Development of a web application for examining climate data of global lake basins: CGLB

Tosiyuki Nakaegawa1,2, Shinichiro Horiuchi1 and Hyungjun Kim2
1Climate Research Department, Meteorological Research Institute, Tsukuba, Japan
2Institute of Industrial Science, The University of Tokyo, Japan

Abstract:
A newly developed web application, Climates of Global Lake Basins (CGLB), combines existing datasets and interactively displays geographical, hydrological, and climatological information for hundreds of lakes around the world. CGLB also provides photographs containing vegetation information as well as quasi-real time monitoring of lake water levels. CGLB can interactively create and animate time series of climatological data in a one-dimensional or two-dimensional (geographical) form. These functions are useful for education, expedition planning, and scientific research. As an example of the application’s use, links between water levels in Kenya’s Lake Turkana and sea surface temperature (SST) and regional precipitation were tested by time-lag correlation analysis. Precipitation showed no significant correlation with the lake water level for time lags of 0, 1, and 2 months. This suggests that variation in land-surface hydrological processes are key to the interannual variability in water levels in the lake. In contrast, SST in the central tropical Pacific was strongly correlated with the lake water level for all time lags. No significant correlation was found between the lake water level and SST in the Indian Ocean adjacent to Kenya, although significant correlations were found between regional precipitation and SST in the Indian Ocean.

KEYWORDS global lakes; lake information; climate; web application; Lake Turkana; water level

INTRODUCTION

Lakes occupy about 0.5% of the global land area (e.g., Nakaegawa, 2012) and hold more than 90% of the world’s liquid freshwater (Oki and Kanae, 2006) in readily accessible bodies. Because of freshwater extraction or sequestration for human uses, some lakes, such as the Aral Sea (e.g., Micklin, 1988) and Lake Chad (e.g., Coe and Foley, 2001), have experienced disastrous shoreline retreat and water quality deterioration. Satellite observations document the impact of anthropogenic practices on those two lakes. Growing awareness of these undesirable situations, and of the fact that sustainable use of lakes and their basins requires special consideration, has given rise to a methodology known as Integrated Lake Basin Management (International Lake Environment Committee (ILEC), 2007).

There are several global lake databases, each of them having unique features. The World Lake database (WLDB; ILEC, 1999) includes comprehensive information about lakes but lists only about 650 lakes. The LakeNet global lake database (http://www.worldlakes.org/lakes.asp) includes comprehensive information that is especially focused on lake protection and management. The Global Lakes and Wetlands Database (GLWD; Lehner and Döll, 2004) includes information about the geography and geomorphology of 250,000 lakes, including shoreline vector data and 1-km gridded data on water-related land cover types. The Global Lake Database (GLDB; e.g., Kourzeneva et al., 2012; Choula et al., 2014) includes data on the mean depths of slightly more than 13,000 lakes, global 1-km gridded data on mean depths, and bathymetric data.

Lake types, such as glacial lakes and meromictic lakes, are often characterized by climate. Under the Global Framework for Climate Services (World Meteorological Organization, 2014), national meteorological and hydrological services and research institutions have developed several climate analysis tools to respond to public demands for climate information and to promote the use of climate information (Timofeyeva-Livezey et al., 2015), although similar web applications date to the beginning of the century (e.g., Ikoma et al., 2000). GLWD and GLDB do not include climatological data because the databases were developed for different purposes. WLDB includes climatological data at stations near lakes, but for only about one-third of the lakes and for various time spans.

We developed a web application, named Climate of Global Lake Basins (CGLB; http://hydro.iis.u-tokyo.ac.jp/CGLB), to interactively display lake information and climatological information together by integration of existing datasets. CGLB can draw time series of climatological data in geographical as well as one-dimensional form. We demonstrate these functions with an example for Lake Turkana in the East African Rift Valley, the largest permanent desert lake in the world (Velpuri et al., 2012), which displays wide variation in water level on interannual and seasonal timescales (Ricketts and Johnson, 1996).

APPLICATION DESIGN

The primary concept of CGLB is the provision of climatological information for a target lake basin so that members of the public who are interested in the lake basin as well as researchers can obtain information interactively for their individual purposes. Existing datasets such as GLWD and GLDB are provided in binary or text forms preventing most
people from readily using them. WLDB and LakeNet provide the text data in web form but do not provide sufficient climatological variables or an interactive display. Therefore, we focused on the development of a web application with an interactive figure display including climatological variables and user-friendly interfaces. We made the most of existing data and web pages by integrating various existing information relevant to lake basins. We did not yet implement any mathematical operations among variables, such as correlation or the four basic arithmetical operations, to facilitate the rapid use of CGLB as a first version.

