Lake-Depth Data Set

Global Lake Database (GLDB) represents global lake depth information on the fine grid with the resolution of 30 arc sec (approximately 1 km). Depth information for the individual lakes from different sources – list of lakes – is in text form, and contains information about approximately 14 000 freshwater lakes. Information in gridded form is customary for the atmospheric modeling community and gives the possibility to use it for global models. As a land-use dataset ECOCLIMAP2 is chosen. For some large lakes the bathymetry is included. For mapping data for individual lakes a novel method (complicated, though justified mathematically) solving an optimization problem for search algorithm is used.

Which land cover dataset is used in GLDBv.1 and GLDBv.2 for lake masking and why?

Raster map ECOCLIMAP2 is used for lake mapping.

A practical argument for this choice was that ECOCLIMAP2 (resolution approximately 1 km) is widely used in numerical weather prediction (NWP) and climate modeling, and thus less consistency problems will appear in the future. In respect of inland water ECOCLIMAP2 classify separately seas, lakes and rivers. Note that rivers are recognized quite poorly, and very often they are mixed up with lakes. When the grid size of a numerical atmospheric model is larger, the lake fraction can be easily calculated from this land cover dataset in a standard way.

What data the individual lake list contains?

For each individual lake in the list the following information is collected:

- geographical coordinates (latitude, longitude) of some point on the lake water surface (degrees);
- the mean depth of the lake (meters);
- the maximum depth of the lake (meters);
- the surface area of the lake (square km);
- the lake name (both international and local);
- the name of the country where the lake is located.

Where the data were missing, the no data symbol ‘9999.0’ is used.

What type of lakes the individual lake list contains and why?

The individual lake list contains information of natural and manmade lakes, as well of freshwater and brackish lakes (lakes with low salinity – less then 10 ‰, and with stable size and shape). Saline lakes and endorheic basins form the additional dataset.

Thermodynamically, the behavior of natural and man-made lakes is similar, hence, it was not distinguished between natural and artificial lakes. Brackish lakes (with low salinity and stable surface area) behave similar to freshwater lakes, so they were not separated from the main dataset.

Saline lakes and endorheic basins behave differently from freshwater lakes and they can hardly be simulated by freshwater lake models. Hence, information for saline lakes and endorheic basins is kept in an additional dataset.

Which lakes have bathymetry data in GLDBv.1 and GLDBv.2?

For large lakes (surface area is more than 400 square km) that were not too shallow (maximum depth is more than 10 m) or too deep (mean depth is less than approximately 70 m) or with a flat (U-shape) bottom, whose bathymetry data could be found in ETOPO1 or already digitized navigation and topographic maps, bathymetry data is added.

Following 36 large lakes have bathymetry data in GLDBv.1 and GLDBv.2: Albert (Zaira), Athabasca (Canada), Balkhash (Kazakhstan), Beloe (Russia), Biwa (Japan), Bratskoe1 (Russia), Bratskoe2 (Russia), Champlain (USA), Chudskoe (Russia), Edward (Zaira), Erie (Canada), Great Bear (Canada), Great Slave (Canada), Hjalmaren (Sweden), Huron (Canada), Kujbishevskoe (Russia), Ladoga (Russia), Lough Neagh (UK), Malaren (Sweden), Michigan (USA), Onego (Russia), Ontario (Canada), Pskovskoe (Russia), Pyramid (USA), Rybinskoe (Russia), Saint Clair (Canada), Sevan (Armenia), Skadar (Armenia), St Jean (Canada), Superior (Canada), Taymyr (Russia), Ustllimskoe (Russia), Vanern (Sweden), Vattern (Sweden), Victoria (Tanzania), Winnipeg (Canada).

Why only selected types of large lakes have bathymetry data in GLDBv.1 and GLDBv.2?

For large lakes (covering several grid boxes of an atmospheric model) information about the bathymetry is useful. A 1D lake model, running in every atmospheric model grid box over the large lake using specific lake depth, makes it possible to reproduce the surface temperature patterns for such lakes. However bathymetry data are not so easy to obtain. That’s why types of large lakes that can be characterized by their mean depth (don’t need bathymetry data) in models were determined. They are following:

- deep lakes (deeper than approximately 70 m) – the surface temperature annual cycle and surface temperature patterns are mainly controlled by atmospheric forcing. The long-term mean surface temperature patterns do not necessarily follow the bathymetry pattern. In lake models, the sensitivity of the surface temperature to the depth parameter is also rather low for very large depth values. This makes it possible in NWP to apply lake models that were developed for medium lakes, using for deep lakes an artificial limitation in depth;
- very shallow lakes (with a maximum depth of less than 10 m) – the surface temperature annual cycle and surface temperature patterns are controlled by atmospheric forcing;
- lakes with a flat bottom (U-shape) – in such cases, variations in depth are too small to influence the surface temperature patterns. It is considered that large lakes pertain to this category if the difference between the mean and maximum lake depths is less than 6 m.

