A spatial classification and database for management, research, and policy making: The Great Lakes aquatic habitat framework

CrossMark

Lizhu Wang [a,*], Catherine M. Riseng [b], Lacey A. Mason [b], Kevin E. Wehrly [c], Edward S. Rutherford [d], James E. McKenna Jr. [e], Chris Castiglione [f], Lucinda B. Johnson [g], Dana M. Infante [h], Scott Sowa [i], Mike Robertson [j], Jeff Schaeffer [k], Mary Khoury [l], John Gaiot [j], Tom Hollenhorst [m], Colin Brooks [n], Mark Coscarelli [o]

[a] International Joint Commission, Great Lakes Office, P.O. Box 32869, Detroit, MI 48232, United States
[b] School of Natural Resources and Environment, University of Michigan, 440 Church St., Ann Arbor, MI 48109-1041, United States
[c] Institute for Fisheries Research, Michigan Department of Natural Resources and University of Michigan, 1109 N. University, Ann Arbor, MI 48109, United States
[d] Great Lakes Environmental Research Laboratory, 4840 S State Road, Ann Arbor, MI, United States
[e] USGS Great Lakes Science Center, 3075 Gracie Road, Cortland, NY 13045, United States
[f] U.S. Fish and Wildlife Service, Lower Great Lakes Fish and Wildlife Conservation Office, 1101 Casey Road, Basom, NY 14013, United States
[g] Natural Resources Research Institute, University of Minnesota, Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811, United States
[h] Department of Fisheries and Wildlife, Michigan State University, 480 Wilson Road, East Lansing, MI 48824, United States
[i] The Nature Conservancy, 101 E Grand River Ave., Lansing, MI 48906, United States
[j] Ontario Ministry of Natural Resources and Forestry, 300 Water St., Peterborough, ON K9J 8M5, Canada
[k] USGS Great Lakes Science Center, 1451 Green Rd, Ann Arbor, MI 48105, United States
[l] The Nature Conservancy, 8 South Michigan Ave., Chicago, IL 60603, United States
[m] U.S. EPA — Mid-Continent Ecology Division, 6201 Congdon Boulevard, Duluth, MN 55804, United States
[n] Michigan Tech Research Institute, 3600 Green Ct., Suite 100, Ann Arbor, MI 48105, United States
[o] Great Lakes Fishery Trust, 230 N. Washington Square, Suite 300, Lansing, MI 48933, United States

ARTICLE INFO

ABSTRACT

Managing the world's largest and most complex freshwater ecosystem, the Laurentian Great Lakes, requires a spatially hierarchical basin-wide database of ecological and socioeconomic information that is comparable across the region. To meet such a need, we developed a spatial classification framework and database — Great Lakes Aquatic Habitat Framework (GLAHF). GLAHF consists of catchments, coastal terrestrial, coastal margin, near-shore, and offshore zones that encompass the entire Great Lakes Basin. The catchments captured in the database as river pour points or coastline segments are attributed with data known to influence physicochemical and biological characteristics of the lakes from the catchments. The coastal terrestrial zone consists of 30-m grid cells attributed with data from the terrestrial region that has direct connection with the lakes. The coastal margin and nearshore zones consist of 30-m grid cells attributed with data describing the coastline conditions, coastal human disturbances, and moderately to highly variable physicochemical and biological characteristics. The off-shore zone consists of 1.8-km grid cells attributed with data that are spatially less variable compared with the other aquatic zones. These spatial classification zones and their associated data are nested within lake sub-basins and political boundaries and allow the synthesis of information from grid cells to classification zones, within and among political boundaries, lake sub-basins, Great Lakes, or within the entire Great Lakes Basin. This spatially structured database could help the development of basin-wide management plans, prioritize locations for funding and specific management actions, track protection and restoration progress, and conduct research for science-based decision making.

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

Introduction

The Laurentian Great Lakes comprise the largest freshwater ecosystem in the world. Their immense surface area (about 246,000 km^2) and

water volume (about 23,000 km^3) support diverse physical, chemical, and biological components that exhibit complex ecosystem functions and processes (USEPA and Government of Canada, 1995; Wehrly et al., 2013). The Great Lakes drainage basin (about 765,000 km^2) spans a large geographic extent that encompasses a diversity of climatic conditions, soils and vegetation types, streams, inland lakes, wetlands, and wildlife. The Great Lakes waters exhibit diverse habitats from

* Corresponding author. Tel.: +1 519 562 4980.
E-mail address: Wangl@windsor.ijc.org (L. Wang).

shallow water in bays/estuaries in Western Lake Erie to deep waters of Lake Superior, freshwater spawning reefs, extensive length of coastlines, and complex physical processes such as circulation and upwelling patterns. The Great Lakes and their drainage basin provide water and other natural resources for urban, industry, agriculture, transportation, fisheries, and other recreational needs for more than 33.5 million people in the basin (U.S. Census Bureau, 2000; Statistics Canada, 2002).

Managing such a large and complex ecosystem is challenging. Management authority of the Great Lakes is spread across multiple organizations including the federal governments of the U.S. and Canada, eight U.S. states, two Canadian provinces, and many local entities. Current management efforts are hampered by the lack of consistent information that can be shared and easily accessed by federal, state, and local organizations (Wehrly et al., 2013). The development of a basin-wide database of comparable ecological and socioeconomic data information across the region is, therefore, highly desirable. A consistent database could be used to develop basin-wide management plans, prioritize locations for funding and specific management actions, and to conduct research for science-based decision making. In addition, managers and policy makers are faced with the challenge of making management decisions at multiple spatial scales from an individual beach to a particular lake sub-basin, and from a specific lake to the entire Great Lakes Basin (McKenna and Castiglione, 2010a,b). Consequently, there is a need to organize information in a hierarchical spatial framework that allows managers and policy makers to apply information and make decisions at a variety of spatial scales.

The need for consistently managed and spatially comprehensive Great Lakes data and information has long been recognized and many efforts and resources have been invested in data collection and synthesis, database development, information delivery, habitat classification, and mapping. For example, the Great Lakes National Program Office (GLNPO) has coordinated with Environment Canada and the Province of Ontario, and the U.S. Great Lakes states to collect physical, chemical, and biological data from open waters of the Great Lakes since the 1960s and to manage those and other sampled environmental data in the Great Lakes Environmental Database (GLENDA) since 2003 (http://www.epa.gov/glnpo/monitoring/data_proj/glenda/). The Great Lakes Observing System (GLOS) provides access to near real-time and archived observations and modeled forecasts for water levels, wave heights, air and water temperatures, and other lake conditions (http://glos.us/). The Great Lakes Environmental Assessment and Mapping (GLTranslated by Susanne and Hans Ulrich RollEAM) project synthesized basin-wide coarse level anthropogenic data and assessed the Great Lakes health conditions (Allan et al., 2013; http://greatlakesmapping.org/). Great Lakes Environmental Indicator (GLEI; Niemi et al., 2007) and the Great Lakes Wetlands Consortium (GLWC) projects have developed environmental indicator data and assessed condition of much of the Great Lakes coastal wetlands (Niemi et al., 2007). Several efforts also have been devoted to the development of Great Lakes habitat and ecological classification systems (Johnson et al., submitted for publication; Rutherford and Geddes, 2007; McKenna and Castiglione, 2010a,b).

These efforts have met many critical information needs and helped answer management questions that could not be answered otherwise. However, each of these efforts had specific objectives and focused on particular aspects of information needs related to those objectives. For example, most Great Lakes data collection programs (e.g., offshore focused GLNPO, coastal focused GLEI) have collected physicochemical and biological point data at selected locations. Extrapolation of those data or inferences made from those data to unsampled areas is difficult due to the lack of spatial linkage (how a spatial unit connects to, and is influenced by other spatial units) and a common spatial framework. Extrapolation of information from sampled data to unsampled areas is important because scientists and managers lack the time or resources to sample all areas of the Great Lakes ecosystem. Many of the existing databases and data delivery systems serve as data downloading sources for geographic information system (GIS) data layers that are either available for one side of the U.S.–Canadian border or available for both sides but lack data consistency (e.g., the Institute for Fisheries Research's Great Lakes GIS [http://ifrgis.snre.umich.edu/projects/GLGIS/] and the Great Lakes Commission's Information Network [http://www.great-lakes.net/]). Other existing databases and data delivery systems serve as data portals for locally synthesized data to provide links to other web-based databases and information (e.g., GLOS). These portals may be limited by database sources that are discontinued, out of date, or haven't been linked by a common spatial framework. Few of the databases mentioned above provide the ability to scale spatial data as management needs dictate. Although those efforts serve well for their specific purposes, they do not satisfy the increasing need for a Great Lakes basin-wide integrated database and information system with a mechanism that allows spatial information linkage and hierarchical stratification.

The need for Great Lakes basin-wide information integration and spatial linkages has been widely recognized. The Protocol of Great Lakes Water Quality Agreement (GLWQA) of 2012 identified a need for an integrated approach to managing information to achieve adaptive management objectives for nearshore health improvement, nonpoint source pollution reduction from urban and agricultural sources, aquatic invasive control and prevention, species and habitat restoration and protection, nutrient load and concentration reduction, and climate change prediction and adaptation (GLWQA, 2012). Emerging issues, such as record low water levels for Lakes Huron and Michigan in December 2012 and January 2013 (Clites et al., 2014), the dramatic changes in offshore productivity and food web composition likely related to aquatic invasive species in lakes Huron and Michigan during the last decade (Fahnenstiel et al., 2010; Cha et al., 2011; Barbiero et al., 2012; Bunnell et al., 2013), the extraordinarily high nuisance algal blooms in Lake Erie during 2011 (Michalak et al., 2013), and the Lake Erie algal toxins that resulted in more than 400,000 people without tap water for two days in 2014 (http://ecowatch.com/2014/08/03/toxic-algae-bloom-500000-without-drinking-water-ohio/), require basin-wide and lake-wide binational strategies and management actions.

To address these ongoing and emerging challenges, both Canadian and U.S. governments need a spatial framework and database that provides basin-wide information on the location, characterization, status, and quantity of Great Lakes physical, chemical, biological, and human ecosystem components (e.g., GLWQA 2012 Annex 2 — Lakewide Management Plans and Annex 10 — Science). Ideally a basin-wide spatial framework and database would: (1) incorporate an integrated, objective standard for basin-wide and lake-wide condition comparison; (2) provide access to key available data and spatial information to decision makers to enable rapid identification of high priority, cost-effective locations for protection, enhancement, and rehabilitation; and (3) provide a spatial framework for reporting that allows the priority activities and progress of multiple government agencies to be synthesized, assessed, and reported at regional and basin-wide scales. Such a spatial framework and database is the foundation for the most efficient allocation of resources and management actions by binational government agencies (Riseng et al., 2008; Wang et al., 2011).

