Report 41

GRDC Report Series

Derivation of watershed boundaries for GRDC gauging stations based on the HydroSHEDS drainage network

Technical Report prepared for the GRDC

Bernhard Lehner
Department of Geography
McGill University, Montreal, Canada

Global Runoff Data Centre

GRDC operates under the auspices of the World Meteorological Organization (WMO) with the support of the Federal Republic of Germany within the Federal Institute of Hydrology (BfG

Global Runoff Data Centre

in the
Federal Institute of Hydrology (BfG)

Am Mainzer Tor 1
56068 Koblenz, Germany

P.O.Box 20 02 53
56002 Koblenz, Germany

Phone: +49 261 1306-5224
Fax: +49 261 1306-5722
E-Mail: grdc@bafg.de
Internet: http://grdc.bafg.de

About the Global Runoff Data Centre (GRDC):

The GRDC is acting under the auspices of the World Meteorological Organization (WMO) and is supported by WMO Resolutions 21 (Cg XII, 1995) and 25 (Cg XIII, 1999). Its primary task is to maintain, extend and promote a global database on river discharge aimed at supporting international organizations and programs by serving essential data and products to the international hydrologic and climate research and assessment community in their endeavour to better understand the earth system. The GRDC was established at the Federal Institute of Hydrology in 1988. The National Hydrological and Meteorological Services of the 187 member states of WMO are the principal data providers for GRDC.

January 2012

Foreword

Hydrologic processes und the resulting phenomena are studied and reported in logic „water“ units at local, regional, national or international levels. A catchment is commonly recognised as the abstract unit of study and reporting in hydrology, the basin as the physiographic unit where hydrologic processes take place.

Determined by the topographical and geological conditions, a basin is bounded by a watershed (line). Within the basin, all waters flow to a common outlet, which is determined by the lowest point on the bounding watershed. Geometrically, a basin may be described by its watershed polygon and a planar basin area.

The widespread use of GIS in hydrology and environmental sciences, led to an increasing demand for basin polygons. Modern GIS technology allows for the delineation of basins for almost every point on the Earth’s surface. Using the HydroTools of ArcGIS, GRDC generated the “Major River Basins of the World”, a set of shape files created 2009 for the generation of GRDC map products. Against this background, GRDC is repeatedly asked for the provision of watershed boundaries for the gauging stations represented in the Global Runoff DataBase.

The recently completed HydroSHEDS drainage network (Lehner et al., 2008) offers the unique opportunity to generate watershed boundaries for GRDC gauging stations using a proofed dataset and applying a consistent methodology. GRDC is happy to have engaged Bernhard Lehner for the creation of the watershed boundaries for more 7500 GRDC stations.

The GRDC likes to thank Bernhard Lehner for his work and the permission to publish the results in the GRDC Report Series. We believe that the watershed boundaries of GRDC stations will attract wide interest.

The work documented in this report is a good example of how recent developments in GIS technology help to make the GRDC data set more public and its access more attractive. GRDC invites scientists to assist the centre in the scientific exploitation of its database. A couple of valuable cooperation’s and reports arose from these invitations in the past, at last this interesting report. Therefore, GRDC would like to encourage others to follow this proved tradition.

Contents

List of Tables included

1. Background

GRDC requires explicit watershed boundaries corresponding to the GRDC gauging stations for many applications. Until now, only few GRDC watersheds have been delineated, and the quality of the outlines has been inconsistent due to the use of different sources. The recently completed HydroSHEDS database (Lehner et al., 2008) provides hydrographic data layers and information that allow for the derivation of watershed boundaries for any given location based on the near-global, high-resolution SRTM digital elevation model. Using this hydrographic information, GRDC stations were linked to HydroSHEDS and watersheds were delineated in a consistent manner.

For all following processes, the HydroSHEDS river network model was applied at 500 m (15 arc-second) resolution. It should be noted that the quality of the HydroSHEDS data is significantly lower for regions above 60 degrees northern latitude, as there is no underlying SRTM elevation data available and thus a coarser scale DEM has been inserted (EROS, 2008).