How the mapping method in GLDBv.1 and GLDBv.2 works?

An automatic mapping method to combine depth data for individual lakes with a raster map is used. The method is probabilistic: it is assumed that all data (coordinates of lakes and lake coastlines) have random errors. For more detailed information of how the mapping method works, please see article of Kourzeneva E. (2012).

What happens with the saline lakes data in GLDBv.1 and GLDBv.2?

Data on freshwater lakes are processed and data on saline lakes are skipped.

What happens if lake depth data in GLDBv.1 are missing?

When there is a lake on the map, but its depth is unknown (no data in the dataset for individual lakes), a ‘default’ depth of 10 m is used.

How data in GLDBv.1 are represented?

GLDBv.1 consists of:

- list of individual lakes, which contains information about approximately 13 000 freshwater lakes, and is presented in a text form;
- bathymetry data of 36 large lakes;
- global gridded lake depth data set with a resolution of 30 arc sec. (approximately 1 km), what means that it consists of 43200 columns and 21600 rows. It contains the mean lake depth in decimeters, and for 36 large lakes it provides the bathymetry. In areas with missing data, the map shows the 'default' depth of 10 m;
- additional global gridded dataset with a resolution of 30 arc sec. (approximately 1 km), what means that it consists of 43200 columns and 21600 rows. It contains the coded information about the origin, thus reliability, of data provided in global gridded lake depth data set. The origin code is given for every pixel according to the following legend:
    - (0) no lake or river (sea or land), depth=0 m;
    - (1) no information about this lake in the list of individual lakes, depth=10 m;
    - (2) missing lake depth in the list of individual lakes, depth=10 m;
    - (3) information about lake depth is taken from the list of individual lakes or bathymetry;
    - (4) river, depth=3 m.

What does geological approach used in GLDBv.2 means?

Geological approach, according to the depth estimation of uninspected freshwater lakes, assumes that water bodies of the same origin and the same age should have a similar size. For example, it is widely known that tectonic lakes are deeper than glacial, karst lakes usually have small surface area, but are rather deep, and eolian lakes are usually shallow.

What does bottom-up approach used in GLDBv.2 mean?

The bottom-up approach was used to outline regions with kindred (typical) geological origin of lakes. The main idea of this approach is to start from a description of the deepest layers (in case of GLDBv.2, deepest layers of the Earth crust – tectonic plates), gradually elevating to the surface.

What information was used to obtain regions with kindred origin of lakes in GLDBv.2?

First of all information about the deepest layers of the Earth crust from Tectonic Plates Map – information about tectonic plates and cratons – was used. Secondly, data about middle layers from the geological map was applied – different types of rocks were distinguished, namely igneous and sedimentary. Thirdly, information about upper layers from the map of Quaternary deposits was used – deposits in the last geological period were outlined. Finally information about the top layer from the geomorphologic map about boundaries of the permafrost was applied.

<u>How regions with kindred origin of lakes were outlined in GLDBv.2?</u>

An innovative algorithm, which combines together both information about lake location and geological information was developed and applied in GLDBv.2. For every lake with known mean depth from the individual lake list by the instrumentality of the polygon test using lake coordinates were determined: the tectonic plate, the craton which the lake in question belongs to, presence/absence of the permafrost, Quaternary deposits type, type of rocks on the territory where the lake in question is located. Different combinations of determined information formed an individual code. These codes corresponded to intersections of contours, which were assumed to be the regions with kindred (similar) geological conditions of lake origin. Altogether, 141 regions were outlined.

<u>How typical depths for regions with kindred origin of lakes were obtained in GLDBv.2?</u>

To find the typical lake depth for allocated regions, statistics from the individual lake list and global gridded lake depth data set was gained and analyzed. The core of statistical analysis was building histograms of the lake depth for each region. Maximum of the histogram was considered to be a guess for the typical lake depth for the region in question. Later on, starting from this guess, an expert decision about the typical lake depth was reached.

Steps of performed analysis were following. First, full statistics collected from the list of individual lakes and from the pixels of the gridded lake map were analyzed. If there were enough data and maximums of histograms for individual lakes and for lake pixels agreed well, the expert decision about the typical lake depth could be arrived. If differences between maximums of histograms were significant special filter was used (lakes with surface area more then 200 km^2 were skipped). If it still was difficult to obtain a confident lake depth estimation on the basis of statistical analysis, method of analogies was used. Also, geographical and improved geomorphologic methods were used if necessary or possible.