In 2010, a research team was formed to address the need for Great Lakes basin-wide information integration and spatial linkages. The team was composed of Great Lakes researchers and managers with extensive experience in collecting and synthesizing Great Lakes regional data, conducting regional assessments, or developing basin-wide habitat classifications. The goal of this team was to develop an operational, integrated basin-wide database and a hierarchical spatial classification framework with basic spatial mapping units for the entire Great Lakes Basin and their associated catchments in both Canada and U.S. The resulting geospatial classification framework and database, the Great Lakes Aquatic Habitat Framework (GLAHF), has a common spatial framework attributed with sampled, calculated, and modeled data and a flexible nested structure that enables aggregation of data into larger units characterized by specific criteria or constraint. This hierarchical structure allows data to be synthesized, utilized, and reported at any

required spatial scales. The GLAHF database includes key available basin-wide data that describe natural and anthropogenically modified variation in catchments, coastal, nearshore, and offshore conditions.

The objectives of this paper are to (1) describe the spatial classification framework, hierarchical spatial database structure, data types and sources, and the processes used to attribute such data into the database; (2) report briefly on the Great Lakes aquatic habitat resources and their patterns in natural variation and levels of human disturbances at different spatial scales; and (3) discuss the utility of the spatial framework and associated database.

Methods

Development of hierarchical spatial database structure and classification framework

Spatial database structure

The database structure of GLAHF was developed with input from managers and policy makers from Great Lakes states/provinces, Canadian and U.S. federal management and research agencies, and nongovernmental organizations. This binational database development process was intended to ensure that state-of-the-art knowledge and techniques were used, and the spatial framework was acceptable and useful for multiple users and uses. The database structure is a hierarchical network of grid cells across the Great Lakes aquatic and coastal areas with attributed physicochemical and biological georeferenced data. The

GLAHF database incorporates five types of target spatial zones including river pour points that summarized their associated catchments variables, coastal terrestrial, coastal margin, nearshore, and offshore regions.

The spatial structure of GLAHF consists of 30 × 30 m grid cells in the coastal terrestrial and aquatic regions (depth < 30 m) and 1.8 × 1.8 km grid cells in the offshore (depth > 30 m) region (Fig. 1). Each pour point consists of multiple 30 × 30 m cells (wider river mouths have more cells) in the coastal margin zone where a tributary connects with the coastal margin. The smaller cell size in the coastal regions was chosen to provide data with relatively high resolution for measuring the variable conditions and to be consistent with the size of relevant spatial data in the coastal terrestrial area (e.g., land use, digital elevations). The larger cell size in the offshore region was chosen to provide the data resolution for measuring the relatively less variable conditions in the offshore waters and for reducing database size and data processing time. Each 30-m or 1.8-km cell has a unique spatial identifier that locates the cell within a lake, a lake sub-basin, and management or political boundary units. The data attributed to each 30-m or 1.8-km cell are linked to the unique cell identifier in the GLAHF relational database structure. This spatial structure provides the mechanism for data attributed to each cell to be aggregated, synthesized, analyzed, scaled up or down, and reported at various spatial scales depending on research and management needs.

River pour point and catchment connection with the Great Lakes

An important component of GLAHF is the incorporation of influences from terrestrial catchments to the coastal and offshore regions of the
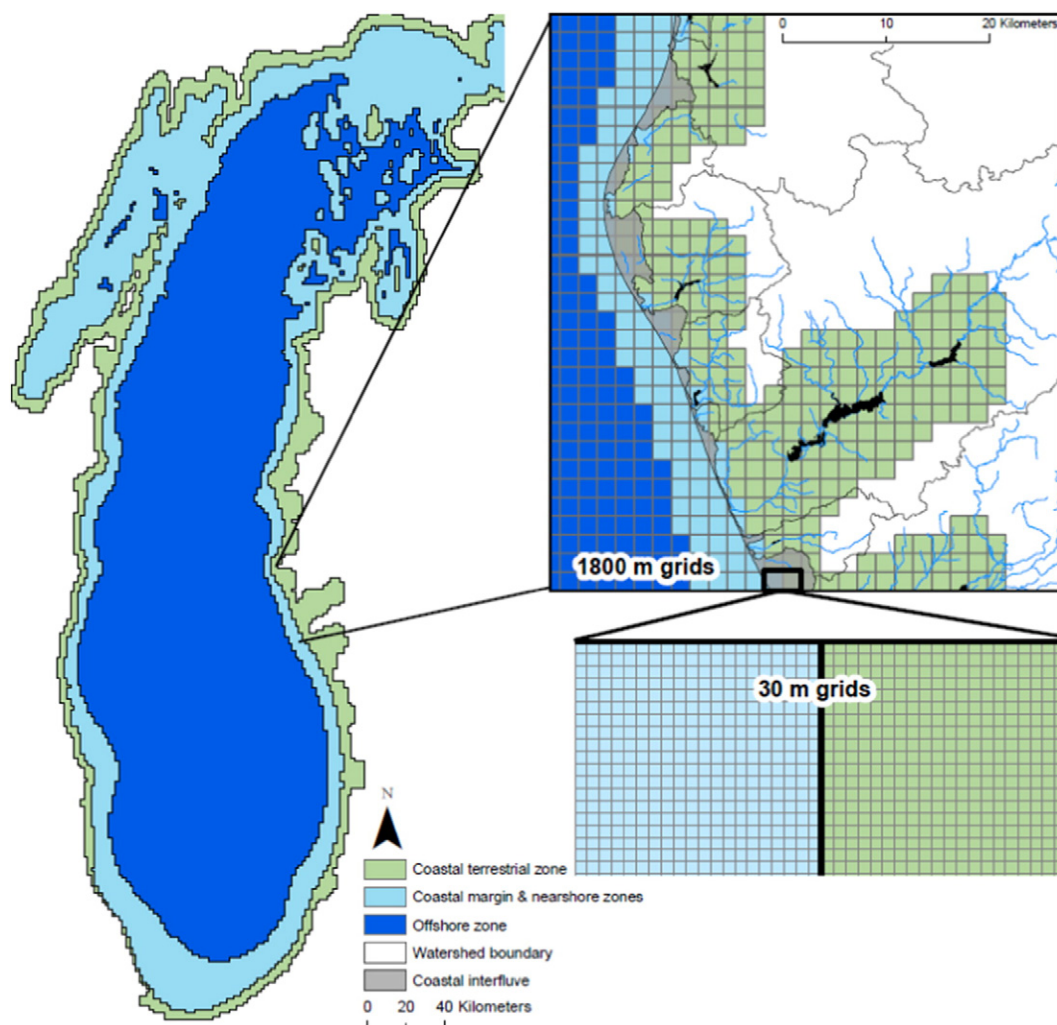

Fig. 1. Data attribution spatial structure consists of catchments, 30-m grid cells in coastal terrestrial and coastal aquatic zones, and 1.8-km grid cells in offshore zones.

Great Lakes. Catchment influences on the Great Lakes are quantified at lake–river connection pour points or coastline segments and further transferred into the coastal and offshore lake regions.

A river pour point is the downstream end of the downstream-most reach of a river network at the coastal region of the Great Lakes. River pour points were identified for river reaches flowing directly into the Great Lakes. The majority of these river reaches have cumulative catchment areas greater than or equal to an 8-digit Hydrologic Unit Code (HUCs, one of the six nested spatial scales of hydrological units used by the U.S.), but small catchments that drain directly into the Great Lakes also occur. We identified consistent basin-wide tributary catchment boundaries using the most current hydrography datasets developed by the U.S. and Canada and a documented repeatable methodology (Danielle et al., submitted for publication) modified from Hollenhorst et al. (2007). For U.S. tributaries, we used the 1:100,000 National Hydrography Database Plus Version 2 (NHD + V2) developed by the U.S. Geological Survey and the U.S. Environmental Protection Agency (NHD Plus V2, 2012) that includes attributed data of physical, chemical, biological, and anthropogenic characteristics for Great Lakes tributaries. For Canadian tributaries, we used the Ontario Integrated Hydrology Dataset (OIHD), hydrologically-enforced flow direction grids (OMNR, 2013) to develop a synthetic stream network and catchments and to generate a stream reach, local catchment, and aggregated catchment dataset that has a similar spatial structure to the NHD + V2 (Hollenhorst et al., 2007; Danielle et al., submitted for publication).

A coastline segment is identified as a shoreline length bordered by interfluve catchments with drainage areas that are between catchments of two river pour points along the Great Lakes coastline. Interfluve catchments which drain local landscapes could have no stream drainage network connections to the shoreline. Data and information from such interfluve catchments were attributed to their corresponding coastline cells, and the connection between the interfluve catchments and the Great Lakes is through coastline segments rather than river pour points.

Spatial classification framework

Our spatial classification framework identifies a hierarchy system with zones that have similar environmental characteristics based on a few dominant processes that shape the physicochemical and biological characteristics in the Great Lakes: tributary inflows, high-energy littoral processes, circulation patterns, and thermal stratification. These zones integrate previously established, ecosystem process defined areas of the Great Lakes (Rutherford and Geddes, 2007; Johnson et al., submitted for publication; LEHTG, 2012).

We defined the largest three levels of GLAHF spatial classification hierarchy as the entire Great Lakes Basin, individual lakes, and lake sub-basins within a lake (Fig. 2). These three spatial scales also include their corresponding catchments and connecting channels (Fig. 3). We delineated four lake sub-basins within each of the lakes Erie, Ontario, and Superior, and five lake sub-basins for each of the lakes Michigan and Huron, including Lake St. Clair and the St. Mary River, St. Clair River, Detroit River, and Niagara River connecting channels (Fig. 3). These sub-basins were manually delineated based on physical and biological processes of dominant spring and summer circulation patterns, bathymetric features, and characteristic biological communities. The sub-basin delineation also considered existing sub-basin boundaries (McKenna and Castiglione, 2010a,b; Johnson et al., submitted for publication) and incorporated reviews from Great Lakes region scientists (personal communication of Drs. J. Kelly and P. Yurista, USEPA Mid-Continent Ecology Division, Duluth MN, and E. Reavie, Natural Resources Research Institute, Duluth MN; July 2013).

Our next level of spatial delineation divided each lake sub-basin into zones: coastal terrestrial, coastal margin, nearshore, and offshore. The coastal terrestrial zone begins at the lake shoreline and extends landward for 5 km (Fig. 1). We defined the lake shoreline as the ordinary high water mark elevation, or the terrestrial edge of hydrologically connected coastal wetlands (Gronewold et al., 2013). The coastal margin zone is defined as lake areas with water depth between 0 and 3 m for all lakes, and the nearshore zone was defined as lake areas with water depth between 3 and 15 m for Lake Erie and 3 and 30 m for the other four lakes. The offshore zone is defined as the lake areas with water depths greater than 15 m for Lake Erie and greater than 30 m for the other four lakes.

