A validation of river routing networks for catchment modelling from small to large scales

Chantal Donnelly, Jörgen Rosberg and Kristina Isberg

ABSTRACT

Underpinning all hydrological simulations is an estimate of the catchment area upstream of a point of interest. Locally, the delineation of a catchment and estimation of its area is usually done using fine scale maps and local knowledge, but for large-scale hydrological modelling, particularly continental and global scale modelling, this level of detailed data analysis is not practical. For large-scale hydrological modelling, remotely sensed and hydrologically conditioned river routing networks, such as HYDRO1k and HydroSHEDS, are often used. This study evaluates the accuracy of the accumulated upstream area in each gridpoint given by the networks. This is useful for evaluating the ability of these data sets to delineate catchments of varying scale for use in hydrological models. It is shown that the higher resolution HydroSHEDS data set gives better results than the HYDRO1k data set and that accuracy decreases with decreasing basin scale. In ungauged basins, or where other local catchment area data are not available, the validation made in this study can be used to indicate the likelihood of correctly delineating catchments of different scales using these river routing networks.

Key words | drainage basins, drainage networks, river routing, watershed delineation

Chantal Donnelly (corresponding author)
Jörgen Rosberg
Kristina Isberg
Swedish Meteorological and Hydrological Institute,
Folkborgvägen 1,
Norrköping,
Sweden
E-mail: *chantal.donnelly@smhi.se*

INTRODUCTION

Global river routing networks (RRNs) are used at all scales, from routing within the global circulation model land surface schemes, to local scale hydrological models in data poor regions (e.g. Li *et al.* 2009; Xie *et al.* 2010). They are frequently used in creating global or continental scale hydrological models (e.g. Arnell 1999; Widén-Nilsson *et al.* 2007) including high-resolution continental scale models which aim to deliver data at all catchment scales (e.g. Donnelly *et al.* 2011). A RRN is a raster data set which includes information on the elevation, flow directions, and links between grid elements. For each gridpoint, the RRN includes the accumulated upstream area or number of accumulated upstream gridpoints. The RRN is created by analysing a digital elevation model (DEM) using the premise that water flows from higher to lower elevations and often refined using information on the locations of known rivers and streams, for example from maps, using the 'stream-burning' method (Maidment 1996; Wesseling *et al.* 1997). Stream burning is used to enforce known river courses into an elevation surface so that errors in the elevation surface do not affect the known course of these streams. Other conditionings of the elevation surface that can be made include deepening of open water surfaces, weeding of the coastal zone (to avoid a coastal barrier to water flow), moulding of valley courses, sink filling and carving through barriers (e.g. Lehner *et al.* 2006).

Validation of RRNs is often restricted to comparison of the resulting drainage network to vectorised rivers on available digital maps at varying scales, as well as digitised lake and wetland locations (e.g. Verdin & Greenlee 1996; Döll & Lehner 2002; Lehner *et al.* 2006). For the DDM30 product, Döll & Lehner (2002) published values for upstream area at a number of gauging stations, major lakes and reservoirs are also taken into account in the manual conditioning. Other RRNs use comparisons with major river basin outlines, vectorised rivers and existing RRNs as part of their validation (e.g. Verdin & Greenlee 1996;

doi: 10.2166/nh.2012.341

Renssen & Knoop 2000; Lehner *et al.* 2006; Yamakazi *et al.* 2009).

It can generally be concluded that the validation of the RRNs is restricted to ensuring the calculated drainage network correlates with known drainage networks, for example river locations vectorised on maps or by matching the outlines of regional and large-scale catchments. It is impractical to collate quality, locally checked catchment outlines for all rivers globally. The increasing use of readily available global RRNs to provide input data for hydrological modelling at smaller scales in data poor regions and high-resolution modelling at large scales demonstrates the need for an evaluation of frequently used RRNs at varying catchment scales.

In this study, the accumulated upstream area calculated by two RRNs, HYDRO1K and HydroSHEDS (Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales) was compared with published values of upstream areas at 1,007 gauging stations within Europe. This gave an indication of how well the RRNs simulate the actual catchment area for catchments ranging from 200 to 800,000 km^2.

METHODS AND DATA

The accumulated upstream area in each grid point of two gridded RRNs, HYDRO1k (Verdin & Greenlee 1996) and HydroSHEDS (Lehner *et al.* 2006), were evaluated by comparing the accumulated upstream area calculated by the RRN with actual values of the upstream area published for gauging stations. Because the purpose of this analysis was to set up a pan-European hydrological model (E-HYPE) (Donnelly *et al.* 2010), the evaluation of the RRNs was restricted to continental Europe. The analysis was also restricted to latitudes below 60 degrees as HydroSHEDS data north of this latitude were not available at the time the study was made.