2. Description of executed project tasks

At the beginning of the project (October 2010), the BfG provided the most recent database of GRDC stations containing 7532 records for which watershed outlines should be derived. Of these, 47 stations had to be excluded as there were no point coordinates available. For all other stations, the provided geographic locations in terms of x- and y-coordinates were considered to be of mixed quality, with various uncertainties and likely errors. For this reason, the following two-fold strategy was designed to link the gauging stations to the HydroSHEDS river network. First, an automated process was applied: all stations were linked to the HydroSHEDS river network within a defined radius around the stations while attempting to optimize the agreement between the reported watershed area in the GRDC database and the modeled watershed area derived from HydroSHEDS. If no acceptable location could be detected within the applied search radius, the station was manually inspected in a secondary procedure. The following detailed steps were performed:

2.1 Automatic procedures for station allocation

- For each station, an individual search radius of 5 km was defined.
- Within this search radius, the watershed area was calculated for every pixel of the HydroSHEDS gridded river network.
- The modeled watershed areas (HydroSHEDS) were then compared to the reported watershed areas of the corresponding stations as provided in the GRDC database.
- All pixels with area differences of more than 50% (positive or negative) were excluded from further steps. All other pixels were coded with the absolute value of their area difference (in %); i.e. a pixel with plus or minus 10% error received the value „10“, etc.
- This procedure provided a ranking scheme according to area discrepancies (RA) with values between 0 and 50, where 0 indicates perfect agreement in watershed area.
- Next, for every pixel the distance to the original location of the station was calculated (i.e. the distance from the center of the search radius). The distance values were normalized to reach 50 at the maximum distance of 5 km; i.e. a pixel at a distance of 1 km received a value of „10“, etc.

- This procedure provided a ranking according to distance (RD) with values between 0 and 50, where 0 indicates perfect agreement in station location.
- Both the area and distance rankings were then combined in an additive way to derive a total ranking (R), whereby distance was weighted double (see „note“ below):
- R = RA + 2RD
- This procedure provided a combined ranking with values between 0 and 150, where 0 indicates perfect agreement in both area and distance, and a higher value indicates increasing discrepancies.
- Finally, from all possible pixels that corresponded to a station, the one showing the lowest ranking value was chosen.
- Note: The distance ranking (RD) was weighted double so that further away pixels would quickly increase in their ranking values and thus become less likely to be chosen. More precisely: a pixel that is 1 km further away (2x10 ranking points) will only be chosen if the area agreement improves by more than 20%. These settings were applied after several tests showed that many stations with high precision in their coordinates showed a difference in watershed area of 5-10%, hence this magnitude of area disagreement should not immediately trigger a large movement of the station.

2.2 Manual procedures for station allocation

- All stations for which no area agreement of less than 50% existed within the 5 km search radius were manually inspected. This also included 230 stations that had no reported area in the GRDC database.
- First, the stations were visualized on Google Maps, and it was attempted to verify the river and station names (typically the name of the nearest settlement) in close vicinity to the given location (~10 km).
- If a station could not be verified within this vicinity, the search was extended along the longitude and latitude lines of the given coordinates (for ~50-100 km). This strategy was applied as in many cases the location was incorrect due to errors in either the longitude or latitude coordinate, but not both. Typical errors included: simple typos in one digit (e.g. 11.58ºN instead of 12.58ºN); logical errors in the original coordinates (e.g. -20.4ºW instead of -19.6ºW for a location that is 0.4º to the right of -20ºW); or a swapped order of the coordinate digits (e.g. 10.35ºN instead of 10.53ºN).
- If still no location was found that matched the river and/or station name, the station name was queried in Google Maps to see whether a location with this name existed anywhere in acceptable distance.
- In all cases, the final decision on whether a station was moved to a new and “reliable” location depended on whether at least two out of the following four indicators could be matched reasonably well: a) river name; b) station name; c) watershed area (match between reported GRDC value and modeled HydroSHEDS value); and d) long-term annual discharge (match between reported GRDC value and modeled HydroSHEDS value). This decision was obviously subjective, and difficult combinations could arise (e.g. multiple agreements yet also disagreement(s) in the different indices). If a station was moved, a quality indicator and comment for the decision was added to the record.

- Typically, the agreement in watershed area had highest priority for the final decision on whether to move a station. In some cases, however, e.g. if river and station names could be clearly verified, and also the discharge values matched, it was concluded that the reported GRDC area was possibly erroneous, and the station was moved to the new location despite the area discrepancy (see comments in Table 2).
- In some cases, the GRDC stations were at the correct location but the HydroSHED river network could not represent the situation correctly. These cases included artificial canals, braided rivers, or stations within river deltas (see comments in Table 2).
- For areas above 60 degrees northern latitude the reliability of the results is generally limited due to the low quality of the HydroSHEDS river network. These records should be interpreted with care, even if a high quality is assigned due to well matching areas.
- Similarly, very small catchments (<10-50 km2) are not very reliable, even if the areas match well within a short distance, as small watersheds are found within close proximity to any location (even incorrect locations).
-

2.3 Calculation of watershed polygons and delivery of results

The watersheds for all re-allocated stations were derived based on the HydroSHEDS drainage network using standard GIS tools and procedures. Basin outlines were produced in two versions: with gridded edges (i.e. exactly following the HydroSHEDS raster cells), and with smoothed edges. The resulting polygons (one for each station) were attributed with the corresponding GRDC station records. Both the re-allocated GRDC stations (points) and corresponding watersheds (polygons) were delivered in ESRI shapefile format.