These spatial classification zones were empirically derived by the research team through consideration of a combination of physical variables (circulation patterns, thermocline depths, water depth, wave energy, and water levels) and incorporation of knowledge from other
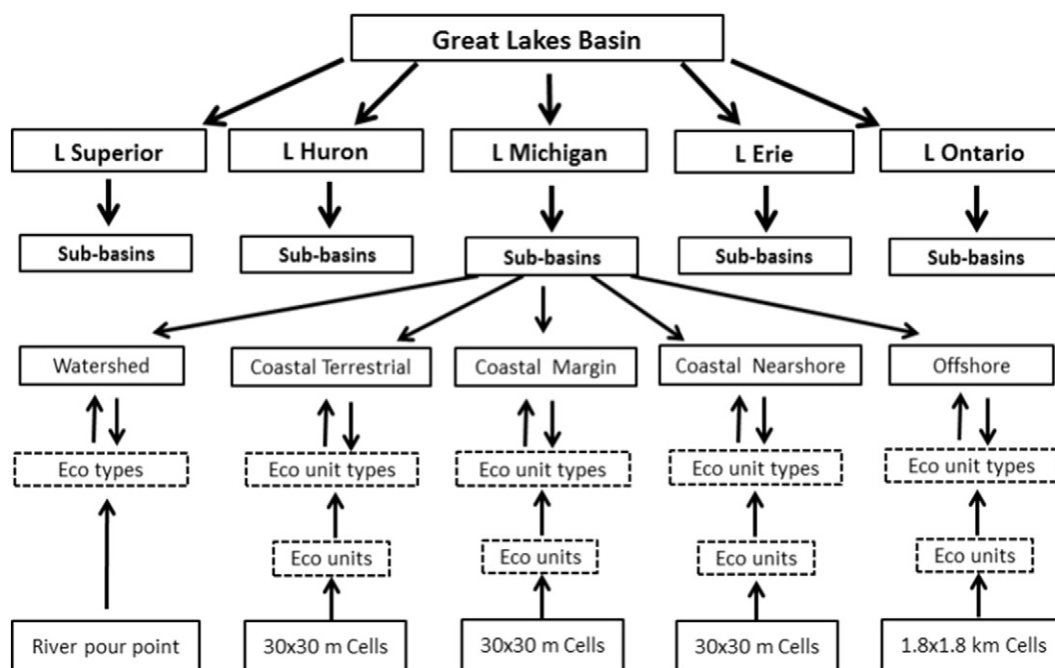

Fig. 2. Great Lakes spatial classification framework. The sub-basin here refers to lake sub-basins and their associated catchments. The single upward arrows illustrate the merging of grid cells into ecological units and classifying such units into ecological types (indicated by dashed-line boxes), which is in progress and not addressed in this paper.

Fig. 3. Great Lakes lake sub-basins and their associated catchments. WSU — Western Lake Superior; CSU — Central Lake Superior; ESU — Eastern Lake Superior; WB — Whitefish Bay; SMR — St. Marys River; NCGeB — North Channel and Georgian Bay; NHU — Northern Lake Huron; EHU — Eastern Lake Huron; CHU — Central Lake Huron; SB — Saginaw Bay; NMI — Northern Lake Michigan; NCMI — North Central Lake Michigan; CMI — Central Lake Michigan; SMI — Southern Lake Michigan; GrB — Green Bay; LSC — Lake St. Clair; WER — Western Lake Erie; CER — Central Lake Erie, EER — Eastern Lake Erie; NR — Niagara River; WON — Western Lake Ontario; CON — Central Lake Ontario; and EON — Eastern Lake Ontario.

studies on Great Lakes habitat classifications that considered similar geophysical variables as drivers for structuring aquatic habitat (Rutherford and Geddes, 2007; McKenna and Castiglione, 2010a,b; LEHTG, 2012). Because the factors determining aquatic habitat are a combination of lake geomorphology and physicochemical processes that occur in transition across multiple spatial scales, we used water depth as a surrogate factor to capture the variability in geomorphic and physicochemical components across a lake (Barton and Schelske, 1982; Yurista et al., 2006; Howell et al., 2012). The coastal margin zone, bounded by the 0–3 m isobaths, is the area where water interacts with the shoreline and reflects wave and seiche energetics and tributary influences most strongly (Johnson et al., submitted for publication). This coastal margin captures embayments, coastal wetlands, rivermouths, and erosional habitats that have been frequently targeted for biological sampling to better understand the importance of coastal ecosystems in life histories of invertebrates, fish, and birds of the Great Lakes region (Jude and Pappas, 1992; Wei et al., 2004). The nearshore zone, bounded by 3–30 m (or 15 m) isobaths, captures the water quality gradient between coastal margin and offshore waters resulting from physical processes such as lake wide circulation patterns, geomorphology, tributary inputs, and alongshore currents that can be consistently delineated around the 30-m isobath (Yurista et al., 2006; Rao and Schwab, 2007). The nearshore zone is also historically the area of the photic zone and is generally not thermally stratified. The offshore zone beyond the 30 m (or 15 m) isobath includes the waters that are influenced by thermal stratification and limited light penetration into deep water.

Data attribution

The GLAHF database is an extensive dataset of over 300 variables organized in a hierarchical folder structure by variable categories such as landscape, climate, physicochemical, and biological data, including their lengthy temporal coverage (Electronic Supplementary Material (ESM) Table S1; also available at http://ifrgis.snre.umich.edu/projects/GLAHF/). GLAHF data exist in spatial (raster and vector) and tabular formats and are stored in ESRI geodatabases and Microsoft Access databases. These data cover minimally a lake sub-basin and are from many sources. Because the data attributed to GLAHF are largely from government and academic research and monitoring programs, we incorporated the metadata and QA/QC source information into the GLAHF metadata. We used this source information to validate the QA/QC process and uncertainty associated with the data to guide our error checking as data were processed. Because GLAHF attributes only summarized data (not raw data), we developed GLAHF metadata following Federal Geographic Data Committee standards and include source references, data-processing and summarization methods, and constraints and uncertainties. Here we briefly summarize methods used to standardize, integrate, and attribute those data into the spatial framework.

River pour point, coastline segment, and coastal terrestrial zone data attribution

We attributed vector and raster data from various sources that are available for the Great Lakes catchments to the tributary pour points, coastline segments, and coastal terrestrial zone (ESM Table S1). For catchments, we summarized and attributed upstream catchment data of Great Lakes tributaries to the pour points or interfluve catchments to coastline segments. Vector data (e.g., surficial geology; shoreline geomorphology) were summed by length or area using the maximum combined area method that assigned a grid cell with the value that covers the most area. Raster data (e.g., land use, including inland lakes) were summed by area for each data category using the averaging method that assigned the average value from the source data to the grid cells.

Some variables were derived from multiple sources with various temporal and spatial resolutions and thus were cross-walked to arrive at consistent classes across the basin. For example, land use/cover data were available from the 2001 U.S. National Landcover Database (NLCD) with 8 level-I classes and 16 nested level-II classes, and from the 2000 Provincial Landcover Ontario (PLO) dataset with 28 classes. These land cover classes were reclassified into 12 common categories and mosaicked to create a seamless basin-wide dataset.

A series of available data were attributed to the tributary pour points, coastline segments, and coastal terrestrial zone grid cells as proportion (e.g., land use), average (e.g., population density), or total amount (e.g., road length). These data included surficial and bedrock geology types, soil types, land use/cover types, imperviousness, farm animal density (Agricultural Census, ESM Table S1), human population density, length of road, wastewater treatment plants, and 20-year means of air temperature and precipitation (ESM Table S1). The catchment size and slope, total river length, number of dams per river length, and percentage of groundwater contribution were calculated by pour points or coastline segments. The lengths of rivers accessible from the Great Lakes were calculated from pour points upstream along the NHD + V2 or OIHD to the first barrier. The barriers were included in the National Anthropogenic Barrier Dataset for the U.S. and the Dams and Barriers layer for Ontario (USACE and OMNR; ESM Table S1). The coastline segments were also attributed with number of boating docks and marinas, shipping ports, and shoreline development index values.

Coastal margin, nearshore, and offshore zones data attribution

Data attributed to the coastal margin, nearshore, and offshore zone grid cells include modeled water thermal, wave, and movement variables; mapped depth, substrate, and wetland variables sampled chemical and biological variables; and occurrence of aquatic nuisance species and shipping traffic variables (ESM Table S1). Data were available in vector, raster, and point format. Raster data were summarized using resampling methods and vector data using maximum combined area methods similar to the catchment data described above. Point data were attributed to a grid cell by geographic location.

Vertical water temperature, circulation, and wave data were available from existing models (NOAA GLCFS; USACE WIS) (see ESM Table S1) for all aquatic areas of the Great Lakes at fine temporal resolution (hourly or daily records). These data required intensive data processing to be summarized at spatial and temporal resolutions useful for characterizing trends. Temporal data with hourly or daily records were summarized across daily or monthly time-steps from 1990s through the most recent period available (ESM Table S1) to provide measures of seasonal and annual variation. Monthly average vertical water temperatures and circulation were modeled every three hours for 20 or 21 vertical levels (Beletsky et al., 2006; Schwab et al., 2009, NOAA's Great Lakes Coastal Forecasting System, ESM Table S1). We summarized the temperature data for three vertical depth classes of 0–20 m, 20–40 m, and >40 m to represent the average conditions of the epilimnion, metalimnion, and hypolimnion strata within the Great Lakes (Cline et al., 2013). We also calculated mean cumulative degree-days by summarizing average daily water temperatures for the 0–20 m water column for all days that were above 0 °C from January 1 through December 31 for 2006–2012. The circulation data were averaged monthly and seasonally across all depth bins for magnitude (m/s) and direction (degrees). A binary term describing the seasonal averaged on/offshore direction was derived by comparing the circulation direction angle to the generalized shoreline angle in a clockwise pattern around each lake. The mean and maximum wave height and period were calculated monthly and seasonally from the modeled data of U.S. Army Corps of Engineers (USACE) and Wave Information Studies and buoy data of Environment Canada (for Lake St. Clair hourly from 2000 to 2012). The USACE modeled hourly waves for 1950 locations around the shoreline of the Great Lakes from 1979 to 2012.