What is the main idea of ‘method of analogies’?

This approach was used in case of scarce statistics about the lake depth for one region, and sufficient statistics for another region with similar geological and geomorphologic structure. So, the main idea of this method is extrapolation.

What is the application of ‘method of analogies’ to GLDBv.2?

The main idea of this method is extrapolation of statistics from one region-analogue to another using certain geological knowledge. From this knowledge, regions-analogues were specified as follows:

- regions with glacial, marine and fluvioglacial Quaternary deposits of one tectonic plate;
- cratons of one tectonic plate;
- Precambrian shields of different tectonic plates;
- orogenies of one tectonic plate.

The orogenies of one tectonic plate were not combined, this information was used just to assess reliability of depth estimates.

What is the main idea of ‘improved geomorphologic method’?

In this method, using the large dataset of in situ measurements, the relations (statistical dependencies) between the lake water volume (from which the mean lake depth may be derived) and the surface area for each region with different geological situation are established.

What is the application of ‘improved geomorphologic method’ to GLDBv.2?

This method was developed by hydrologists basing on information about geological origin of lakes only for Northwest Russia for the middle size lakes (> 10 km^2). In GLDBv.2 this method is considered to have high reliability, so it is given the highest priority on the appropriate territory – Northwest Russia.

What is the main idea of ‘geographical method’?

This method implies that the mutual distribution of lake parameters is specific for each geographical zone. Geographical zones of tundra, northern taiga, middle taiga and mixed forest were considered. This approach is not related to lake origin, although it may implicitly account for morphology of the territory through the dependency of vegetation on lithology of a rock or permafrost conditions. Statistical dependencies distributions of the mean lake depth for different zones depending on the lake area were presented in tables.

What is the application of 'geographical method' to GLDBv.2?

This method was developed only for the boreal zone. The analytical equations approximating statistical dependencies distributions of the mean lake depth for different zones depending on the lake area were generated and used instead of tables. This method has very low reliability. So, it is used for the territories where neither statistical expert evaluations, nor improved geomorphologic method are possible.

What happens if lake depth data in GLDBv.2 are missing?

In absence of direct measurements (no mean depth data in the individual lake list), indirect estimates (though less precise) obtained by an expert evaluation method, or the geographical method, or the geomorphologic method are used. If even then mean lake depth is unknown, a 'default' depth of 10 m is used.

How data in GLDBv.2 are represented?

GLDBv.2 consists of:

- list of individual lakes, which contains information about approximately 14 000 freshwater lakes, and is presented in a text form;
- bathymetry data of 36 large lakes;
- global gridded lake depth data set with a resolution of 30 arc sec. (approximately 1 km), what means that it consists of 43200 columns and 21600 rows. It contains the mean lake depth in decimeters, including indirect estimates (or predictions) and for 36 large lakes it provides the bathymetry. In areas with missing data (no measured or estimated depth), the map shows the 'default' depth of 10 m;
- additional global gridded dataset with a resolution of 30 arc sec. (approximately 1 km), what means that it consists of 43200 columns and 21600 rows. It contains the coded information about the origin, thus reliability, of data provided in global gridded lake depth data set. The origin code is given for every pixel according to the following legend:

(0) no lake or river (sea or land), depth=0 m;

(1) no information about this lake in the list of individual lakes, depth=10 m;

(2) missing lake depth in the list of individual lakes, depth=10 m;

(3) information about lake depth is taken from the list of individual lakes or bathymetry;

(4) river, depth=3 m;

(5) lake depth is estimated by expert evaluations;

(6) lake depth is estimated by the geographical method;

(7) lake depth is estimated by the geomorphologic method.

What are the differences between GLDBv.1 and GLDBv.2?

There are several differences between GLDBv.1 and GLDBv.2, namely:

- The individual lake list was increased by approximately 500 lakes;
- The global gridded lake depth data set was completed with indirect estimates of the mean lake depth from the geological origin of boreal zone lakes;
- The additional global gridded data set containing coded information about sources of data was also updated.

References:

Kourzeneva, E., Asensio, H., Martin, E. and Faroux, S. 2012. Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modeling. Tellus A. 64, 15640. DOI: 10.3402/tellusa.v64i0.15640.

Choulga, M., Kourzeneva, E., Zakharova, E. and Doganovsky, A. 2014. Estimation of the mean depth of boreal lakes for use in numerical weather prediction and climate modeling. Tellus A 2014, 66, 21295, http://dx.doi.org/10.3402/tellusa.v66.21295