HYDRO1k is a freely available 1 km resolution RRN developed by the USGS to provide a globally consistent hydrographic data set. It is based on a 30' DEM (GTOPO30) and includes a hydrologically corrected DEM, derived flow directions, flow accumulations, slope, aspect, and a compound topographic (wetness) index. Hydrological conditioning included identification of natural sink features and correction of non-existing sinks, verification of waterways against the Digital Chart of the World (DCW) drainage cover (Defence Mapping Agency 1992) and other available mapping sources (Verdin & Greenlee 1996).

HydroSHEDs is based on a newer, higher resolution DEM called the Shuttle Radar Topography Mission (SRTM) (NASA/JPL 2005) and was developed by the Conservation Science Program of the World Wildlife Fund and partners to provide both regional and global scale hydrographic information for a range of applications. Data made available include a void-filled DEM, the hydrologically conditioned elevation model, drainage directions, flow accumulation, the river network and drainage basin delineations at 3 arc-sec (~90 m at equator), 15 arc-sec (~500 m at equator) and 30 arc-sec resolutions. Flow accumulation and the river network are only available at the larger resolutions and for this study the 30 arc-sec resolution data were analysed rather than the 15 arc-sec due to the large extent of the study area. Also since the large sub-basin area to be extracted for the hydrological model made the finer details in 15 arc-sec resolution insignificant. As the larger resolution data sets from HydroSHEDS are based on the finer scale analysis, this is not expected to affect the results for the scale of catchments considered in this study. Hydrological conditioning included stream burning using the ARCWORLD river layer (ESRI 1992) and the Global Lakes and Wetlands Database (GLWD) (Lehner & Döll 2004), deepening of open water surfaces, weeding of the coastal zone (to avoid a coastal barrier to water flow), moulding of valley courses, sink filling and carving through barriers (Lehner *et al.* 2006). HydroSHEDS was expected to be more accurate than HYDRO1k due to the higher resolution of the underlying DEM and the more rigorous conditioning made to the DEM.

The HydroSHEDS database gives the number of upstream grid points rather than upstream area for each grid point. This needed therefore to be converted to an area. Since the extent of the watershed is not known at this point, the latitude at the actual grid square (i.e. the downstream outlet of the catchment) is used for the conversion. This introduces a potential error in that the area of a grid square changes depending on latitude. For the Rhine River, which had the largest variation in latitude over the catchment, the error was approximately 5% of the actual area. The error introduced by this assumption was therefore considered acceptable for this study. It should be noted that outside the study area, Europe, there exist many rivers with much larger variations in latitude (e.g. Nile, Murray-Darling) for which the error would be larger.

The accumulated area in the RRN was evaluated against known points (in this case, gauging stations) measuring catchments of scales ranging from 200 km^2 to over 800,000 km^2. Gauging stations were chosen as validation points because (a) large international databases of gauging station data and metadata which include upstream catchment area exist, (b) these databases generally consist of nationally approved data which are then collated internationally (i.e. from local sources), and (c) a wide range of catchment scales is represented in these databases. Lehner (2012) derived watershed boundaries from HydroSHEDS for 7,500 Global Runoff Data Centre (GRDC) (Global Runoff Data Centre 2009b) stations worldwide using the GRDC metadata and HydroSHEDS. In that study, the focus was on making a correct catchment delineation for the GRDC gauging stations. Here, a much larger number of gauging stations covering a wider range of catchment sizes across Europe are included in the analysis and the hypothesis that the RRN correctly estimates the upstream catchment area is tested.

Gauging station locations and metadata including upstream catchment area were downloaded from three international databases:

- GRDC, an international archive of runoff data and metadata operating under the auspices of the World Meteorological Organization. Data provided to the GRDC should be verified and released by the regional or national services that collect the data, but data from other sources may be accepted (Global Runoff Data Centre 2009b). Due to the station selection criteria (Global Runoff Data Centre 2009b), catchments represented in this database are often larger in scale. A total of 1,225 stations were downloaded.
- European Water Archive (EWA), a European archive of runoff data and metadata created by the EURO-FRIEND network of the UNESCO International Hydrological Program. It is hosted by the GRDC, but is a separate database to the GRDC database. Data are provided voluntarily by various providers including regional and national data collection services, but also other sources (Global Runoff Data Centre 2009a). A total of 3,834 stations were downloaded.
- Baltex Hydrological Database (BHD) is an archive of runoff data and metadata for rivers with discharge to the Baltic Sea. These data were originally collected for the BALTEX project and are hosted by the Swedish Meteorological and Hydrological Institute. Data were provided by national hydrological authorities (BHDC 2009). A total of 593 stations were downloaded.