Averaged monthly estimates of surface water temperature were obtained from remote sensing estimated daily data (NOAA Great Lakes CoastWatch, http://coastwatch.glerl.noaa.gov/) from 1994 to 2013. From these surface water data we calculated two thermal metrics relevant to Great Lakes biota: upwelling and spring rate of warming. We calculated daily areas of upwelling where surface temperatures are lower than a threshold value from the same data. The threshold value was calculated by taking the median surface temperature minus 4 °C as a constant (Plattner et al., 2006; Wegscheider, 2006). We then summarized these daily values annually from 1994 to 2013 as an upwelling index for each of the coastal aquatic and offshore grid cells (Fig. 4). We generated an index of spring warming to identify areas of potentially higher spring and early summer productivity (Fig. 4). This index is the difference between June 1 and March 1 surface water temperatures (NOAA CoastWatch data from 1995 to 2012) divided by the number of days during that period to estimate the averaged change in temperature per day and was calculated for each grid cell.

The mapped depth, substrate, and wetland variables were also available for all areas of the three aquatic zones at a relatively coarse spatial resolution and required substantial data acquisition and assembling. Water depth was derived from NOAA's topobathymetry and bathymetric grids with 90-m horizontal resolution (ESM Table S1). Geological landforms of lake bottom were derived for the Great Lakes subaqueous terrain features by incorporating relief and depth into an integrated measure of bottom topology (Gallant et al., 2005; Fig. 4). Substrate compositions were aggregated from various publications (Lewis, 1966; Powers and Robertson, 1968; Moore and Meyer, 1969; Thomas et al., 1972, 1973; Bertrand et al., 1976; Dell, 1976; Thomas and Dell, 1978; Robbins, 1986; Thomas, 1988; Rasul et al., 1997; Boase, 2003; Eyles and Mohr, submitted for publication; LEHTG, 2012). Substrate data included the dominant substrate type and percentage of each substrate type. Substrate types were cross-walked across the five lakes into six substrate types (clay, silt, sand, mud, hard, rock) although the original substrate types for each lake were maintained in the database because studies reported different levels of substrate specificity among lakes. Fish spawning habitats were summarized from a variety of sources that were generated by examining catch records of spawning females, and from survey data on adult spawners, eggs, or larvae locations (Goodyear et al., 1982; also P. Chow Fraser and E. Rutherford, unpublished data). Coastal wetland data were obtained from the Great Lakes Coastal Wetland Consortium (Albert et al., 2005).

Chemical and biological data were included for areas where basin-wide sampling occurred and were available. The integration of these data into our database required significant processing efforts due to sampling method differences. The nitrogen and phosphorus concentrations, organic toxic contaminants, heavy metal concentrations, and compositions of phytoplankton, zooplankton, and benthic macroinvertebrate assemblages were obtained from the Great Lakes Environmental Database and Environment Canada sources. Fish species occurrences were obtained from the U.S. Geological Survey Great Lakes Science Center, Ohio Department of Natural Resources, and Ontario Ministry of Natural Resources.

Occurrence data for aquatic nuisance species were obtained from NOAA's Great Lakes Aquatic Nonindigenous Species Information System. The shipping and vessel traffic data were from the U.S. Coast Guard and made available through Marine Cadastre for 2009–2011 (ESM Table S1). The U.S. Coast Guard requires commercial and private ships of 30+ gross tons and passenger ships and tankers of 150+ gross tons to use the automatic identification system to locate the vessels during their voyage in U.S. waters and records the information on a one minute time step. These locations were then compiled into buffered shipping paths to represent the dominant shipping lanes in the Great Lakes. ESM Table S1 provides more details on data sources and variable types.
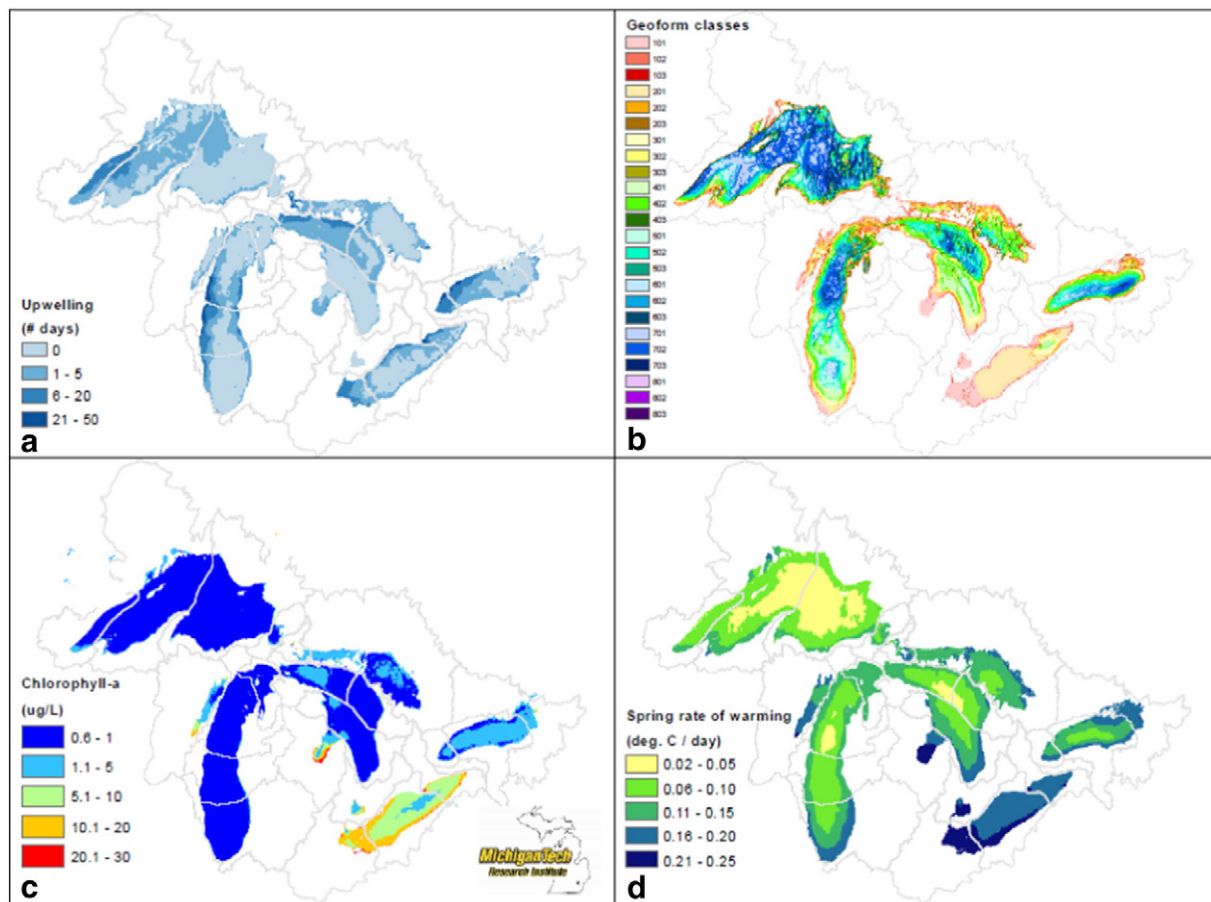
Fig. 4. Mean (2006–2012) upwelling (a), geoform class (b), mean (2008–2012) chlorophyll-*a* concentrations (c), and mean (2008–2012) spring rate of warming (d). The geoform classes are calculated by classifying the combinations of lake bathymetry depth and lake bottom relief into 24 geoform classes labeled using standard marine classes (Dietrich et al., 1980).

Results

Spatial classification framework and associated database

The resulting product of the binational research team is the GLAHF classification framework and database. The classification framework consists of catchments, coastal terrestrial zone, coastal margin zone, nearshore zone, and offshore zone from each lake ecosystem's outermost boundary toward the center of the lake. The catchments captured in the database as river pour points or coastline segments were attributed with available data and information that are known to influence the lakes; the coastal terrestrial zone was attributed with data from the terrestrial region that have direct impacts on lakes; the coastal margin zone was attributed with data about coastline conditions, coastal human disturbances, and highly variable physicochemical and biological characteristics; the nearshore zone was attributed with physicochemical and biological data that are moderately variable; and the offshore zone was attributed with data that are relatively spatially less variable compared with the other water zones.

These five spatial classification zones are further linked or nested within lake sub-basins and political boundaries, such as state/provincial or management boundaries. The hierarchical structure of the database and spatial classification allows information to be synthesized following ecological hierarchy from grid cells to classification zones, a lake sub-basin, a lake, and all the five Great Lakes (Fig. 5). It also allows information to be synthesized following a political boundary hierarchy from grid cells to classification zones, a management unit, a state or province, multiple states/province, and the boundary between U.S. and Canada (Fig. 5).

The database contains over 300 variables (ESM Table S1) with hundreds of thousands of data records. The data sources include direct

website downloading, data acquired from partners (e.g., Rutherford and Geddes, 2007; McKenna and Castiglione, 2010a,b), and personal contacts with government agencies, research institutions, nongovernment organizations, and individual data holders. GLAHF database consists of three types of data in addition to its hierarchical spatial unit boundaries. The first type of data is GIS layers, such as land cover/use and bathymetry, which are data largely generated by government agencies or government funded projects and are available for all spatial units. The second type of data is model predicted data, such as daily water temperature and water current, which are projections by researchers based on relationship among sampled data and predictors and are largely available for all spatial units. The third type of data is monitored or sampled data, which is generally available only for some spatial locations even if they are sampled to represent an entire lake or a lake sub-basin.

