awareness that developing and applying a new generation of state-of-the-art land surface models to the entire Great Lakes basin is an important area for future research (for further discussion, see Gronewold and Fortin, 2012; Fry et al., 2014).

We also acknowledge that estimating over-lake precipitation using shoreline precipitation estimates (particularly over water surfaces as large as the Great Lakes) can lead to both annual and monthly biases. However, for much of the historical period of record of the GLM-HMD-I, alternative sources of information (including over-lake stations, see Fig. 2) were not available (Gronewold and Stow, 2014b). Nonetheless, we view modification or replacement of some or all of the historical GLM-HMD-I over-lake and over-land precipitation estimates using relatively recent regional climate model (RCM) simulations (see, for example, Lofgren et al., 2011; Holman et al., 2012) and very recent over-lake measurements (Blanken et al., 2011; Spence et al., 2013) as a high priority for future research.

**Period II (early 20th century)**

Precipitation estimates in the second time period (i.e., 1931–1947) were developed by Quinn and Norton (1982) at NOAA-GLERL using monthly station data and a modification of the conventional Thiessen polygon interpolation scheme (Thiessen, 1911). Their approach, hereafter referred to as the GLERL-MTP (monthly Thiessen polygon) method, used historical meteorological stations that recorded data on a daily or monthly basis (daily values were subsequently aggregated to monthly values) and interpolated them across a basin-wide domain of 5 km × 5 km grid cells.

The relative weight assigned to the precipitation value from a particular station is equivalent to the number of grid cells for which that particular station is the closest of any station. All stations within each lake
basin, as well as stations outside each lake basin but within 25 km of its boundary, were used to compute monthly over-land precipitation values for this period. Similarly, all stations within 25 km of the lake shoreline, and island-based stations on the lake itself, were used to compute over-lake precipitation. The weights applied to measurements from each station were recomputed for each month to accommodate changes in the spatiotemporal resolution of the monitoring network, as well as periods in which measurements from individual stations were unavailable. While the precipitation records from this period are updated as part of the standard GLM-HMD-I maintenance protocol. For further details regarding the evolution of data for this time period, we direct readers to Quinn and Norton (1982).

**Fig. 4.** Historical NOAA-GLERL subbasin delineations (not to scale) that serve as a basis for spatial interpolation of precipitation, temperature, and runoff measurements for the 1948–present period of record.

Period III (turn of the 20th century)
Over-land and over-lake precipitation estimates for the oldest time period in the GLM-HMD-I precipitation record were computed by the Lake Survey District of the United States Army Corps of Engineers (USACE) using an areally-weighted “district” (AWD) approach (Quinn and Norton, 1982). We hereafter refer to this method as USACE-AWD.
The beginning of the time period for this set of over-land precipitation estimates varies by lake system. For Lakes Superior and Erie, the record begins in 1882, while for Lakes Michigan, Huron, and Ontario, it begins in 1883. For Lake St. Clair, it begins in 1900. For all lakes, this period of precipitation records ends in 1930.

The USACE-AWD method established districts (relatively large areas) and subdistricts (relatively small areas) and subsequently calculated the arithmetic mean of daily precipitation station values within each subdistrict (Fig. 5). Daily subdistrict precipitation values were then combined using an areal-weighting scheme to compute daily precipitation values for each district. Finally, district-wide values were combined, again using an areal-weighting scheme, to compute basin-wide over-land precipitation values. Over-lake precipitation estimates for this time period are based on interpolating measurements from nearshore gaging stations. As with precipitation records for period II, precipitation records from period III are described in greater detail in Quinn and Norton (1982) and are distributed to (but no longer updated for) the general public as part of the GLM-HMD-I.

**Over-lake evaporation**

We estimate total monthly evaporation over each lake by aggregating daily simulations from NOAA-GLERL's one-dimensional Large Lake Thermodynamics Model (LLTM, also referred to as the Lake Evaporation Model, or LEM, as described in Croley, 1989; Croley, 1992; Croley and Assel, 1994). We calculate evaporation first in units of mm (expressed as a depth over the surface of each lake) and then in units of cms, converting between the two using the areas in Table 1. Here we describe three recent improvements to the LLTM that are reflected in the latest set of over-lake evaporation estimates in the GLM-HMD-I.