After excluding stations occurring in multiple databases, catchments <200 km^2, those with time-series less than 10 years (the resulting database was also to be used for hydrological model validation and calibration) and closely located stations (a threshold of 80% unique catchment area was used to ensure a uniform distribution of validation stations), a total of 1,386 stations remained for use in the validation database. The minimum catchment size was set to 200 km^2 to limit the analysis to a realistic resolution of a continental hydrological model for which the RRN might be used, in this case the HYPE model (Donnelly *et al.* 2010). The largest scale represents the catchment size of the largest river basin, in this case the Danube River.

The gauging station metadata used to verify the RRNs is the station location and the catchment area upstream of the station. Most of the data in these databases come from the regional and national services which collect the data. The quality of the data therefore varies depending on the data provider. Station locations are requested from the GRDC and the EWA database (Global Runoff Data Centre 2009a, b) in decimal degrees stored at up to 6 decimal places and the reference system used; however, sometimes this location is determined directly from a map of unknown scale or projection, rather than local survey or GPS. The method by which station area is determined also varies and typographic errors also occur in the databases (Looser, personal communication). A procedure to determine which data points are useful in analysis was therefore required.

Plotting the station points at their given locations against the RRNs showed that there were inaccuracies in many of the coordinate sets. The locations of station points were therefore checked to ensure that the coordinates given for the point are located correctly, e.g. on the waterway of interest or correctly located upstream or downstream of a confluence. This was done in the three steps outlined below.

1. To quality check the catchment area given by the databases, the mean annual discharge volume was divided by the given catchment area to give an estimated mean annual precipitation minus evapotranspiration over the catchment. This value was then plotted on a map using a colour scale for each gauging station point. Any points which deviated significantly from those surrounding it were checked manually. The manual check involved checking if the precipitation and evapotranspiration for the location may actually be correct and if not, searching national databases for upstream catchment area. If these verifications could not be made, the points were discarded.
2. The published catchment area was then compared with the accumulated area in the RRN. HydroSHEDS was used for this comparison as it was expected to be more accurate than HYDRO1k. If the discrepancy between the catchment area and the accumulated grid area was larger than an acceptable threshold, the point was preliminarily moved to the grid location that has the smallest difference between this known value and the area given by the flow accumulation grid, within a certain radius that can be set for each point. In this case, 1.5 arcminutes was used. Any corrections made are printed to a separate shapefile for manual check and log.
3. A manual check was made, both of all the stations moved in step 2, but also those for which large discrepancies between the published catchment area and accumulated grid area remained. Therefore, no stations were permanently moved until the relocation could be manually verified. This was done by visually checking the old and new station location on a map with a vectorised river map for Europe from the European Environmental Agency called Named Rivers (Crouzet & Simonazzi 2008), river locations on Google Maps and Google Earth (http://maps.google.com and http://earth.google.com) and where necessary, searching national databases to find more information to verify the location of the station. If the location could not be manually verified, the point was discarded.

Note that as a result of this procedure, only stations for which a secondary check of the station location could be made remain in the database. It is acknowledged that there may also remain some stations incorrectly located but for which both the published area and the HydroSHEDS area are similar but incorrect, however, the number of incorrect truthing points in relation to correct truthing points should be low following these checks.

Following these checks, a database of 1,007 stations remained to validate the accumulated upstream area in the RRNs. Figure 1 shows the geographical distribution of these stations and the distribution of catchment sizes among the stations. Most of the stations are from catchments less than 50,000 km^2 with the majority of the order 500–5,000 km^2.

In order to assess the accuracy of the upstream area suggested by the RRNs, a geometrical relative error (Equation (1)) comparing the accumulated upstream area in the RRN with the published upstream area was calculated.