Great Lakes aquatic resource distribution patterns

Below we summarize several key factors using both political and ecological boundaries from the attributed data to describe general patterns. The intent of this description is to illustrate how information can be summarized and used at different spatial scales, rather than to exhaustively synthesize the characteristics of the Great Lakes using the many physicochemical and biological factors in our database. Although some information has been reported sporadically in the scientific literatures or government reports, integrating all relevant important information in one place provides a holistic view of spatial patterns and driving forces that control the characteristics of the Great Lakes ecosystem.
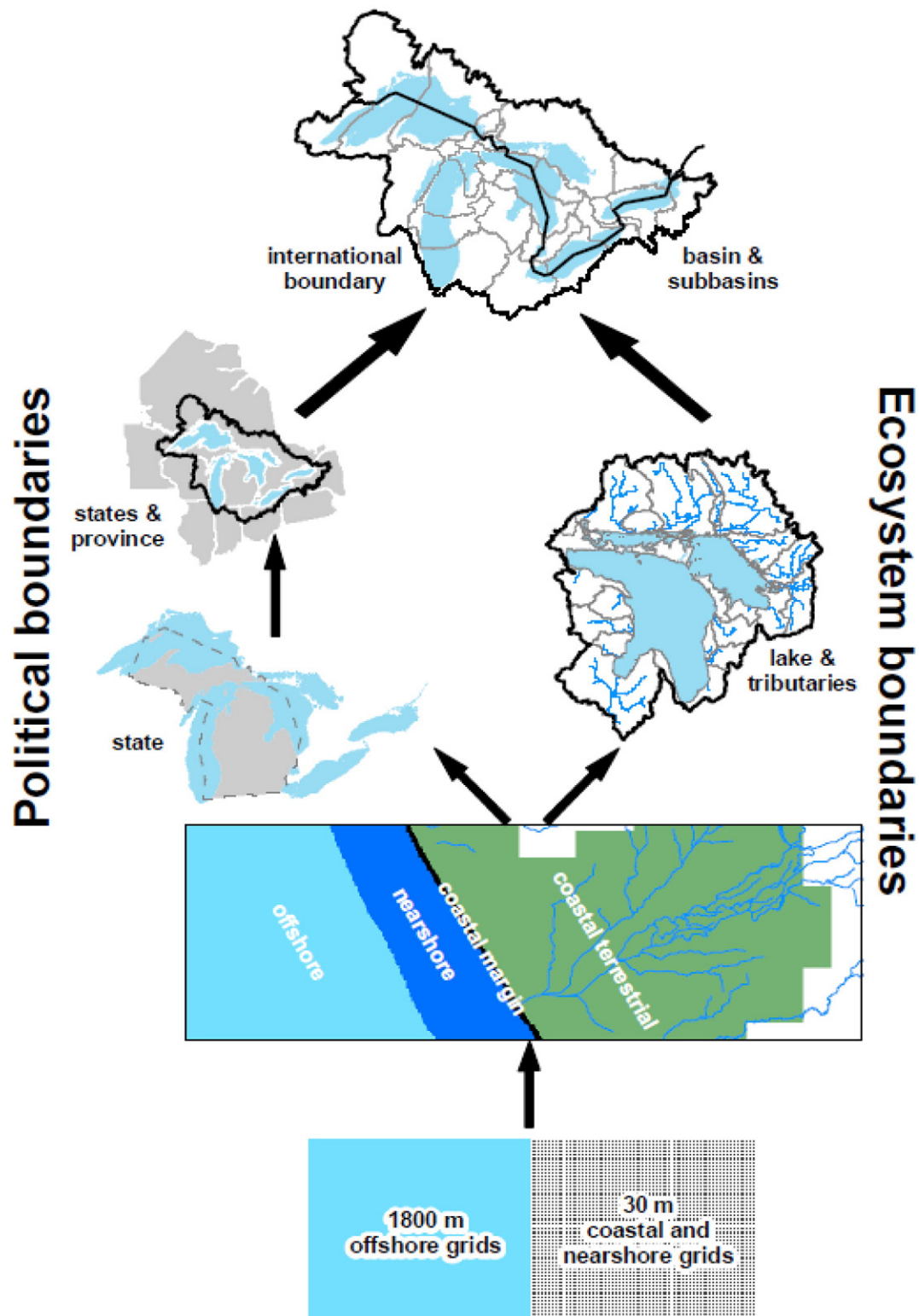
Fig. 5. Great Lakes information synthesis and reporting spatial hierarchy system. The information can be summarized, analyzed, and reported from the grid cells up to the entire Great Lakes Basin and any spatial scales in between. Conversely, the information reported at the Great Lakes Basin scale can be analyzed down to individual grid cells and any spatial scales in between.

Lake characteristics by political boundaries

The U.S. has greater Great Lakes surface water area (64%), coastal wetland area (71%), offshore area (68%), and developed shoreline length (82%) than Canada (Table 1). Canada has more shoreline length than U.S. (67% vs 33%), primarily due to the sinuous shoreline of the north side of Lake Huron (North Channel and Georgian Bay; 1:24,000 scale). The coastline lengths reported here are greater than previously published (Gronewold et al., 2013) because the GLAHF coastline was developed from finer resolution hydrography data (NHD + V2c 2012, OMNR 2013) and included connecting channels and major estuaries (EPA, 1992).

Table 1
Selected Great Lakes ecosystem characteristics summarized by U.S. states and Ontario of Canada.

	IL	IN	MI	MN	NY	OH	PA	WI	ON	Total	US%
Lake characteristics											
Great Lakes surface water area (km^2)	4080	606	100,035	6621	10,495	9233	1955	24,291	88,713	246,028	64
Coastal terrestrial region area (km^2)	356	462	21,214	1497	3922	2169	424	4959	27,786	62,789	56
Coastal aquatic region area (km^2)	1500	606	25,378	835	3180	9233	1535	6529	36,532	85,328	57
Coastal wetland area (km^2)	12	3	1110	1	115	75	4	194	576	2090	72
Offshore region area (km^2)	2580	0	74,806	5786	7346	0	419	17,762	52,127	160,827	68
Length of shoreline (km)	150	104	6321	397	1128	748	149	1520	20,919	31,436	33
Length of beach (km)	40	31	2347	17	203	142	37	319	2171	5307	59
Length of developed shoreline (km)	83	55	501	46	120	182	20	205	267	1480	82
Catchment characteristics											
Total catchment area (km^2)	260	9067	150,297	15,912	38,491	30,164	1571	44,457	236,893	527,112	55
Total length of rivers (km) 100 k	82	5610	88,674	9494	32,989	28,587	1276	27,872	164,150	358,733	54
Catchment − agriculture (km^2)	3	5379	39,415	420	13,893	19,210	519	14,881	46,466	140,187	67
Catchment − urban (km^2)	211	1790	16,362	636	3311	5179	246	3794	5902	37,431	84
Catchment − forest (km^2)	14	872	55,879	9543	16,245	4327	716	15,507	144,573	247,675	42
Catchment − wetland (km^2)	18	720	28,650	5750	3447	693	5775	8460	6291	54,087	88
Catchment − grassland & herbaceous (km^2)	3	157	5615	56	304	453	24	267	9255	16,134	43
Catchment − water (km^2)	10	148	3472	644	1208	278	4	1509	22,934	30,208	24
Population (1000 persons)	696	1329	9883	219	3468	4016	243	2799	11,168	33,822	67
Number of wastewater treatment plants	0	55	370	22	199	274	12	183	351	1466	76
Highway length (km)	2961	32,050	329,370	20,873	89,042	110,239	5131	91,055	158,522	839,243	81

Among the states and province, Michigan (41%) and Ontario (36%) have the highest percentages of surface water area compared with the other states (Wisconsin ~10%, remaining states <5%). Michigan (52%, 47%, 20%) and Ontario (29%, 32%, 67%) also have the most coastal wetlands, offshore water areas, and shoreline length. Although Michigan and Ontario have the highest length of developed shoreline of the Great Lakes, Illinois (55%), Indiana (52%), and Ohio (24%) have the highest percentage of developed shoreline relative to their state or provincial shoreline length.

Catchment characteristics by political boundaries

The U.S. has slightly more catchment areas (55%) and river lengths (54%) than Canada (Table 1). Michigan has more catchment areas and tributary lengths (29%, 25%) than does New York (7%, 9%), Wisconsin (8%, 8%), or Ohio (6%, 8%). In contrast, Indiana, Illinois, Minnesota, and Pennsylvania together have less than 5% of total catchment areas or tributary lengths of the Great Lakes.

The U.S. has more agricultural (67%) and urban (84%) lands in catchments than Canada. The proportion of agricultural land was highest in Ontario (33%) and Michigan (28%) catchments. Urban land use was highest in Michigan (44%), followed by Ontario (16%), Ohio (14%), Wisconsin (10%), and New York (9%), and lowest (<5%) in the other states. About 33.8 million people inhabit the Great Lakes ecosystem (Table 1), with 67% in the U.S. and 29% in Michigan. Based on a preliminary survey (Arvai et al., 2014), there were 1466 wastewater treatment plants in the Great Lakes ecosystem, of which 1115 were in the U.S. and 351 in Canada. In 2010–2011, there were about 839,200 km of highways in the Great Lakes ecosystem and 81% was in the U.S.

Lake characteristics by lakes

Surface area variation represents the major distinguishing characteristics of the Great Lakes. Lakes Superior, Huron, and Michigan account for the majority of the surface area of the Great Lakes (33%, 24%, and 24%) and lakes Erie and Ontario together accounted for less than 19% (Table 2). Offshore zone areas are largest in Lake Superior (43%) and smallest in Lake Erie (1%), but the trend is the opposite for coastal margin and offshore zones (13% vs 30%). Lake Huron has the longest (15,365 km) and Lake Ontario has the shortest (2648 km) shorelines, while Lake Michigan has the longest (1696 km) and Lake Ontario has the shortest (352 km) total length of beaches.

Wave and thermal conditions also vary considerably among the Great Lakes (Table 2). The mean wave height and duration, summarized over the growing season when coastal areas are more vulnerable to wave erosion, was 0.5 m and 3.2 s, respectively for all lakes. Lake Huron had the highest wave height (0.8 m) and longest wave duration (4.0 s), while Lake Erie had the smallest wave height (0.4 m) and shortest wave duration (2.7 s). Lake Superior had the deepest (19.2 m) and Lake Ontario had shallowest (15.0 m) August thermocline depth. Similarly, Lake Superior had the lowest (2252 degree-days) and Lake Erie had the highest (4028 degree-days) degree-days.

The five Great Lakes experienced different thermal characteristic changes during the past 18 years (Table 2). The difference in average growing season (May 1–October 31) surface water temperatures between 1995–1999 and 2008–2012 was greatest for lakes Superior (1.6 °C) and Ontario (1.3 °C) and least for lakes Huron (1.1 °C) and Michigan (1.0 °C). The mean spring warming rate (surface water temperature increase from May–June) averaged over 5-year periods (from 1995 to 1999 and from 2008 to 2012) increased the most for Lake Ontario (0.4 °C/10 days) and the least for lakes Huron and Michigan (0.2 °C/10 days). However, the relative change in spring warming between these two periods was the most dramatic for Lake Superior (91%) and least for Lake Erie (21%). An index of five-year mean ice cover, estimated from sequential days with over 30% coverage, decreased by 8.2 days from 1995-1999 to 2008-2012 for all the Great Lakes considered together, but declined most for lakes Superior (18.2 days) and least for Lakes Ontario (0.8 days).

Catchment characteristics by lakes

The catchment areas, as proportion of total Great Lakes basin catchment, for lakes Superior (26%), Huron (25%), and Michigan (23%) are much larger than that of lakes Erie (14%) and Ontario (12%, Table 2). In contrast, the catchment area per km^2 of water surface area for lakes Superior (1.7), Huron (2.3), and Michigan (2.1) are much smaller than those for lakes Erie (2.8) and Ontario (3.3).

The total proportions of agricultural land in catchments are greater for lakes Ontario (34%) and Erie (28%) than for lakes Michigan (20%), Huron (17%), and Superior (1%, Table 2). The agricultural land per km^2 of water surface area is much greater for lakes Ontario (3.07) and Erie (1.80) than for lakes Michigan (0.61), Huron (0.49), and Superior (0.02). Total proportions of urban land in the catchment are higher for lakes Michigan (32%) and Erie (31%) than for lakes Huron (19%), Ontario (13%), and Superior (5%). The density of urban land area per km^2 of water surface area is higher for Lake Erie (0.43) than for Lake Superior (0.02).