The datasets included in CGLB are listed in Table I, including relevant URLs. Table I also provides the type of operations available for each dataset in CGLB. The basic lake data for WLDB are displayed on the lower right of the computer display. Interactive figures are provided for two datasets in CGLB and for two datasets in hyperlink web pages. Only hyperlinks are provided for the other five datasets. Details on each dataset is described in the next section.

**DATA**

**Existing data and web pages**

Our initial step was to compile information about individual lakes. We adopted WLDB as the basis of our global lakes dataset because it is a reference for other lake databases. It is a comprehensive database containing not only hydrological and water resources information but also social and economic information. Although for many lakes WLDB only contains information on geography and geomorphology, it includes 14 data categories such as location, water depth, water quality, water temperature, climate conditions, annual fish catch, and population density. In this study, we used 15 data types from these 14 categories. Three other information sources were added as hyperlinks, including Wikipedia, the Global Reservoir and Lake Monitor (GRLM), and supplemental information from WLDB. Wikipedia is an Internet encyclopedia with open access and free content policies. Everybody who can access the site and follow its rules can edit most of its articles. This allows rapid revision of the contents, and therefore Wikipedia can provide the latest information about the lakes. GRLM provides 10-day lake level variations observed by satellite radar altimeters (Birkett, 1995). A total of 18 data types are included for each lake in this application.

We incorporated a climatological monthly mean dataset into CGLB named CRU developed by the Climate Research Unit, University of East Anglia (Harris et al., 2014). This dataset was constructed from ground-based observations.

<table>
<thead>
<tr>
<th>Name</th>
<th>Items</th>
<th>Number of items</th>
<th>Interactive/ hyperlink</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLDB</td>
<td>Country, state, latitude, longitude, altitude, shoreline length, surface area, volume, maximum depth, minimum depth, water level control, annual water level fluctuation, shoreline, residence time, catchment area</td>
<td>15 displayed on CGLB</td>
<td>ILEC (1999)</td>
<td><a href="http://wldb.ilec.or.jp">http://wldb.ilec.or.jp</a></td>
</tr>
<tr>
<td>Wikipedia</td>
<td>Descriptions</td>
<td>1 hyperlink</td>
<td></td>
<td><a href="http://en.wikipedia.org/wiki/Main_Page">http://en.wikipedia.org/wiki/Main_Page</a></td>
</tr>
<tr>
<td>GRLM</td>
<td>Water level</td>
<td>1 hyperlink</td>
<td>Birkett et al. (2011)</td>
<td><a href="http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/">http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/</a></td>
</tr>
<tr>
<td>CRU</td>
<td>Cloud cover, diurnal temperature range, precipitation, daily mean temperature, monthly average daily minimum temperature, monthly average daily maximum temperature, vapor pressure, wet day frequency, potential evapotranspiration</td>
<td>9 interactive</td>
<td>Harris et al. (2014)</td>
<td><a href="http://www.cru.uea.ac.uk/data">http://www.cru.uea.ac.uk/data</a></td>
</tr>
<tr>
<td>ClimatView</td>
<td>Monthly mean temperature, monthly mean maximum temperature, monthly mean minimum temperature, precipitation, monthly mean temperature anomaly, precipitation ratio, climatological mean temperature, climatological mean precipitation</td>
<td>8 interactive at hyperlink web page</td>
<td>JMA (2014)</td>
<td><a href="http://ds.data.jma.go.jp/gmd/tcc/tcc/products/climate/climatview/frame.php">http://ds.data.jma.go.jp/gmd/tcc/tcc/products/climate/climatview/frame.php</a></td>
</tr>
<tr>
<td>Google Maps</td>
<td>Global map as Web GIS</td>
<td>1 interactive</td>
<td></td>
<td><a href="https://www.google.co.jp/">https://www.google.co.jp/</a></td>
</tr>
<tr>
<td>DCP</td>
<td>In-situ landscape information</td>
<td>1 hyperlink</td>
<td><a href="http://confluence.org/confluence.php">http://confluence.org/confluence.php</a></td>
<td></td>
</tr>
</tbody>
</table>
and interpolated at a horizontal resolution of 0.5°. This dataset only covers those land areas of the globe without missing data. CRU includes nine major hydroclimatic variables, but we added more from other sources. For hydroclimate analyses, atmospheric data are required for running land-surface models (Nakagawa and Sugi, 2001) and for estimating atmospheric water balances. These data were obtained from the Japanese 55-year Reanalysis (JRA-55; Kobayashi et al., 2015) produced by the Japan Meteorological Agency (JMA). This dataset was produced by assimilating observations into the reanalysis system by a 4-dimensional variational method. JRA-55 provides higher-quality and more homogeneous climate data for a variety of meteorological variables that were not available from observations, but it includes some biases in locations where no observations were available. The monthly mean values in this application, obtained by averaging 3-hourly data, cover the entire globe including the oceans, at a horizontal resolution of about 55 km. We incorporated 11 variables from JRA-55.