For

$$\begin{aligned} A_{calc} \geq A_{pub}, \quad E = A_{calc}/A_{pub} - 1 \\ A_{calc} < A_{pub}, \quad E = 1 - A_{pub}/A_{calc} \end{aligned} \tag{1}$$

where A_{calc} = calculated catchment area from the RRN, A_{pub} = published catchment area and E = error. This symmetrical error term gives the same value of the error (with opposite sign) regardless of whether the calculated or the published area is larger. It was chosen over the common relative error because this relative error is limited to –100% for underestimated station areas, but may approach infinity for overestimated areas.

RESULTS AND DISCUSSION

Figure 2 shows the distribution of the geometrical relative error for stations located in the mouths of Europe's 54 largest rivers. This represents the ability of the RRNs to delineate entire river basins ranging in size from 1,000 to 800,000 km^2. This subset of the database is of interest

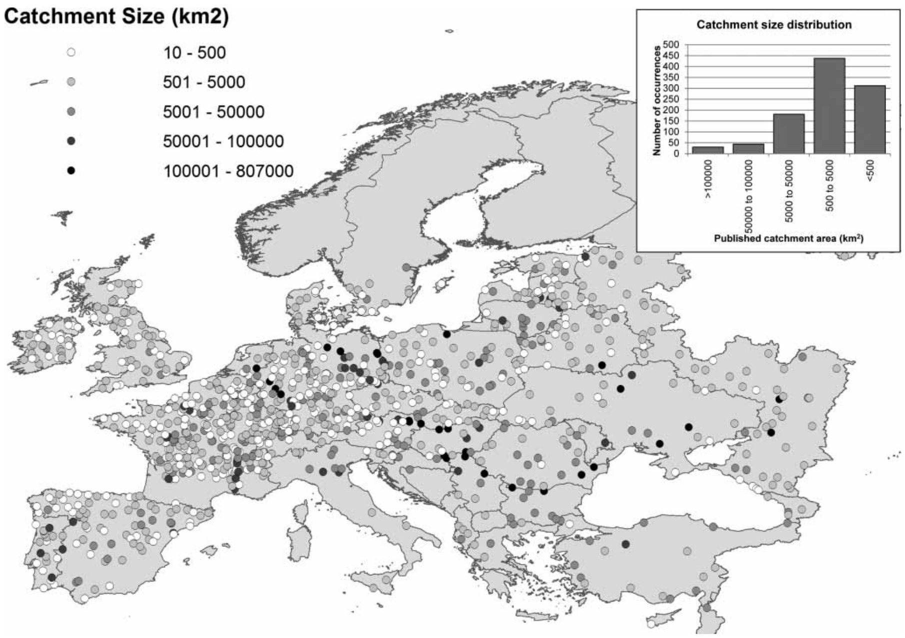

Figure 1 | Map of gauging station distribution and histogram of catchment size distribution in final validation database.

because the hydrological conditioning and validation of the RRNs was most likely made for at least some of these rivers which can easily be located on most European maps. The HydroSHEDS RRN estimates 70% of these river catchments to within 5% and over 90% to within 10% of the published area. HYDRO1k only manages to estimate about 60% of the river catchment areas to within 10% of the published area.

Figure 3 shows the geometrical error vs the catchment size for all of the gauging stations in the database as compared to the HYDRO1k and HydroSHEDS RRNs. Note that only stations with errors between +200 and

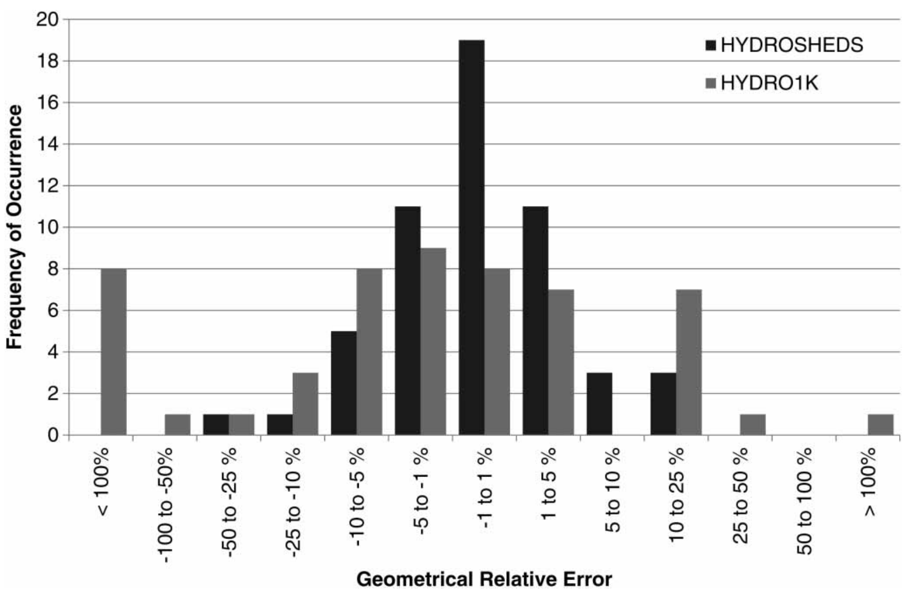

Figure 2 | Distribution of error for stations in river mouths.

