Vector dataset for river systems originating in Eurasia to the Arctic Ocean

Liang Wenshan\textsuperscript{1,2}, Qiu Yubao\textsuperscript{2*}, Xie Pengfei\textsuperscript{3}, Shi Lijuan\textsuperscript{2}, Wang Xingxing\textsuperscript{2}

1. Jiangxi University of Science and Technology School of Architectural and Surveying & Mapping Engineering, Ganzhou 341000
2. Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094
3. Hebei University of Technology School of Artificial Intelligence Tianjin 300400

(qiuyb@radi.ac.cn)

Abstract. Arctic rivers are part of the freshwater resources for the Arctic and are of great significance to research into changes in the Arctic Ocean. However, global river datasets do not describe rivers in Eurasia in sufficient detail. Based on a 30-m resolution land surface water dataset in 2010 together with global river width data (GRWL) and reference SRTM 30-m digital elevation data, the river systems in each Arctic basin were identified. Google Earth images were used for verification and the resulting accuracy was 93.96%. The areas of the rivers in each country studied were also analyzed and it was found that Russia has the largest area of rivers and Mongolia has the highest ratio of river area to country area. Based on these river datasets, the river ice phenology for the Ob River basin was analyzed using cloud-free binary snow-cover data obtained by Moderate Resolution Imaging Spectroradiometer (MODIS). The result showed that the open-water period for the Ob River is from June to October.

Keywords: Arctic; Eurasian; River; Basins; Vector dataset

1. Introduction

In the past 50 years, water bodies such as rivers and lakes in the Eurasian Arctic and their interactions with energy have been greatly affected by global warming, which has caused great changes in the surface water system in the Arctic region. Many links between the Arctic and the global climate involve the hydrological cycle. River discharge is the main source of freshwater in the Arctic, and changes in this have the potential to change the global ocean circulation (Rahmstorf S., 2002) and the North Atlantic thermohaline circulation (McClelland J W et al., 2006). Basic data concerning the Arctic rivers originating from Eurasia is of great significance to studies of changes in the Arctic sea ice and Arctic Ocean. There are many global water-body datasets in existence; for example, M. L. Carroll used the SWBD (SRTM Water Body Dataset) dataset and MODIS (Moderate Resolution Imaging Spectroradiometer) data to produce the Global raster water mask, which provided global water data at a resolution of 250 m (Carroll M L et al., 2009). Jean François, P. conducted an analysis of global water change using Landsat data from 1984 to 2015 and provided six sets of related aquatic products that reflected the changes in water bodies according to different indicators (P. Jean-F et al., 2016). L.Y. Ji used MODIS data to generate daily global water, ice/snow and land raster datasets with a resolution of 500 m, covering the period from 2001 to 2016 (L.Y. Ji et al., 2018). However, it is clear that the current global river datasets do not describe the rivers in Eurasia adequately. Moreover, the inter-annual variations in the Eurasian rivers are large and due to the problem of the phase of the dataset, there may be unidentified rivers, including rivers that have been partially interrupted or newly added water body data.

In this study, global 30-m resolution land surface water data (J. Chen et al., 2014) were used to study the Arctic rivers originating from Eurasia. Following vector data preprocessing and river identification, vectorized Arctic river data for Eurasia was produced. Based on this, the area and
the proportion of the rivers flowing through each country were statistically analyzed, thus providing a basic understanding of the utilization of freshwater resources together with an assessment of the sustainability of water resources in different countries. In addition, the river data were combined with cloud-free MODIS binary snow-cover data (Y. B. Qiu et al., 2017). The Ob River Basin was selected as an example, and the river ice phenology in the basin was analyzed, thus providing support for river basin climate analysis and river navigation.

2. Data and Method

2.1 Data

The dataset used in this study was based on a global 30-m resolution land surface water dataset (J. Chen et al., 2014), which was based on 30-m LandsatTM5\ETM+ multispectral data covering a period of around 1-2 years in 2010 and 2011 together with data acquired by the Chinese environmental disaster reduction satellite. Because it is assumed that the imagery will be less cloudy, this dataset is based mainly on image data from the vegetation growing season. The data are subjected to geometric and radiometric correction. The computer classification algorithm then classifies the global water bodies, uses object-oriented processing methods to remove paddy fields and wetlands, and then performs data integration and mapping on water bodies including rivers, lakes and reservoirs.