Monthly in-situ surface observations from meteorological stations can be useful because the horizontal resolutions of CRU and JRA-55 (~50 km) are too coarse to describe the climatology around small lakes. We used the ClimatView web application developed by JMA (JMA, 2014) for monthly mean in-situ meteorological surface observations. ClimatView provides figures and the original data underlying the monthly values. The data originate as CLIMAT messages sent through the Global Telecommunication System from members of the World Meteorological Organization. ClimatView also provides climatological or 30-year mean values as well as time series of monthly mean values. We incorporated eight variables from ClimatView.

Google Maps is used as the base map for displaying geographical information about lakes. Google Maps also offers satellite imagery, street views, and other features such as My Maps letting users create their own map that are useful for obtaining visual information about lakes. In addition, the Degree Confluence Project (DCP; http://confluence.org) has hyperlinked local data for many integer-degree intersections of latitude and longitude (“confluence points”) located on land, providing land cover information and environmental conditions from these locations.

Developed data

Lake location information in WLDB was quality checked. The information has a coarse spatial resolution of at most 1 minute, or about 1.85 km. This horizontal resolution may be enough for large lakes but not for small lakes. In addition, when we set the display region for each lake on Google Maps, we found that some lake location information contained inappropriate values. In this case, we computed the center of gravity of each lake water surface area with the 1-km horizontal resolution of GLWD Level 3 in order to objectively evaluate the lake location information and compare it with the one in WLDB.

Three parameters are required for good display of each lake on Google Maps: the center location of the display region; longitude and latitude; and the map scale. First guess values for the three parameters are obtained from 1-km horizontal resolution of GLWD Level 3 by searching the maximum and minimum of longitude and latitude of a lake water surface. We determined the best values by visual observation by correcting the first guess values.

We prepared climatological mean values for all variables of CRU and JRA-55 by taking 30-year means from 1981 to 2010 as a standard of the World Meteorological Organization since the current version of CGLB does not implement any mathematical operations among variables except for anomalous values or an instantaneous value minus climatological mean value.

SYSTEM OF WEB APPLICATION

The structure of CGLB is depicted schematically in Figure 1 and described below.

System environment

Web server

The server uses Linux as the operating system plus three general-purpose Linux applications: Apache 2.2, a commercial-grade HTTP server with freely available source code (http://apache.org/); PHP for web development and embedment into HTML (http://php.net/manual/en/); and ImageMagick for creation, editing, composition, and conversion of bitmap images (http://imagemagick.org/). In addition, it uses the Grid Analysis and Display System (GrADS; http://iges.org/grads/) for drawing figures as a background task. The data from CRU and JRA-55 were prepared in GrADS default format: gridded binary data files in “stream” or “direct access” mode with GrADS data descriptor files or control files containing filenames for the binary data, missing or undefined data values, mapping between grid coordinates and world coordinates, and descriptions of variables in the binary dataset.

Client PC

The following browsers are supported: Firefox, Google Chrome, and Internet Explorer 9 or later. JavaScript is enabled to execute the Google Maps API, and cookies are used to store information about figures and GrADS scripts from the user’s previous session.

System specification

The contents of the system are designed to be flexible by using templates and style sheets so that they can be easily updated. The system is also designed to easily add and delete information about individual lakes and climatological data.

Connection with Google Maps

Google Maps is used for displaying a target lake and the 18 types of associated information mentioned above (Supplement Figure S1). The target lake can be selected by choosing a country and lake from a pull-down menu. The displayed area information in Google Maps is transferred to GrADS for drawing a two-dimensional or spatial map (Supplement Figure S2).