Among the 359,085 km of tributaries in the Great Lakes Basin, about 50% of the length is distributed in the catchments of lakes Huron and Superior (each 25%) and the other 50% are in the catchments of the other

Table 2
Selected Great lakes ecosystem characteristics summarized at a lake scale.

Great Lakes	Erie	Huron	Michigan	Ontario	Superior	All lakes
Lake characteristics						
Surface water area (km^2)	27,033	59,699	57,879	19,262	82,139	246,013
Coastal terrestrial area (km^2)	8543	18,610	12,216	7773	14,948	62,090
Coastal aquatic area (km^2)	25,317	23,804	18,018	5437	10,971	83,547
Coastal wetlands area (km^2)	385	717	467	234	278	2081
Offshore area (km^2)	1934	36,090	39,923	13,595	69,284	160,827
Length of shoreline (km)	4372	15,365	2980	2648	5948	31,312
Length of beach (km)	532	1507	1696	352	1155	5242
Mean depth (m)	19	57	84	87	150	92
Maximum depth (m)	63	224	275	245	405	405
Area with depth < 3 m (%)	3.7	7.3	1.4	5.3	2.0	3.6
Area with depth 3–30 m (%)	87.8	26.2	25.0	19.4	7.6	29.1
Area with depth > 30 m (%)	8.5	66.5	73.5	75.3	90.4	54.1
Mean June 1–Nov 30 wave height (m)	0.4	0.8	0.5	0.5	0.5	0.5
Mean June 1–Nov 30 wave duration (s)	2.7	4.0	3.0	3.4	3.1	3.2
Mean August thermocline depth (m) 1996–2010	18.0	17.4	18.3	15.0	19.2	17.6
Mean August temperature (°C) at thermocline 1996–2010	15.1	12.7	12.4	13.2	8.9	12.4
Mean spring (May–June) rate of warming (°C/day) 1995–2012	0.17	0.09	0.09	0.10	0.05	0.10
Mean spring (May–June) rate of warming (°C/day) 1995–1999	0.14	0.07	0.07	0.08	0.02	0.08
Mean spring (May–June) rate of warming (°C/day) 2008–2012	0.17	0.09	0.09	0.12	0.05	0.11
Mean surface temperature May 1–Oct 31 (°C) 1995–2012	18.4	14.0	15.1	15.6	10.0	14.6
Mean surface temperature May 1–Oct 31 (°C) 1995–1999	18.2	13.7	14.8	15.5	9.4	14.3
Mean surface temperature May 1–Oct 31 (°C) 2008–2012	19.1	14.8	15.8	16.8	11.0	15.5
Mean 0–20 m cumulative degree-days (2008–2012)	4028	3109	3414	3357	2252	3232
Mean lake level (m, IGLD85) 1995–2012	174.19	176.20	176.20	74.79	183.25	NA
Mean lake level (m, IGLD85) 1995–1999	174.41	176.62	176.62	74.81	183.42	NA
Mean lake level (m, IGLD85) 2008–2012	174.17	176.07	176.07	74.78	183.16	NA
Days > 30% lake area covered by ice (1995–1999)	45.8	39.4	10.0	1.0	30.4	25.3
Days > 30% lake area covered by ice (2008–2012)	43.6	26.8	3.0	0.2	12.2	17.2
Substrate clay (%)	8.7	22.1	32.5	18.5	34.8	NA
Substrate silt (%)	1.8	0.0	24.2	0.0	0.0	NA
Substrate mud (%)	59.6	33.0	3.0	52.8	41.0	NA
Substrate sand (%)	26.0	22.5	35.3	4.2	1.2	NA
Substrate hard (%)	0.0	22.4	5.0	24.6	15.7	NA
Substrate unknown (%)	0.7	0.0	0.0	0.0	7.3	NA
Catchment characteristics						
Total catchment area (km^2)	74,843	134,375	120,609	63,305	141,715	534,846
Land use − agriculture (km^2)	48,693	29,252	35,570	59,155	1529	174,200
Land use − urban (km^2)	11,528	7139	12,181	4788	1987	37,622
Land use − forest (km^2)	12,420	74,068	37,165	26,389	98,546	248,589
Land use − wetland (km^2)	2619	9964	23,186	4867	9851	50,487
Land use − grassland (km^2)	716	4050	3410	579	7857	16,612
Land use − water (km^2)	926	10,049	3450	2828	14,476	31,729
Number of wastewater treatment plants	532	247	361	247	81	1468
Total length of rivers (km) (100 k)	66,363	88,564	61,853	51,875	90,430	359,085
Length of river accessible from Great Lakes (km)	30,713	28,502	16,211	11,900	38,189	125,515
Costal terrestrial zone land use and population						
Coastal agriculture (%)	49.2	21.4	20.5	48.1	0.0	23.9
Coastal urban (%)	25.5	10.1	20.6	18.0	1.9	14.7
Coastal forest (%)	15.2	49.5	29.5	23.1	5.6	46.2
Coastal wetlands (%)	6.1	11.0	21.4	6.9	81.1	30.5
Coastal grassland (%)	0.9	1.9	4.3	2.2	9.7	4.5
Coastal water (%)	3.3	6.0	3.6	1.8	1.7	4.8
Coastal population (#/km^2)	320	26	183	408	19	148.3
Aquatic zone mean (2008–2012) chlorophyll-a (µg/L)						
Coastal margin zone (0–3 m)	14.7	4.3	4.6	6.0	1.9	6.3
Nearshore zone (3–30 m or 3–15 m)	12.4	2.2	1.5	2.4	1.2	3.9
Offshore zone (>30 m)	7.6	0.8	0.6	1.2	0.6	2.2

lakes (each 14–18%, Table 2). About 33% of the tributary lengths have no dams known to substantially block access from Great Lakes to tributaries. Lake Superior has the most length of accessible tributaries (30%), followed by lakes Huron (24%), Erie (21%), Michigan (16%), and Ontario (9%).

Lake and catchment characteristics by classification zones

Coastal terrestrial zone

Among the five Great Lakes, lakes Erie and Ontario have the highest (49% and 48%), while Lake Superior has the lowest percentages of

agricultural land in the coastal terrestrial zone (Table 2). Lake Erie also has the highest percentage of urban land (26%), followed by lakes Ontario (21%), Michigan (18%), and Superior (5%). In contrast, Lake Ontario has the highest ($164/km^2$) and Lake Erie has the second highest ($101/km^2$) human population densities, while Lake Huron ($8/km^2$) and Superior ($3/km^2$) have the lowest.

Coastal margin, nearshore, and offshore zones

Although many factors in our database can be used to describe the characteristics of the classification zones, we chose chlorophyll *a* to indicate the difference in primary productivities among those zones.

Among the five Great Lakes, Lake Erie coastal margin zone had the highest (15 μg/L) chlorophyll *a* concentrations, Lake Ontario coastal margin had the second highest (6 μg/L), and that of the other three lakes had the lowest (2–5 μg/L) chlorophyll *a* concentrations (Table 2). Similarly, Lake Erie nearshore zone had much higher chlorophyll *a* concentration (12.4 μg/L) than the other lakes (1.2–2.4 μg/L). Lake Erie offshore zone also had much higher chlorophyll *a* concentration (7.7 μg/L) than the other lakes (0.6–1.2 μg/L). Among the zones for all five lakes, the coastal margin zone had the highest, the nearshore zone has the next highest, and the offshore zone had the lowest chlorophyll *a* concentrations (Table 2).

Discussion

Value of the hierarchical spatial classification framework

Our hierarchical spatial classification framework helps resolve several challenges in data and information management and delivery of the Great Lakes. First, it provides the basic spatial units for data attribution that facilitates incorporation of data from multiple sources into a common spatial network, and integrates data across data types, scales, and ecosystem and political boundaries. Each of the five Great Lakes is a large (surface area and volume) and complex ecosystem with different physicochemical and biological characteristics that naturally vary within and among lakes. The feature of our framework can potentially be used to answer how big an area that a sampling site can represent, how such data can be synthesized to represent larger spatial units, and how the variation within and among lake sub–basins can be properly measured.

Second, it provides a mechanism to quantify connectivity among spatial units and their positions within the system. This feature is extremely important because the natural physicochemical and biological habitat characteristics of a spatial unit are largely determined by the bathymetry, current, waves, geomorphic, and climate conditions at the spatial scales of interest. Those habitat characteristics are also strongly influenced by their connectivity with major system features, such as distance to and condition/size of river mouths, wetlands, shipping routes, point pollution sources, and preserved areas. The importance of connectivity in riverscape ecology has been well established (e.g., Frissell et al., 1986; Poff, 1997; Allan et al., 2004; Wang et al., 2006, 2012) and GLAHF includes multiple measures of connectivity related to Great Lakes habitats and organisms. These features enable ecological modeling and forecasting of changes to Great Lakes food webs and fisheries resulting from stressors and/or management actions. The ability to forecast changes allows management to move from a reactive to a more proactive state.

Last, our hierarchical spatial classification framework and database provides a mechanism for synthesizing data and information at local, regional, and basin-wide scales, or political jurisdictions. It enables managers and researchers to access data summarized for different spatial units to better understand, conserve, and restore the lake ecosystems.

Uses of the spatial system and database

One of the important uses of our spatial system and database is that the data attributed at the smallest spatial units (30-m or 1.8-km grids) can be synthesized into information at multiple ecological, management, and political units for policy making and management planning and implementation. These units may include a lake sub-basin, a lake, or the entire five-lake ecosystem; lake sections within a specific local conservation area, a planning district, a state or province, or a multi-state/provincial region; or the entire Great Lakes basin in the U.S. and Canada. Such information can meet the needs of local stakeholders who are interested in only the water quality of their beaches, coastal line protection, drinking water quality, or harbor operations; local governments and planners who are managing cities or districts; state/provincial governments who are responsible for a lake area within their political boundaries; partnerships who have common interests in protection and rehabilitation of a particular lake or all the five lakes; and the governments of U.S. and Canada who are responsible for implementing the GLWQA.