First, in 2012, we began implementing an alternative formulation of over-lake cloud cover (one of the inputs to the LLTM) that, relative to the pre-2012 methodology, draws from the relatively broad range of meteorological stations in the NOAA NCDC Integrated Surface Hourly Database (Smith et al., 2011). Second, we recalibrated (also in 2012) the LLTM using the most recent set of lake surface water temperature estimates from NOAA’s CoastWatch Great Lakes Surface Environmental Analysis (GLSEA, see Leshkevich et al., 1996; Schwab et al., 1999, for details), and the most recent set of ice cover measurements from the NOAA-GLERL Great Lakes Ice Atlas (Assel and Norton, 2001; Assel, 2005; Wang et al., 2012). Third, and finally, we changed the beginning of the period of record for the over-lake evaporation estimates from 1948 to 1950, setting aside simulations from 1948 and 1949 as a model initialization period.

**Runoff**

We estimate historical monthly runoff by extrapolating daily streamflow measurements (Fig. 6) from both the USGS and WSC across NOAA-GLERL subbasins (Fig. 4) using a conventional flow-per-unit area ratio approach. We hereafter refer to this method as the NOAA-GLERL area ratio method, or GLERL-ARM.

We begin by identifying, for each lake basin, a set of USGS and WSC gages that have a relatively long (roughly five years or more) uninterrupted period of record and that are far downstream but are not influenced by significant “backwater” effects. For every day in our period of
Spatial distribution of USGS and WSC streamflow gages across the Great Lakes basin (boundary represented by gray line) reporting daily measurements in 1880, 1910, 1940, 1970, and 2000. Yellow dots represent the subset of stations that meet GLERL-ARM selection criteria (blue dots represent stations that do not meet the criteria). Bottom right-hand panel indicates total number of gages reporting daily values from 1840 to present.

Interbasin diversions and impacts on runoff estimates

Of the multiple diversions of water into and out of the Great Lakes basin (Fig. 1), the Ogoki diversion, the Long Lac diversion, and the Lake Michigan diversion at Chicago are those that have a relatively significant impact on basin-scale runoff. Importantly, alternative approaches to accounting for these particular diversions may lead to different estimates of runoff and, subsequently, to different estimates of net basin supply.

The Ogoki River (Figs. 1 and 7) is located in the Hudson Bay basin to the north of the boundary between the Hudson Bay and Lake Superior basins, and flows to the northeast before joining with the Albany River and discharging into Hudson Bay. In 1943, the Ogoki River diversion project was completed in an effort to increase hydroelectric power production within the Great Lakes basin (for details, see Day et al., 1982; Clites and Quinn, 2003; Heinmiller, 2007). As a result of this project, water from the Ogoki River can be diverted into the Lake Superior basin via Lake Nipigon and, ultimately, into Lake Superior through the Nipigon River (left-hand panel, Fig. 7). Estimation of runoff from the Nipigon River watershed, therefore, must account for both the magnitude and timing of the Ogoki River diversion as well as the regulation of outflows from Lake Nipigon. In the GLM-HMD-I, we account for these impacts by employing measurements from the Nipigon River flow gage (located roughly 10 km upstream of the mouth of the Nipigon River) in our ARM-based estimate of flow from the entire Nipigon subbasin (Lake Superior subbasin 19 in Fig. 4).

While we recognize that the ungaged portion of the Nipigon subbasin is relatively small, we also acknowledge that our approach to accounting for the regulated flows out of Lake Nipigon can lead to biased estimates of flow from the Nipigon River subbasin. These biases can arise because the “true” (but unobserved) flow-to-area ratio of the ungaged portion is quite different from the flow-to-area ratio calculated from the measurements at the Nipigon River gage. Alternative approaches to accounting for flow diversions into the Nipigon River system include subtracting interbasin flow diversion estimates from the gaged portion of the Nipigon River before applying them (via the GLERL-ARM, for example) to the ungaged portion.

The Long Lac diversion (right-hand panel, Fig. 7) is located to the east of Lake Nipigon, and connects the headwaters of the Kenogami River (which drains to the north towards Hudson Bay) with the Aguasabon River (which drains to the south towards Lake Superior). Unlike our approach to quantifying flow from the Nipigon River, we estimate flow from the Aguasabon River (which is ungaged) using the methodology for other ungaged subbasins (as described in the opening paragraph of Runoff section).