Web interface

The web interface was developed using PHP and JavaScript, based on information in the GrADS control files, and contains a dropdown menu and radio buttons. The user’s choice of items from the menu and buttons generates a GrADS script that draws the desired figure (Supplement Figure S3). The GrADS script is run in batch mode and produces the figure in Encapsulated PostScript (EPS) format, then ImageMagick converts it to Portable Network Graphics.
EXAMPLE ANALYSIS FOR LAKE TURKANA

First, we demonstrate what kinds of climate and lake information CGLB can easily provide to users by a simple click during preliminary investigation. Second, we demonstrate a preliminary time-lag analysis after downloading time-series data from CGLB.

Study area and climate

Lake Turkana, the world’s largest permanent desert lake, lies mostly in northwestern Kenya with a small portion in southwestern Ethiopia (Table II). CGLB displays a variety of information about Lake Turkana, as shown in Figure 2 and Supplement Figure S4. Figures S4 and S5 illustrate what can be done in CGLB. Supplement Figure S4 shows information about Lake Turkana available at the web pages hyperlinked from CGLB, while Supplement Figure S5 shows interactive display of climatological variables. Figure 3 shows the landscape of the confluence point 2°N, 37°E, which is 13 km southeast of the southern tip of Lake Turkana and the bottom left panel of Supplement Figure S4 provides the satellite image of the surrounding area of Lake Turkana. The vegetation at this point is scrub and the surface soil is very dry, typical of the Lake Turkana basin. Supplement Figure S5 provides a schematic showing the interactive assembly of climate figures for the Lake Turkana Basin in

Table II. Hydrological features of Lake Turkana (ILEC, 1999)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface area</td>
<td>67,500 km²</td>
</tr>
<tr>
<td>Catchment area</td>
<td>131,000 km²</td>
</tr>
<tr>
<td>Volume</td>
<td>204 km³</td>
</tr>
<tr>
<td>Elevation</td>
<td>360.4 m</td>
</tr>
<tr>
<td>Max depth</td>
<td>109.0 m</td>
</tr>
<tr>
<td>Mean depth</td>
<td>30.2 m</td>
</tr>
<tr>
<td>Residence time</td>
<td>12.5 yr</td>
</tr>
<tr>
<td>Water level control</td>
<td>Unregulated</td>
</tr>
</tbody>
</table>

(PNG) format. GrADS scripts and figures are downloadable by the originating user for 14 days after their production.
CLIMATES OF GLOBAL LAKE BASINS

Figure 3. Screenshot from the Degree Confluence Project (DCP) showing land cover at a point 37°E, 2°N near the southern tip of Lake Turkana (http://confluence.org/confluence.php?lat=2&lon=37).

CGLB. The top and second-row panels show a time-series of CRU surface air temperature and precipitation, and divergence of water vapor flux. The lower panels of Supplement Figure S5 show the geographical distribution of climatological monthly mean precipitation. Very small amounts of precipitation were observed in Lake Turkana but considerable precipitation was observed in the area surrounding the lake basin which is clearly responsible for maintenance of the world’s largest permanent desert lake.

Although some registered CLIMAT stations exist in the Lake Turkana basin, their data are not available. The nearest station with available data is Moyale, Kenya (Figure 4). There the surface air temperature has a typical seasonal cycle with a maximum in February and minimum in July. Precipitation has a biannual cycle, with a primary maximum in April and a secondary maximum in November that coincide with the northward and southward migrations of the intertropical convergence zone over Lake Turkana, respectively. This seasonal cycle is typical of most tropical lakes, such as Gatun Lake in Panama (Nakaegawa et al., 2015). The two rainy seasons are called the long rains season (March to May) and the short rains season (October and November; e.g., Nakaegawa et al., 2012). Climatological monthly mean divergence of water vapor flux over Lake Turkana shows a biannual cycle similar to that of precipitation (second row center panel of Supplement Figure S5.).

Figure 5 shows a similar record for Lake Turkana based on CRU data. The seasonal cycles of surface air temperature and precipitation are quite similar at Moyale and Lake Turkana, but the absolute values differ because the two places differ in elevation by 752 m. The temperature difference of 7.4°C under a dry adiabatic lapse rate assumption...
corresponds satisfactorily with the difference of 7.0°C in annual mean surface air temperature between the two locations.

The water level in Lake Turkana varies on multiple timescales (Figure 6). It varies on a decadal scale, with maxima in 1998 and 2015 and minima in 1996 and 2006. A seasonal cycle in water level is apparent in Figure 6, but the phases of the seasonal cycle differ in the high and low phases of the decadal-scale variability (Figure 7). High water levels occur from October to December during the decadal minimum and from December to February during the decadal maximum. The peak in the climatological seasonal cycle falls in September owing to the averaging of these different phases.