Another important use of our database is to classify the spatial ecological units into habitat classes. Classifying spatial units and establishing physicochemical and biological expectations are critical for water resource policy-making, regulation, and scientific hypothesis testing. Presently, such classifications and expectations have been well developed for inland lakes (e.g., Cheruvelil et al., 2008; Wang et al., 2010; Wehrly et al., 2012) and river segments (e.g., Brenden et al., 2008; Wang et al., 2012), but for the Great Lakes it has been developed only for specific resources or uses, such as Quaternary deposits of the Great Lakes (Karrow et al., 2000), coastal wetlands (Albert et al., 2005), fish habitat (McKenna and Castiglione, 2010a,b; Chu et al. 2014), and human disturbances (Allan et al., 2013; Danz et al., 2005). Ecological classifications of Lake Erie were conducted by Johnson et al. (submitted for publication) and a physical habitat classification of Great Lakes nearshore and offshore waters was conducted by Rutherford and Geddes (2007). These efforts have provided classifications of spatial units by type or status based on one or a composite set of variables. The GLAHF framework has been structured to allow multiple user-defined classifications using specific sets of variables such as those described above based on an objectively defined set of aquatic ecological units. The next phase of GLAHF will be the development of these core or base units of classification within the hierarchical spatial classification framework. Aquatic ecological units (AEUs) will be formed by merging neighboring data-attribution units of the 30-m or 1.8-km grid cells into AEUs that have relatively homogenous physicochemical and biological characteristics (Fig. 2). These AEUs are the building blocks for habitat classification of the Great Lakes system, which are nested within the coastal margin, nearshore, and offshore zones (Fig. 2). Such a process allows for the classification and mapping of ecological units for multiple uses or purposes across the entire binational Great Lakes ecosystem.

An important potential use of the database is to serve the needs of managers and policy makers trying to achieve many of the GLWQA objectives. The hierarchical structure of the database and spatial framework will facilitate the assembling of critical information needed for developing and implementing lake-wide management plans, and tracking the progress of implementing the plans at local, lake sub-basin, lake basin, and Great Lakes basin-wide scales. The GLAHF database can provide data needed to quantify beneficial use impairments and locations, track distributions and invasion rates of aquatic non-indigenous species, and identify amounts and sources of point and non-point source pollutants. Our database provides a spatial classification framework that is needed for overall assessment of condition of the nearshore waters, for identifying nearshore areas that may become subject to human disturbance stress and areas of high ecological value, and for establishing priorities of protection and restoration using local and large scale factors. Our hierarchical classification framework and database provides the essential mapping tools and information for conducting a baseline habitat survey, assessing net habitat gain, and developing and implementing lake-wide habitat and species conservation strategies. The GLAHF spatial framework and database contains the data needed for calculating major indicators and assessing status, trends, and progress toward achieving the general and specific objectives of the GLWQA (GLWQA, 2012). It also provides a Great Lakes system-wide tool for the development and implementation of coordinated data accounting, management, sharing, and delivery strategies.

Although spatial structure of GLAHF was subjectively decided by a bi-national team of resource managers and scientists, it was reviewed by many scientists, managers, and policy makers in the Great Lakes region to ensure each spatial zone is distinct from the other zones and is useful for management and research. It is built on current knowledge

to ensure that the zones reflect the lake ecosystem spatial components and functionality. The 5-km landward coastal terrestrial zone (including river mouths and catchments above lake coastal line) was delineated to capture the more important roles of natural and anthropogenic environment adjacent to the lakes compared to those at distance (Hollenhorst et al., 200; Wang et al, 2010; Wehrly et al, 2013). Although the 30-m cell size at the coastal terrestrial, coastal aquatic margin, and coastal nearshore zones and the 1.8-km cell size at the offshore zone meet the majority data resolution needs and is an efficient way to manage the Great Lakes Basin-wide data, at times the users may wish to incorporate their application at finer spatial resolution. In such a case, the users can use data in GLAHF and further divide the cells into smaller sizes.

Limitations and future management of GLAHF

The intent of GLAHF database was not to include all data collected from the Great Lakes. The goal of GLAHF spatial database has been the integration of available basin-wide data to address the needs for data rectified to a common spatial framework. The outcome is a spatial database of consistent data available at the Great Lakes Basin scale and minimally at a lake system scale. For example, we did not attempt to assemble and harmonize data that were collected for specific objectives at a local scale. For local users, our database provides a data backbone (spatial hierarchical framework) to link localized data with the readily available broad scale data (e.g., bathymetry, temperature, wave, substrate, and ecological unit classifications, etc.) to meet their own needs. Because the spatial resolution, type, and availability of lake system-wide data could vary depending on data source, our data attribution process has simplified some of the data that were inconsistent across the scales into common descriptors or common spatial units. Hence, some of the data in our database may have lost their fine detail or resolution.

Some of the data currently in our database are incomplete and can be improved in the near future. Nearshore areas of the Great Lakes have only recently been sampled and studied basin-wide and these data need to be incorporated into GLAHF when they become available. The ballast water loading and discharge locations in the Great Lakes are available only for the United States portion of the waters. The coastline typology for the U.S. is much newer and more detailed than those for the Canada, which will need improvement when newer data become available. The approach of the first iteration of GLAHF is to include as much of the best available data on a lake/system-wide scale as possible. Given the fast development of information technology and increasingly availability of regional data, our database will require periodic update and improvement.

The GLAHF database is extensive (100 s GBs), which could be cumbersome and difficult to manage. The advancement in faster computing power, larger hard drives, the advent cloud computing, and large database operation systems (e.g., SQL Server, ArcSDE) make the management of our database more efficient and easier. The GLAHF spatial data is stored in the ESRI proprietary format. Although this format requires specific software licenses and continued file maintenance and updating, it is the most commonly used GIS platform to facilitate multiple users. Data in this format can be stored in a compressed format and can be exported in open source spatial data format standards (e.g., shapefile, csv, kml, GeoTIFF).

The GLAHF database and spatial framework have been developed through a research project with the goal of developing a publically accessible database and classification system. The plans for long-term database management, data delivery, and data update were beyond the scope of this research project but are currently under development. In collaboration with Institute for Fisheries Research, University of Michigan, and U.S. Geological Survey Great Lakes Science Center, we have begun the work toward this direction.

Overall, our database provides the essential data for achieving some of the GLWQA objectives and for meeting the needs of many other management and research programs. The utility of the database can be improved by incorporating additional detailed localized data that are not available at a lake or system-wide scale. Presently, many additional data layers and data layers at a higher resolution are available only at a regional or local scale. Adding of these data to our lake system-wide database by regional agencies or partnerships will improve the utility of GLAHF for local users by providing information that otherwise could not be supplied by the lake system-wide database and by placing that data within a lake- or basin-wide framework. This database also can be improved by incorporating updated or new data layers, newly designed systematic collected information, and modeled physicochemical and biological conditions under projected climate changes.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jglr.2015.03.017.

References

Albert, D.A., Wilcox, D.A., Ingram, J.W., Thompson, T.A., 2005. Hydrogeomorphic classification for great lakes coastal wetlands. J. Great Lakes Res. 31 (Supp. 1), 129–146.

Allan, J.D., McIntyre, P.B., Smith, S.D.P., et al., 2004. Landscape and riverscapes: the influence of land use on river ecosystems. Annu. Rev. Ecol. Evol. Syst. 35, 257–284.

Allan, J.D., et al., 2013. Joint analysis of stressors and ecosystem services to enhance restoration effectiveness. Proc. Natl. Acad. Sci. U. S. A. 110 (1), 372–377.

Arvai, A., Klecka, G., Jasim, S., Melcer, H., Laitta, M.T., 2014. Protecting our great lakes: assessing the effectiveness of wastewater treatments for the removal of chemicals of emerging concern. Water Qual. Res. J. Can. 49 (1), 23–31.

Barbiero, R.P., Lesht, B.P., Warren, G.J., 2012. Convergence of trophic state and the lower food web in Lakes Huron, Michigan and Superior. J. Great Lakes Res. 38 (2), 368–380.

Barton, C.R., Schelske, C.L., 1982. Lake-wide seasonal changes in limnological condition s in Lake Michigan in 1976. J. Great Lakes Res. 8 (3), 413–427.

Beletsky, D., Schwab, D., McCormick, M., 2006. Modeling the 1998–2003 summer circulation and thermal structure in Lake Michigan. J. Geophys. Res. 111, C10010. http://dx.doi.org/10.1029/2005JC003222.

Bertrand, G., Lang, J., Ross, J., 1976. The Green Bay Watershed Past/Present/Future. Technical Report No. 229. University of Wisconsin Sea Grant Program, Madison, WI (300 pp.).

Boase, J.C., 2003. Integrating Sonic Tracking and GIS to Determine Habitat Selection and Benthic Prey Distribution of Adult Lake Sturgeon in Lake St. Clair. (Master's thesis). University of Michigan, Ann Arbor.

Brenden, T., Wang, L., Seelbach, P.W., 2008. A river valley segment classification of Michigan streams based on fish and physical attributes. Trans. Am. Fish. Soc. 137, 1621–1636.

Bunnell, D.B., Barbiero, R.P., Ludsin, S.A., et al., 2013. Changing ecosystem dynamics in the Laurentian Great Lakes; bottom-up and top-down regulation. Bioscience 64 (1), 26–39.

Cha, Y.K., Stow, C.A., Nalepa, T.F., Reckhow, K.H., 2011. Do invasive mussels restrict offshore phosphorus transport in Lake Huron? Environ. Sci. Technol. 45, 7226–7231.

Cheruvelil, K.S., Soranno, P.A., Bremigan, M.T., Wagner, T., Martin, S.L., 2008. Grouping Lakes for water quality assessment and monitoring: the roles of regionalization and spatial scale. Environ. Manag. 41, 425–440.

Chu, C., Koops, M.A., Randall, R.G., Kraus, D., Doka, S.E., 2014. Linking the land and the lake: a fish habitat classification for the nearshore zone of Lake Ontario. Freshwater Science. 33, 1159–1173.

Cline, T.J., Bennington, V., Kitchell, J.F., 2013. Climate change expands the spatial extent and duration of preferred thermal habitat for Lake Superior Fishes. PLoS One 8 (4), e62279. http://dx.doi.org/10.1371/journal.pone.0062279.

Clites, A.H., Smith, J.P., Hunter, T.S., Gronewold, A.D., 2014. Visualizing relationships between hydrology, climate, and water level fluctuations on Earth's largest system of lakes. J. Great Lakes Res. 40, 807–811.

Danielle K.F., Riseng C.M., Wehrly K.E. Mason L., Gaiot J., Hollenhorst T., Johnston C., Wyrzykowski C., Annis G., Castiglione C., Todd K, Robertson M., Infante D.M., Wang L, McKenna J.K., Whelan G. A consistent, binationally harmonized watershed boundary dataset – the great lakes hydrographic dataset (submitted for publication, April 1, 2015).

Danz, N.P., Regal, R.R., Niemi, G.J., et al., 2005. Environmentally stratified sampling design for the development of Great Lakes environmental indicators. Environ. Monit. Assess. 102, 41–65.

Dell, C.I., 1976. Sediment distribution and bottom topography of southeastern Lake Superior. J. Great Lakes Res. 2 (1), 164–176.