The Lake Michigan diversions at Chicago (Fig. 8) route water from Lake Michigan into the Chicago Sanitary and Ship Canal. Water in the Chicago Sanitary and Ship Canal is then routed out of the Great Lakes basin (dashed line, Fig. 8) to the Illinois and, ultimately, the Mississippi Rivers. Consequently, runoff from a significant portion of the Chicago metropolitan area that is technically within the Great Lakes drainage basin does not actually reach Lake Michigan. We explicitly acknowledge the Chicago flow diversions in our runoff calculations by using a delineation of the metropolitan Chicago area watershed (Fig. 4, Lake Michigan subbasin 15) that excludes the area draining into the Chicago Sanitary and Ship Canal system.
Net basin supply

Net basin supply (NBS) is the total amount of water that enters or leaves an individual lake through a combination of over-lake evaporation, over-lake precipitation, and runoff. Two common approaches to calculating NBS for the Great Lakes include the component and the residual methods (for details and representative applications, see Croley and Lee, 1993; Mortsch and Quinn, 1996; Quinn, 2002; Assel et al., 2004; Deacu et al., 2012). In the GLM-HMD-I, we calculate component NBS for each of the lake basins (in units of mm) as the sum of over-lake precipitation, over-lake evaporation, and runoff values as described in the previous sections. We then convert NBS estimates from units of mm to cms using coordinated lake surface areas (Table 1).

It is informative to recognize that component NBS values do not include the flows that enter or leave a lake basin through the major interconnecting channels (i.e., the St. Marys River, the St. Clair River, the Detroit River, the Niagara River, and the St. Lawrence River), nor do they include flows through interbasin diversions and groundwater seepage. Adding these values to the NBS would result in an estimate of total basin supply (a value not included in the GLM-HMD) and would employ data sets developed and maintained primarily by other federal agencies. Internationally-coordinated estimates of flows through the interconnecting channels, for example, are maintained by the USACE and Environment Canada, and are available both by request from those agencies, and through on-line interfaces developed through the international Great Lakes water levels and hydro-climate dashboard projects (see, for example Gronewold et al., 2013a; Clites et al., 2014). Estimates of groundwater fluxes are available from the United States Geological Survey (Grannemann et al., 2000).

Spatial aggregation of database components

The GLM-HMD-I was originally developed to support a range of basin-scale Great Lakes hydrological modeling projects, including those focused on estimating and forecasting energy and water fluxes over the surfaces of each of the lakes (see, for example, the evolution of NOAA-GLERL lake thermodynamics models, as described in Quinn, 1979; Croley, 1989; Croley and Assel, 1994). For these projects, hydro-meteorological data sets were developed for each of the lakes as well as the water surfaces (and watersheds) of Georgian Bay, Lake Huron without Georgian Bay, and Lake St. Clair. We recognize that many current hydrological modeling projects, however, require basins-scale data sets for all of Lake Huron, and perhaps even for the entire Lake Michigan–Huron system. The GLM-HMD-I therefore includes not only the data sets developed for disaggregated portions of the Lake Huron basin, but for conventional aggregations across the major basins of the Great Lakes system as well.

Note on quality assurance

The NOAA-GLERL GLM-HMD-I is continuously updated using measurements collected at thousands of stations, most of which are operated by federal agencies that employ formal data quality assurance and quality control procedures. As such, accuracy of the data sets in the GLM-HMD-I depends on (and, we believe, is aided by adherence to) agency-specific quality assurance procedures for each data stream. However, we employ supplementary quality assurance procedures to identify and remove values that appear anomalous or inconsistent. These procedures include, but are not limited to, verifying that the

Fig. 7. Schematic representations (reproduced with permission from Annin, 2009) of the Ogoki River diversion (left-hand panel) and the Long Lac diversion (right-hand panel).
range of data for a particular variable falls within a reasonable range of values, derived mainly from historical records.

For example, reported daily precipitation values greater than 10 in. (25.4 cm) are removed; and while we recognize it is not entirely impossible for the precipitation in a region on a particular day to exceed 10 in., we find (through periodic informal cross-validation) it common for these high values, when reported, to in fact be erroneous based on comparison to other nearby measurements. We also (as another example) compare monthly precipitation totals and averages to expected ranges to identify potential anomalies; this approach periodically helps identify meteorological stations for which most (and in some cases, all) of the daily data met our threshold criteria, but the monthly cumulative totals were extreme (sometimes on the order of several meters). Reasons for these infrequent anomalous records vary, and while we attempt to identify and address all of them, we also recognize that even with robust quality control measures, there are (given the size of the Great Lakes basin and the numerous sources of data that contribute to the GLM-HMD) undoubtedly inconsistencies, biases, and unaddressed sources of variability, uncertainty, and error.

Representative applications

Of the many historical and potential future applications of the GLM-HMD, those that most explicitly leverage its relatively long period of record and its assimilation of direct hydrometeorological measurements from both sides of the US–Canadian border are long-term, basin-scale assessments of trends in the Great Lakes water budget and regional air temperatures (Hartmann, 1990; Assel, 1998; Lenters, 2001; Assel et al., 2004; Ehsanzadeh et al., 2013). Here, we introduce two representative assessments that demonstrate the utility of the GLM-HMD-I and reflect previous applications in which it has been used in conjunction with other preexisting data sets to address pressing Great Lakes water resource research- and management-oriented questions.