HydroBASINS global basin data (Lehner B et al., 2013) depict global basin boundaries and sub-basin divisions. The data provide a hierarchy of nested basins of varying sizes from tens of square kilometers to millions of square kilometers and use a HydroSHEDS database with a 15 arc-second grid resolution to separate the basins in a consistent manner at different scales. Based on the topological concept of the Pfafstetter coding system, different levels of basins are defined. These data were used to obtain information about the basins of the Eurasian Arctic.

2.2 Data processing techniques and methods

Based on the global 30-m resolution land surface water dataset in 2010, in this study the Arctic rivers originating from Eurasia were taken as the research subject. After preprocessing of the vector data, projection transformation, block splicing and data cutting, GRWL global river-width data and SRTM 30m Digital Elevation Data (DEM) (http://gdex.cr.usgs.gov/gdex/) were used to identify rivers in each subdivision. Figure 1 shows a flowchart representing the processing of the river vector data.

Next, the 30-m vector data for the Arctic rivers originating in Eurasia were constructed. This data set covered 20 basins across six countries and corresponded to a total river area of 225,246 km$^2$. Finally, to verify the results, the river vector data were compared with the existing international rivers database and Google Earth high-resolution imagery.
3. Results and validation

3.1 Results

Vector datasets for river systems originating in the Eurasian Arctic include streams originating from Eurasia, as well as rivers such as the Lena, Ob and Yenisei that flow into the Arctic Ocean. Statistics on the Arctic rivers of Eurasia show that the total river area is 225246 km$^2$. Among these, the rivers with the largest discharge into the Arctic Ocean are the Yenisei, Ob, Lena, Kolyma, Pechora and Northern Dvina.

The Yenisei River Basin has the largest river area of 75357.57 km$^2$; the river areas in the Ob, Lena, Kolyma, Pechora and Northern Dvina are 38744.91 km$^2$, 26794.14 km$^2$, 5304.16 km$^2$, 4349.74 km$^2$, and 2567.89 km$^2$, respectively.
Table 1. River areas for the six rivers with the largest discharge into the Arctic Ocean

| Basin          | Yenisei River | Ob River | Lena River | Kolyma River | Pechora River | Northern Dvina River |
|----------------|---------------|----------|------------|--------------|---------------|----------------------|
| Area (km²)     | 75357.57      | 38744.91 | 26794.14   | 5304.16      | 4349.74       | 2567.89              |

3.2 Data Validation

An area within the river data set was randomly selected as the sample area. As shown in Figure 4, the river verification sample for the Kolyma River Basin, the GRWL river width data reference sample and the river vector data verification sample were first obtained. 34898 river reference samples and 9438 river verification samples were collected. If the number of river reference samples and verification samples were the same, the verification samples were considered correct; if the number of river reference samples and verification samples were inconsistent, the samples were imported into Google Earth for verification.

The formula used to calculate the accuracy of the dataset was in equation (1):

\[ a = \frac{b}{b+c} \]  \quad (1)

where \( a \) is the accuracy of the dataset, \( b \) is the correct number of random samples, and \( c \) is the number of random samples.

The mapping accuracy for this data set was calculated to be 93.96% in this study.
4. River spatial distribution pattern

The rivers flowing into the Arctic Ocean flow through China, Russia, Kazakhstan, Mongolia, Finland, and Norway. Ordered by the area of rivers from greatest to smallest, the order is Russia, Mongolia, Kazakhstan, Norway, Finland, and China.

Of the countries listed, Russia has the largest river area and also accounts for the largest fraction of the total river area - nearly 80%; its rivers are mainly distributed in the north of the country. The Yenisei River, which has the highest flow of any river in Russia, also contains the most water of any Siberian river. Mongolia is second to Russia in terms of river area as well as in terms of the proportion of the total river area accounted for by the country; it also has the highest ratio of river area to total land area. There are 3,800 rivers within Mongolia, most of which are distributed in the central and northern regions. These rivers, for example the Selenga River, generally flow towards Russia. The Selenga River originates on the northern slope of Hangay Uul Mountain and flows through northern Mongolia and then into Lake Baikal, Russia. The distribution of rivers in Kazakhstan varies greatly, with most being in the northeast of the country and the number decreasing from north to south. The Ili River and the Irtysh River are the two largest transnational rivers that cross the border between Kazakhstan and China. Among the six countries listed above, China has the smallest river area to country area ratio. The only river that flows into the Arctic Ocean from China is the Irtysh River, which also flows through Kazakhstan and Russia. Norway has rivers flowing into the Arctic Ocean all over the country; these rivers are shorter and the water flow is more rapid. The Kovasselva River in Norway is one of the shortest rivers in the world.