3. Results

In total, 7532 GRDC stations were processed. 7164 point locations were linked to the HydroSHEDS river network and watershed polygons were derived for them. Of these, 6528 stations were automatically linked, while 636 were manually assigned. 368 stations could not be allocated to due various reasons (see Table 2) and no watershed polygon was derived for them. The re-allocated stations were moved by an average distance 2.9 km.

After the stations were assigned to the new locations on the HydroSHEDS river network, the following new attribute columns were calculated:

Table 1: New attribute columns for re-allocated GRDC stations

Column	**Content**
Long_org; Lat_org	Longitude and latitude of original GRDC position in decimal degrees
Long_new; Lat_new	Longitude and latitude of new position on HydroSHEDS river network in dec. degrees
Dist_km	Distance between original and new position in km
Area_hys	Area according to HydroSHEDS in km2
Area_diff	Difference between reported GRDC area and modeled HydroSHEDS area in percent
Disc_hys	Long-term average discharge according to HydroSHEDS in m3/s (based on coarse scale runoff estimates provided by the global hydrological model WaterGAP_2.1)
Disc_diff	Difference between reported GRDC discharge and modeled HydroSHEDS discharge in percent
Elev_hys	Elevation (a.s.l.) according to HydroSHEDS in meters
Quality	Overall quality indicator: High, Medium, Low, or Unassigned (see also Table 2)
Type	Type of procedure: Automatic or Manual (see also Table 2)
Comment	Comment (see also Table 2)

All 7532 stations were attributed with a quality indicator, a type, and a comment indicating the results of the re-allocation process. The following table summarizes the attributes:

Table 2: Quality, type and comments for re-allocated GRDC stations

Quality	**Type**	**Comment**	**Additional explanation**	**Number of occurrences**
High	Automatic	Area difference <= 5% and distance <= 5 km		4697
Medium	Automatic	Area difference 5-10% and distance <= 5 km		806
Low	Automatic	Area difference 10-50% and distance <= 5 km		1025

High	Manual	Good agreement (mostly verified)	At least two of the available indicators (river name, station name, area, and average discharge) could be verified; and area difference <= 5%	263
Medium	Manual	Seems ok (partially verified)	At least two indicators could be verified; and area difference 5-10%	175
Medium	Manual	Seems ok, but area differs (dry parts)	Location seems correct (based on river name, station name, and/or average discharge), but the area is not matching; this is likely due to dry regions (endorheic basins) within the catchment area that are differently treated in GRDC and HydroSHEDS	31
Medium	Manual	GRDC area seems wrong	Location seems correct (based on river name, station name, and/or average discharge), but the area is not matching; this is likely due to an error in the GRDC record	48
Low	Manual	Not sure, but could be ok	Two indicators could be verified, but there remain some discrepancies; and area difference 10-50%	92
Low	Manual	Location ok, but catchment not well represented	Location seems correct, but the catchment is not well represented in HydroSHEDS (yet still acceptable); possible reasons include small errors in HydroSHEDS, or upstream diversions	27
Unassigned	Manual	HydroSHEDS cannot represent situation properly	Location may be correct, but the catchment is incorrectly depicted in HydroSHEDS; possible reasons include artificial canals, braided rivers, or deltas	16
Unassigned	Manual	HydroSHEDS incorrect (above 60 degrees North)	Location may be correct, but the catchment is incorrectly depicted in HydroSHEDS; mostly occurring for regions above 60 degrees North	108
Unassigned	Manual	Unclear (station not assigned to HydroSHEDS)	Unclear situation	244

4. References

Lehner, B., Verdin, K., Jarvis, A. (2008): Hydrological data and maps based on Shuttle elevation derivatives. U.S. Geological Survey. online resource: http://hydrosheds.cr.usgs.gov/

Lehner, B., Verdin, K., Jarvis, A. (2008): New global hydrography derived from spaceborne elevation data. Eos, Transactions, AGU, 89(10): 93-94.