Dietrich, G., Kalle, K., Krauss, W., Siedler, G., 1980. General Oceanography. Translated by Susanne and Hans Ulrich Roll. 2nd ed. John Wiley and Sons (Wiley-Interscience), New York.

EPA (Environmental Protection Agency), 1992. Great Lakes Monitoring and Research Strategy Environmental Monitoring and Assessment Program. Environmental Research Laboratory, Office of Research and Development. United States Environmental Protection Agency, Duluth, MN, p. 55804 (June 1992. EPA/620/R-92/001).

Eyles, N., Mohr, L., 2006. Lake trout/lake whitefish spawning shoal assessment: Georgian Bay. Ontario Ministry of Natural Resources and Office of Research Services, University of Toronto Title: Two Year (2004–6) Collaborative Ontario Ministry of Natural Resources And University of Toronto Project, Project Number: 300771 Final Report: May 2006. Ontario Ministry of Natural Resources, and Office of Research Services, University of Toronto (submitted for publication).

Fahnenstiel, G., Pothoven, S., Vanderploeg, H., Klarer, D., Nalepa, T., Scavia, D., 2010. Recent changes in primary production and phytoplankton in the offshore region of southeastern Lake Michigan. J. Great Lakes Res. 36, 20–29.

Frissell, C.A., Liss, W.J., Warren, C.E., Hurley, M.D., 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environ. Manag. 10, 199–214.

Gallant, A.L., Brown, D.D., Hoffer, R.M., 2005. Automated mapping of Hammond's landforms. IEEE Geosci. Remote Sens. Lett. 2 (4) (October).

GLWQA (Great Lakes Water Quality Agreement), 2012. Great Lakes Water Quality Agreement, Protocol Amending the Agreement between Canada and the United States of America on Great lakes Water Quality, 1978 (as Amended on October 16,1983, and on November 18, 1987. Signed September 7, 2012 and entered into force February 12, 2013).

Goodyear, C.D., Edsall, T.A., Dempsey, D.M., Moss, G.D., Polanski, P.E., 1982. Atlas of the Spawning and Nursery Areas of Great Lakes Fishes vol. 9. U.S. Fish and Wildlife Service, Office of Biological Services.

Gronewold, A.D., Fortin, V., Lofgren, B., Clites, A., Stow, C.A., Quinn, S.F., 2013. Coasts, water levels, and climate change: a Great Lakes perspective. Climate Change 20, 697–711.

Hollenhorst, T.P., Brown, T.N., Johnson, L.B., Ciborowski, J.J.H., Host, G.E., 2007. Methods for generating multi-scale watershed delineations for indicator development in Great Lake coastal ecosystems. J. Great Lakes Res. 33, 13–26.

Howell, E.T., Chomicki, K.M., Kaltenecker, G., 2012. Tributary discharge, lake circulation and lake biology as drivers of water quality in the Canadian nearshore of Lake Ontario. J. Great Lakes Res. 38 (Suppl. 4), 47–61.

Johnson, L., Ciborowski, J., Mackey, S., Hollenhorst, T., Gauthier, R., Button, D.T., 2007. An Integrated Habitat Classification and Map of the Lake Erie Basin: Final Report. National Fish and Wildlife Foundation (submitted for publication, November 30, 2007).

Jude, D.J., Pappas, J., 1992. Fish Utilization of Great Lakes Coastal Wetlands.

Karrow, P.F., Dreimanis, A., Barnett, P.J., 2000. A proposed diachronic revision of late Quaternary time-stratigraphic classification in the eastern and northern Great Lakes area. Quat. Res. 54, 1–12.

LEHTG (Lake Erie Habitat Task Group), 2012. Report of the Lake Erie Habitat Task Group, Presented to the Standing Technical Committee, Lake Erie Committee. Great Lakes Fishery Commission (March 23, 2012).

Lewis, M., 1966. Sedimentation Studies of Unconsolidated Deposits in the Lake Erie Basin. (Ph.D. thesis). University of Toronto.

McKenna Jr., J.E., Castiglione, C., 2010a. Hierarchical multi-scale classification of nearshore aquatic habitats of the Great Lakes: Western Lake Erie. J. Great Lakes Res. 36, 757–771.

McKenna Jr., J.E., Castiglione, C., 2010b. Hierarchical multi-scale classification of nearshore aquatic habitats of the Great Lakes: Western Lake Erie. J. Great Lakes Res. 36 (4), 757–771.

Michalak, A.M., et al., 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proc. Natl. Acad. Sci. U. S. A. 110 (16), 6448–6452.

Moore, J.R., Meyer, R.P., 1969. Progress report on the geological–geophysical survey of Green Bay, 1968. Technical Report No. 1. University of Wisconsin Sea Grant Program, Madison, WI (16 pp.).

NHD Plus V2, 2012. National Hydrography Dataset Plus, Version 2. United States Geological Survey and United States Environmental Protection Agency (June 2012. (Last accessed January 10, 2014: http://www.horizon-systems.com/NHDPlus/index.php)).

Niemi, G.J., Kelly, J.R., Danz, N.P., 2007. Environmental indicators for the coastal region of the North American Great Lakes: introduction and prospectus. J. Great Lakes Res. 33 (Special Issue 3), 1–12.

OMNR (Ontario Ministry of Natural Resources), 2013. Ontario Integrated Hydrology Data: Elevation and Mapped Water Features for Provincial Scale Hydrology Applications (Technical Release) V. 1.1. Spatial Data Infrastructure (formerly Water Resources Information Program). Ontario Ministry of Natural Resources (pp. 32–33).

Plattner, S., Mason, D.M., Leshkevich, G.A., Schwab, D.J., Rutherford, E.S., 2006. Classifying and forecasting coastal upwellings in Lake Michigan using satellite derived temperature images and buoy data. J. Great Lakes Res. 32, 63–76.

Poff, N.L., 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. J. N. Am. Benthol. Soc. 16, 391–409.

Powers, C.F., Robertson, A., 1968. Subdivisions of the benthic environment of the Upper Great Lakes, with emphasis on Lake Michigan. J. Fish. Res. Board Can. 25 (6), 1181–1197.

Rao, Y.R., Schwab, D.J., 2007. Transport and mixing between the coastal and offshore water in the Great Lakes: a review. J. Great Lakes Res. 33, 202–218.

Rasul, N., Coakley, J.P., Pippert, R., 1997. Sedimentary environment of Lake Erie: geologic setting, sediment distribution, and modern evolutionary trends. Canada National Water Research Institute Contributionpp. 97–114.

Riseng, C.M., Wang, L., Wiley, M.J., Rutherford, E., Brenden, T., 2008. State-of-the-art approaches for assessment of Great Lakes nearshore and large river fish habitat. Final Report to the Great Lakes Fishery Trust, United States Environmental Protection Agency, and Great Lakes Fishery Commission.

Robbins, J.A., 1986. Sediments of Saginaw Bay, Lake Huron: elemental composition and accumulation rates. Great Lakes Research Division, Special Report No. 102. USEPA, Environmental Research Laboratory-Duluth, Duluth, Minnesota.

Rutherford, E.S., Geddes, C.A., 2007. Ecological Classification of Nearshore and Open-Water Fish Habitats in the Great Lakes. Final Report to Michigan Department of Natural Resources. State Wildlife Grant Program. Institute for Fisheries Research, Ann Arbor, Michigan.

Schwab, D.J., Beletsky, D., DePinto, J., Dolan, D.M., 2009. A hydrodynamic approach to modeling phosphorus distribution in Lake Erie. J. Great Lakes Res. 35, 50–60.

Statistics Canada, 2002. Population Data for Ontario From 1929 to 2001. Provided by Ontario Ministry of Finance. Ontario, Toronto.

Thomas, R.L., 1988. Distribution and composition of the surficial sediments of Georgian Bay and North Channel. Hydrobiologia 163, 35–45.

Thomas, R.L., Dell, C.I., 1978. Sediments of Lake Superior. J. Great Lakes Res. 4 (3–4), 264–275.

Thomas, R.L., Kemp, A.L.W., Lewis, C.F.M., 1972. Distribution, composition and characteristics of the surficial sediments of Lake Ontario. J. Sediment. Res. 42, 66–84.

Thomas, R.L., Kemp, A.L.W., Lewis, C.F.M., 1973. The surficial sediments of Lake Huron. Can. J. Earth Sci. 10, 226–271.

U.S. Census Bureau, 2000. 2000 Census. Department of Commerce, Washington, D.C.

USEPA (Environmental Protection Agency) and Government of Canada, 1995. The Great Lakes: An Environmental Atlas and Resource Book. Great Lakes National Program Office, US Environmental Protection Agency.

Wang, L., Seelbach, S.W., Hughes, R.M., 2006. Introduction to landscape influences on stream habitats and biological assemblages. Am. Fish. Soc. Symp. 48, 1–23.

Wang, L., Wehrly, K., Breck, J., Kraft, L.S., 2010. Landscape based assessment of human disturbance for Michigan lakes. Environ. Manag. 46, 471–483.

Wang, L., Infante, D., Esselman, P., Cooper, A., Wu, D., Taylor, W., Beard, D., Whelan, G., Ostroff, A., 2011. A hierarchical spatial framework and database for the national river fish habitat condition assessment. Fisheries 36 (9), 436–449.

Wang, L., Brenden, T., Cao, Y., Seelbach, P., 2012. Delineation and validation of river network spatial units for water resources and fisheries management. Environ. Manag. 50, 875–887.

Wegscheider, S., 2006. Investigation of Coastal Upwelling Events in Lakes Michigan, Huron, Superior, and Ontario. Internal Report From the Summer Fellowship Program. NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, MI.

Wehrly, K.E., Breck, J.E., Wang, L., Szabo-Kraft, J., 2012. Classifying fish assemblages in sampled and unsampled lakes: a multivariate regression tree approach. Trans. Am. Fish. Soc. 141, 414–425.

Wehrly, K.E., Wang, L., Infante, D., Cooper, A., Geddes, C., Stanfield, L., Rutherford, E., 2013. Landscape change and its influences on aquatic habitat and fisheries in the Great Lakes Basin. In: Taylor, W., Leonard, N., Lynch, A. (Eds.), Great Lakes Fisheries Policy and Management: A Binational Perspective, 2nd edition Michigan State University Press, East Lansing.

Wei, A., Chow-Fraser, P., Albert, D., 2004. Influence of shoreline features on fish distribution in the Laurentian Great Lakes. Can J Fish Aquat Sci. 61, 1113–1123.

Yurista, P.M., Kelly, J.R., Miller, S.E., 2006. Comparisons of zooplankton community size structure in the Great Lakes. J. Geophys. Res. Oceans 111, C05S08. http://dx.doi.org/10.1029/2005JC002971.