First, we present and evaluate the GLM-HMD-I time series of historical annual over-lake precipitation, over-lake evaporation, the difference between over-lake precipitation and over-lake evaporation, runoff, and annual average air temperature. We then, building on research documenting changes in the Great Lakes seasonal water level cycle.
Historical trends in the annual water budget and annual average surface air temperature of the Great Lakes basin

The time series of annual data from the GLM-HMD-I indicates periods of both significant interannual variability and long-term trends in historical over-lake precipitation and evaporation (top two rows of Fig. 9); however, the variability and trends are not necessarily the same for each of the lake systems. For example, the GLM-HMD-I indicates that over-lake precipitation on Lakes Superior and Michigan–Huron is, on average, less than that on Lakes Erie and Ontario and that, beginning in the early 1970s, precipitation over Lakes Michigan–Huron, Erie, and Ontario transitioned from a period with predominantly below average values to predominantly above average values. Precipitation over Lake Superior for the past ten years, however, has predominantly been below its long-term average.

Simulated annual evaporation rates in the GLM-HMD-I (second row Fig. 9), though they constitute a shorter period of record than the precipitation estimates, also reflect a long-term trend with noticeably less interannual variability than over-lake precipitation estimates. More specifically, the GLM-HMD-I indicates an extended period of below-average evaporation rates across all of the Great Lakes between the early 1950s and the early 1980s, with evaporation rates rising gradually through the 1980s. The GLM-HMD-I then indicates, particularly for Lakes Superior, Michigan–Huron, and Erie, abrupt increases in over-lake evaporation beginning in the late 1990s and relatively persistent above-average evaporation rates since then. Interestingly, the beginning of the current period of above-average evaporation coincides with the 1997–1998 El Niño event (Trenberth, 1997; McPhaden, 1999), one of the strongest on record, and one associated with changes not only in the climatology of the Great Lakes (including changes in lake surface temperature and ice cover, as described in Assel, 1998; Assel et al., 2000; Gronewold and Stow, 2014b) but in other parts of the world as well (see Chandra et al., 1998; Turk et al., 2001; Navarrete et al., 2002, among many others). We note here that while GLM-HMD-I evaporation rates are based on LITM simulations, others have found these estimates to be consistent with observations of both ice cover (Wang et al., 2012), surface water temperature (Austin and Colman, 2007) and, more recently, evaporation estimates from a small network of eddy-covariance stations (Blanken et al., 2011; Spence et al., 2011).

The net effect of variability and trends in over-lake evaporation and over-lake precipitation is reflected by their sum (third row Fig. 9, quantifying precipitation as a positive contribution and evaporation as a negative contribution). This time series of the difference between over-lake precipitation and over-lake evaporation (which we periodically refer to as “net precipitation”) provides insight into some of the drivers behind significant shifts in the water budget of the Great Lakes over time. For example, we observe that decreasing precipitation over Lake Superior for the past 20 years, and increasing evaporation for an even longer period, collectively propagate into a relatively consistent decrease in net precipitation over Lake Superior over the past 50 years. The net precipitation on the other lakes appears to have more of a cyclical pattern, with oscillations between periods of low and high net over-lake precipitation. Of particular note are pronounced recent decreases in net over-lake precipitation on Lakes Superior and Michigan–Huron, that when presented alongside the contributions of both over-lake precipitation and over-lake evaporation, provide insight into drivers behind recent changes in Great Lakes water levels (Sellinger et al., 2007; Lamon and Stow, 2010; Gronewold et al., 2013b). While data for much of 2014 are still preliminary, we expect future updates to the GLM-HMD, coupled with the analysis presented here, to serve as an important stepping stone towards improving understanding of the hydrologic impacts of the recent extreme cold winter of 2013–2014 on the Great Lakes (for details, see Clites et al., 2014b).

The time series of annual runoff into each of the lakes (fourth row in Fig. 9, in units of cms) indicates that rainfall–runoff relationships may be changing over time due, perhaps, to changes in regional geomorphology, land use, and land cover. While this observation warrants further research, here we underscore the fact that the data compiled in the GLM-HMD-I is one of the only resources available for supporting such a broad spatio-temporal assessment (Gronewold and Fortin, 2012; Kult et al., 2014).

Finally, our analysis of GLM-HMD-I Great Lakes regional air temperatures (bottom row Fig. 9) indicates significant warming over the past several decades, though the records also indicate that current air temperatures may not be entirely dissimilar from the temperature measurements in the 1950s over the basins of Lakes Michigan–Huron, Erie, and Ontario. Air surface temperatures in the Lake Superior basin also appear to have been above average in the 1950s, though the duration and intensity of the current warming period is particularly pronounced given the abrupt shift from below- to above-average temperatures in the late 1990s.