USGS – U.S. Geological Survey (2000): HYDRO1k Elevation Derivative Database. USGS EROS Data Center. Sioux Falls, SD. online resource: http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/HYDRO1K

Report No. 1 (May 1993)	Second Workshop on the Global Runoff Data Centre, Koblenz, Germany, 15 - 17 June, 1992.
	(17 pp, annex 73 pp)
Report No. 2 (May 1993)	Dokumentation bestehender Algorithmen zur Übertragung von Abflußwerten auf Gitternetze. (incl. an English abstract in English by the GRDC: Documentation of existing algorithms for transformation of runoff data to grid cells) / G.C. Wollenweber.
	Out of print (71 pp)
Report No. 3 (Jun 1993)	GRDC - Status Report 1992.
	(5 pp, annex 5 pp)
Report No. 4 (Jun 1994)	GRDC - Status Report 1993.
	(16 pp, annex 34 pp)
Report No. 5 (Nov 1994)	Hydrological Regimes of the Largest Rivers in the World - A Compilation of the GRDC Database.
	(275 pp)
Report No. 6 (Dec 1994)	Report of the First Meeting of the GRDC Steering Committee, Koblenz, Germany, June 20 - 21, 1994.
	(10 pp, annex 38 pp)
Report No. 7 (Jun 1995)	GRDC - Status Report 1994.
	(12 pp, annex 20 pp)
Report No. 8 (Jul	First Interim Report on the Arctic River Database for the Arctic Climate System Study (ACSYS).

Global Runoff Data Centre (GRDC)
Am Mainzer Tor 1, 56068 Koblenz, Germany
phone +49 261 1306-5224
fax +49 261 1306-5722
email grdc@bafg.de
web: grdc.bafg.de

1995)	
	(34 pp)
Report No. 9 (Aug 1995)	Report of the Second Meeting of the GRDC Steering Committee, Koblenz, Germany, June 27 - 28.
	(17 pp, annex 34 pp)
Report No. 10 (Mar 1996)	Freshwater Fluxes from Continents into the World Oceans based on Data of the Global Runoff Data Base / W. Grabs, Th. de Couet, J. Pauler.
	Out of print (49 pp, annex 179 pp)
Report No. 11 (Apr 1996)	GRDC - Status Report 1995.
	(16 pp, annex 45 pp)
Report No. 12 (Jun 1996)	Second Interim Report on the Arctic River Database for the Arctic Climate System Study (ACSYS).
	(39 pp, annex 8 pp)
Report No. 13 (Feb 1997)	GRDC Status Report 1996.
	(25 pp, annex 36 pp)
Report No. 14 (Feb 1997)	The use of GRDC - information. Review of data use 1993/1994. Status: January 1997.
	(18 pp, annex 34 pp)
Report No. 15 (Jun 1997)	Third Interim Report on the Arctic River Data Base (ARDB) for the Arctic Climate System Study (ACSYS): Plausibility Control and Data Corrections (Technical Report).

Global Runoff Data Centre (GRDC)
Am Mainzer Tor 1, 56068 Koblenz, Germany
phone +49 261 1306-5224
fax +49 261 1306-5722
email grdc@bafg.de
web: grdc.bafg.de

	(3 pp, annex 20 pp)
Report No. 16 (Aug 1997)	The GRDC Database. Concept and Implementation / J. Pauler, Th. de Couet.
	(38 pp, annex 4 pp)
Report No. 17 (Sep 1997)	Report on the Third Meeting of the GRDC Steering Committee, Koblenz, Germany June 25-27, 1997.
	(30 pp, annex 137)
Report No. 18 (Jul 1998)	GRDC Status Report 1997.
	(13 pp, annex 37 pp)
Report No. 19 (Aug 1998)	Evaluation of Statistical Properties of Discharge Data of Stations Discharging Into the Oceans - Europe and Selected World-Wide Stations / F. Portmann.
	(80 pp)
Report No. 20 (Jul 1998)	Water Resources Development and the Availability of Discharge Data in WMO Region II (Asia) and V (South-West Pacific) W. Grabs, J. Pauler, Th. de Couet.
	(51 pp, annex 68 pp)
Report No. 21 (Sep 1998)	Analysis of long runoff series of selected rivers of the Asia-Pacific region in relation with climate change and El Niño effects / D. Cluis.
	(23 pp, annex 58 pp)
Report No. 22 (April 1999)	Global, Composite Runoff Fields Based on Observed River Discharge and Simulated Water Balances / B. M. Fekete, C. Vörösmarty, W. Grabs.
	(36 pp, annex 77 pp)

Global Runoff Data Centre (GRDC)
Am Mainzer Tor 1, 56068 Koblenz, Germany
phone +49 261 1306-5224
fax +49 261 1306-5722
email grdc@bafg.de
web: grdc.bafg.de