| Country | River area (km²) | Proportion of total river area within the country (%) | Ratio of river area to area of the country (%) |
|---------|-----------------|------------------------------------------------------|---------------------------------------------|
| China   | 1805.92         | 0.80%                                                | 0.01%                                       |

Table 2. Statistics for river area and proportion of the area of each country accounted for by rivers
5. River data application

The Ob River basin is spread across China, Russia, Mongolia and Kazakhstan. Of the rivers studied, the Ob flows through the largest number of countries and has the greatest significance. The Ob River Basin was taken as an example. Using the Eurasian Arctic river data and the MODIS snow products (http://modis-snowice.gsfc.nasa.gov) in 2010, the ice on the Ob River was identified after data clipping, cloud removal and binarization. The results of this are shown in Figure 6. On 21st March 2010, the rivers in the lower reaches of the Ob River basin were in a completely frozen state. On 24th April 2010, as the northern hemisphere entered spring and temperatures rose, the river ice began to break up and open water appeared in the river. From 30th April to 10th May, the ice in the river gradually broke up from south to north. By 21st June, the river was completely open. By 30th October of the same year, the river was frozen again. It can be seen that the lower reaches of the Ob River have a river icing period of November to May, and normal navigation can be carried out in June. By using the vector river data combined with MODIS data, river ice information can be obtained, and the processes of river freezing and ice break-up monitored, thus providing a scientific basis for river navigation.
6. Conclusion

Using the 2010 global 30-m resolution land surface water dataset, this study took the Arctic rivers originating in Eurasia as the research subject. The GRWL river database and 30-m SRTM DEM data were used to identify and to construct vector data for the Arctic rivers of Eurasia. The total area of these rivers and the proportion of river area to the total river area were calculated for each country. In terms of the rivers flowing into the Arctic Ocean, Russia has the largest river area, and the Mongolian river area accounts for the highest proportion of the country area. These results provide data for analyzing the spatial distribution of Arctic rivers and for assessing the use of freshwater resources by the countries concerned. In addition, the river vector data were combined with cloud-free MODIS binary snow-cover data to carry out a preliminary analysis of the phenology of the ice on the Ob River. It was concluded that the navigation period for the lower reaches of the Ob River is from June to October; these results provide support for analyzing the effects of river ice on navigation.

Multi-temporal basic vector data of the Arctic rivers in Eurasia and a large-area, long-term sequence of river ice phenology are still being processed.

Support: This research was supported by the Multi-Parameters Arctic Environmental Observations and Information Services (Grant number 2017YFE0111700), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant number XDA19070201) and the International Cooperation Program of Chinese Academy of Sciences (Grant number 131211KYSB20170041).

Reference
[1] Rahmstorf S 2002 J. Nature. Oceanic circulation and climate during the past 120,000 years. 419 207-214
[2] McClelland J W, Déry S J, Peterson B J, Holmes R M, Wood E F 2006 J. Geophysical Research Letters. A pan-arctic evaluation of changes in river discharge during the latter half of the 20th century.
[3] Carroll M L, Townshend J R, Dimiceli C M, Noojipady P, Sohlberg R A 2009 *J. International Journal of Digital Earth*. A new global raster water mask at 250 m resolution. 2 18

[4] JeanFrançois P, Cottam A, Gorelick N, Belward A S, 2016 *J. Nature*. High-resolution mapping of global surface water and its long-term changes. 540 418-422

[5] Ji L, Gong P, Wang J, Shi J, Zhu Z 2018 *J. Water Resources Research*. Construction of the 500-m Resolution Daily Global Surface Water Change Database (2001–2016). 54 10-270

[6] Yubao Qiu, Huan Zhang, Duo Chu, et al. 2017 *J. Journal of Glaciology and Geocryology*. Snowless and Coudless Algorithm on the Tibet Plateau Based on MODIS 3 65-70

[7] Chen J, Liao A P, Chen L J, Zhang H W, He C Y, Peng S, Wu H, Zhang W W, Li R, Zheng X Y, Mei Y, Lu M, Lu L, Kang R, Xing H Q, Liu L F, Han G, Wang J, Yang A L, Sun L M, Song H, Chen W P, Zhou X M, Xia L, Jiang H B, Huo J, Zhang Yan, Liu W D, Li Y X, Zhai L, Sang H Y 2014 *DB. Global Change Research Data Publishing & Repository*. Global 30m resolution land surface water dataset.

[8] Lehner B, Grill G 2013 *J. Hydrological Processes*. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. 27 2171-2186