Changes in the Lake Superior seasonal water budget

Our estimates of historical (1950 to present) Lake Superior monthly NBS (top row Fig. 10) provide insight into potential origins of some of the more recent trends in annual and decadal NBS (i.e., Fig. 9). More specifically, our estimates indicate particularly pronounced decreases in NBS in the months of December, February, and August.

Visual analysis of individual components of Lake Superior NBS (rows 2–4 of Fig. 10) offers additional insight into potential drivers behind the trends in the Lake Superior seasonal NBS. For example, our estimates indicate that runoff has (relative to the other NBS components) changed little over time. Changes in over-lake evaporation, however, are more pronounced, particularly in the mid-winter (December and February) and late summer (e.g., August) months, and underscore the impact of changes in over-lake precipitation and over-lake evaporation on seasonal changes in the Lake Superior water budget. The tendency for significant increased evaporation in both the mid-winter and late summer months represents not only a profound change in Lake Superior’s water budget, but in the climatology of the Lake Superior basin as well, and is a critical component of broader scale changes taking place across the entire Great Lakes ecosystem.

Concluding remarks

The NOAA GLM-HMD is one of many resources currently available for understanding changes in the water budget of Earth’s largest surface freshwater system. Recent and ongoing research, including expansion of monitoring networks and advances in hydrologic modeling, will likely propagate into continued improvements in not only the GLM-HMD, but other regional databases as well. More specifically, the GLM-HMD could be expanded to include snow cover and soil moisture estimates from regional hydrological simulation models including those developed at NOAA-GLERL (Croley, 1983; Croley and He, 2005) and at NOAA’s National Operational Hydrologic Remote Sensing Center (or NOHRSC; Lee et al., 2005; Rutter et al., 2008). Yet another potential improvement in the GLM-HMD would be explicit quantification of groundwater flows within the Great Lakes basin. These flows have historically been omitted from regional water balance studies, in part because of the complications in calculating groundwater flows across the entire Great Lakes basin, and in part because of the general assumption that groundwater represents a small fraction of the overall
A contribution to the water budget of each lake system. Given the potential for changes in water withdrawal and groundwater exchange rates, we believe that an explicit assessment of the impacts of groundwater demand and changes in groundwater surface elevation (see, for example, Grannemann et al., 2000; Neff et al., 2005) would be a valuable contribution to Great Lakes water budget research.

Despite the many opportunities for improving the GLM-HMD-I, we know of no other database that synthesizes a similarly broad range of

![Figure 9](image_url)

**Fig. 9.** Time series of historical data from the GLM-HMD-I including total annual over-lake precipitation ($P$), total annual over-lake evaporation ($E$), the difference between over-lake precipitation and evaporation ($P-E$), average annual runoff rates ($R$), and average annual over-lake air temperature ($AT$) for each of the four major Great Lake systems. Alternating colors for each data set reflect values either above or below the long-term average for each lake.

![Figure 10](image_url)

**Fig. 10.** Monthly data from the GLM-HMD-I for Lake Superior including net basin supply (NBS), runoff ($R$), over-lake precipitation ($P$), and over-lake evaporation ($E$). The vertical bars in each panel represent monthly total values from 1950 to 2013. Coloring in each panel differentiates positive from negative values. Horizontal black lines in each panel represent the smoothed trend.
hydrological and meteorological data explicitly for the entire Great Lakes basin for a similarly long (i.e., decades) period of record. In addition, while the GLM-HDM-I is primarily intended to document monthly-scale data interpolated across each of the Great Lakes basins, it also includes estimates at a higher spatio-temporal resolution (such as the daily sub-basin-scale runoff estimates from the GLERL-ARM). These data sets are not officially distributed as part of the GLM-HDM-I, but are available upon request. Importantly, while the GLM-HDM has existed and been routinely updated for decades, this study represents the first time it has been formally documented in the peer-reviewed literature. As such, it provides a more formal basis for future studies, including those that might further analyze drivers behind changes in Great Lakes regional hydrometeorology.

Acknowledgments

Funding for this research was provided by the IJC, NOAA, USAACE, and the Great Lakes Restoration Initiative (administered by USEPA). The authors are thankful to Steve Constant, Lauren Fry, Frank Quinn, Cathy Darnell, John Bratton, and Brent Lofgren for providing valuable comments that improved the clarity of the manuscript. This is NOAA-GLERL contribution number 1740.

References