Another difference in phase between precipitation in the Lake Turkana basin and the water level in Lake Turkana results from processes such as soil moisture infiltration, shallow groundwater residence, and river channel routing, which retard the input of water to the lake (Nakaegawa and Hosaka, 2008). The top right panel of Supplement Figure S5 shows a time-series of annual mean divergence of water vapor flux over Lake Turkana. This divergence corresponds to upward water vapor flux from the land surface. The divergence also varies on a decadal time scale with large interannual variability, like the lake water level but with maximum in 2007 and minimum in 1997. This fact suggests that decadal variability of atmospheric circulation affects that of the lake water level. The period of the lake water level available from GRLM is too short for analysis of decadal variability but long for that of interannual variability. Therefore, we chose to analyze the interannual variability of water level in December when water levels were consistently high throughout the period of data.

The world’s largest permanent desert lake, Lake Turkana, is maintained by precipitation in the surrounding area. This means that land-surface hydrological processes in Lake Turkana basin are key to the maintenance of the water surface in this desert region. Hydrological processes such as precipitation-runoff often produce a phase difference between precipitation and runoff. Thus, we performed a time-lag correlation analysis between the water level and two hydroclimatological variables: sea surface temperature (SST) and CRU precipitation which are downloadable from CGLB. Note that the following analysis was performed in offline mode rather than in online mode on CGLB since no mathematical operations are implemented in CGLB.

The time-lag correlation of the December water level with precipitation suggests that interannual variability in precipitation does not control the Lake Turkana water level directly (see Supplement Text S1 and Figures S6d-f). In the very dry environment of the Lake Turkana basin, precipitation returns quickly to the atmosphere through evaporation and does not affect the lake water level. Since SST is more temporally stable and is the primary control of interannual climate variability, the time-lag correlation of the Lake Turkana water level with SST was computed. For time lags of 0, 1, and 2 months, strong correlations were found in the central tropical Pacific (Figures S6a–c) with spatial patterns resembling that of SST anomalies during the central Pacific El Niño (Kao and Ju, 2009) or El Niño Modoki (Ashok and Yamagata, 2009) phenomena. In contrast, no significant correlation was seen in the Indian Ocean except for SST in the Arabian Sea leading by 2 months (shades of Supplement Figure S6c).

CONCLUSION

We have developed a web application, Climates of Global Lake Basins (CGLB), which integrates existing climatological and geographic datasets to display a wide range of information pertaining to lakes of the world. Users can readily
obtain data on geographical features, fundamental climatological mean values and time series, vegetation, and water levels of lakes registered in the database. CGLB can draw time series of climatological data in both one-dimensional and geographical formats and can produce animations of the resulting figures. These functions may be useful for teaching physical geography and earth science in secondary education, for planning expeditions, and for scientific research.

CGLB is suitable for feasibility studies or preliminary studies of lakes. As a demonstration of a preliminary scientific analysis, we investigated the linkage between the Lake Turkana water level in Africa, and SST and (regional) precipitation using time-lag correlation analysis. Precipitation showed no significant correlation with the Lake Turkana water level for time lags of 0, 1, and 2 months. This suggests that variation in land-surface hydrological processes are key to the interannual variability of the water level. In contrast, SST in the central tropical Pacific was strongly correlated with the Lake Turkana water level at all three time lags. No significant correlation was found between the lake water level and SST in the Indian Ocean adjacent to this part of Africa, although significant correlations were seen between regional precipitation and SST in the Arabian Sea leading by 2 months.

The total number of lakes registered in the current version is about 650 based on WLDB (ILEC, 1999), which is small compared to the latest dataset of GLDB (Kourzeneva et al., 2012). No mathematical operations among variables have been implemented, such as correlation or the four basic arithmetical operations mentioned in the Design section. The analysis for Lake Turkana was performed in an offline mode. Thus, a pressing need for future work is to increase the number of lakes in CGLB and to implement mathematical operations.

ACKNOWLEDGMENTS

We are grateful to all the people who allowed us to incorporate their data into CGLB, and to use the snapshots of their web pages as figures. We are also grateful for the freely available software that we used in CGLB. This study was supported by KAKENHI Grant 23226012 from the Japan Society for the Promotion of Science. This research was also supported by KAKENHI Grant 25179–25204. DOI: 10.1029/ 95JC02125.

REFERENCES


Nakaegawa T. 2012. Comparison of water-related land cover types in six 1-km global land cover datasets. Journal of Hydro-