Report No. 23 (Oct 1999)	Report of the fourth Meeting of the GRDC Steering Committee, Koblenz, Germany, 23-25 June 1999.
	(29 pp, annex 140 pp)
Report No. 24 (Nov 1999)	Use of the GRDC Data 1993-1999: A Comprehensive Summary.
	(48 pp)
Report No. 25 (Jun 2000)	GIS-related monthly Balance of Water Availability and Demand in Large River Basins - case study for the River Danube / I. Dornblut.
	Out of print (27 pp, annex 46 pp)
Report No. 26 (Nov 2000)	Modelling raster-based monthly water balance components for Europe / Carmen Ulmen.
	(133 pp)
Report No. 27 (Jul 2002)	Water Resources Management Country Profile Germany. A contribution to the Global Water Information Network WWW.GLOBWINET.ORG / R. Winnegge and T. Maurer.
	(32 pp)
Report No. 28 (Nov 2002)	Report of the Fifth Meeting of the GRDC Steering Committee, Koblenz, Germany, 25-28 June 2001.
	(36 pp, annex 300 pp)
Report No. 29 (Feb 2003)	GRDC Status Report 2002.
	(28 pp, annex 32 pp)
Report No. 30 (Dec	Development of an Operational Internet-based Near Real Time Monitoring Tool for Global River Discharge Data / T. Maurer.

Global Runoff Data Centre (GRDC)
Am Mainzer Tor 1, 56068 Koblenz, Germany
phone +49 261 1306-5224
fax +49 261 1306-5722
email grdc@bafg.de
web: grdc.bafg.de

2003)	
	(23 pp, annex 5 pp)
Report No. 31 (Oct 2004)	Globally agreed standards for metadata and data on variables describing geophysical processes. A fundamental prerequisite to improve the management of the Earth System for our all future / T. Maurer.
	(43 pp, annex 28 pp)
Report No. 32 (Nov 2004)	Detection of change in world-wide hydrological time series of maximum annual flow / Z.W. Kundzewicz, D. Graczyk, T. Maurer, I. Przymusinska, M. Radziejewski, C. Svensson, M. Szwed.
	(36 pp, annex 52 pp)
Report No. 33 (Nov 2004)	Trends in flood and low flow series / C. Svensson, Z.W. Kundzewicz, T. Maurer.
	(26 pp, annex 18 pp)
Report No. 34 (Mar 2005)	Report of the Sixth Meeting of the GRDC Steering Committee, Koblenz, Germany, 11-13 June 2003
	(27 pp, annex 85 pp)
Report No. 35 (Nov 2006)	Report of the Seventh Meeting of the GRDC Steering Committee, Koblenz, Germany, 6 - 8 July 2005
	(36 pp, annex 80 pp)
Report No. 36 (Aug 2007)	The Global Terrestrial Network for River Discharge (GTN-R) : Real-time Access to River Discharge Data on a Global Scale. 1st Interim Report / U. Looser, I. Dornblut, T. de Couet
	(24 pp, annex 42 pp)
Report No. 37 (Dec 2007)	Hydrology of the World's International River Basins: Hydrological parameters for use in global studies of international water-relations / K. Stahl (Oregon State University, Department of Geosciences, Corvallis, USA)

Global Runoff Data Centre (GRDC)
Am Mainzer Tor 1, 56068 Koblenz, Germany
phone +49 261 1306-5224
fax +49 261 1306-5722
email grdc@bafg.de
web: grdc.bafg.de

	(36 pp, annex 16 pp)
Report No. 38 (Apr 2008)	Report of the Eighth Meeting of the GRDC Steering Committee, Koblenz, Germany, 19 - 21 September 2007.
	(32 pp, annex 16 pp)
Report No. 39 (Jul 2009)	Hydrologic Information – Metadata: Semantic structure for the description of hydrologic data (GRDC Metadata Profile – Final draft) / I. Dornblut.
	Under review (April 2011)
Report No. 40 (May 2011)	Report of the Ninth Meeting of the GRDC Steering Committee, Koblenz, Germany, 23 - 25 June 2009.
	(27 pp, annex 9 pp)
Report No. 41 (Jan 2012)	Derivation of watershed boundaries for GRDC gauging stations based on the HydroSHEDS drainage network / B. Lehner (Department of Geography, McGill University, Montreal, Canada)
	(12 pp)

Global Runoff Data Centre (GRDC)
Am Mainzer Tor 1, 56068 Koblenz, Germany
phone +49 261 1306-5224
fax +49 261 1306-5722
email grdc@bafg.de
web: grdc.bafg.de