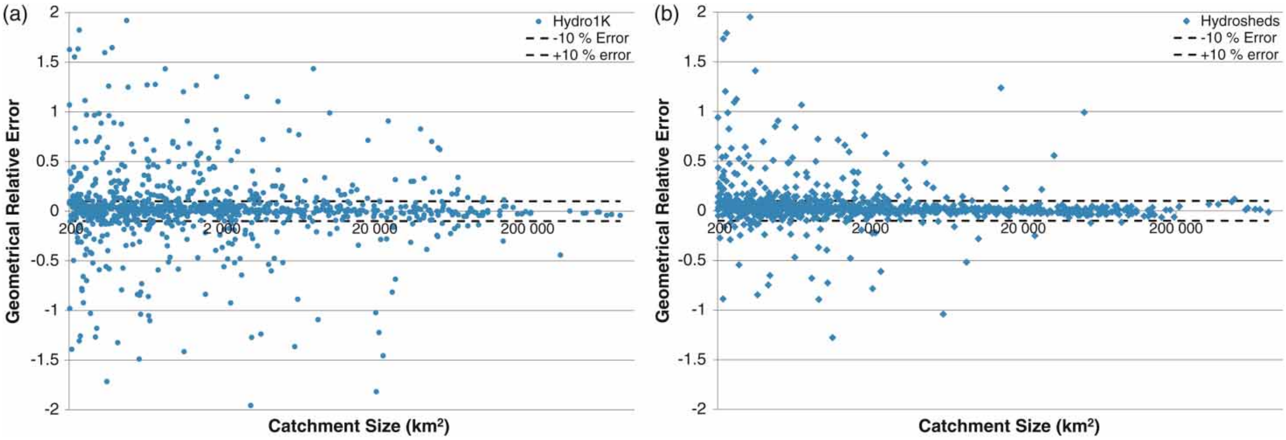

Figure 3 | Geometrical relative error vs catchment size for (a) HYDRO1k and (b) HydroSHEDS.

−200% error are shown to make the results more clearly discernible. There were 34 stations exceeding this error for HydroSHEDS of which the largest had a catchment area of 4,000 km^2. For HYDRO1k the outliers are more significant, with 114 stations not shown including 8 river mouth stations and catchment areas of up to 125,000 km^2. This indicates some shortcomings of the HYDRO1k data set even for large-scale catchment delineation.

For both RRNs the majority of catchment areas are estimated to within ±10%, nevertheless there remain large numbers of stations for which there are significant errors in the calculated catchment area. It can be seen that HydroSHEDS does a better job in estimating catchment area than HYDRO1k. This result was expected because HydroSHEDS uses a higher-resolution DEM, more sophisticated hydrological conditioning techniques and more data for hydrological conditioning.

From Figure 3 it can be concluded that the error in the estimated catchment area decreases with increasing catchment size. It could also be noted that the decrease in error with increasing catchment size was most pronounced for positive errors, i.e. where the RRN overestimates the catchment area.

Errors in the RRN are mostly caused by insufficient resolution (i.e. valleys, ridge tops, gorges are missed) and remain where insufficient hydrological conditioning has been made. Hydrological conditioning is generally made for readily available large-scale data; i.e. it is not practical to collect maps showing small creeks (first and second order rivers) for the entire world and condition to these. Therefore, it is expected that more errors remain for the smaller catchments. Overestimation errors occur where the RRN overestimates the catchment area. This is because upstream grid points which should flow into another catchment are linked into the catchment of interest. As the size of a catchment increases, the relative size of the incorrectly added area becomes smaller.

On the other hand, it is possible for the RRN to estimate a much smaller area than it should, even where the catchment area is large. These underestimation errors therefore have less correlation to the catchment area, i.e. large underestimation errors still occur for large catchment sizes. These underestimation errors occur less often than overestimation errors (e.g. for HydroSHEDS there were three times as

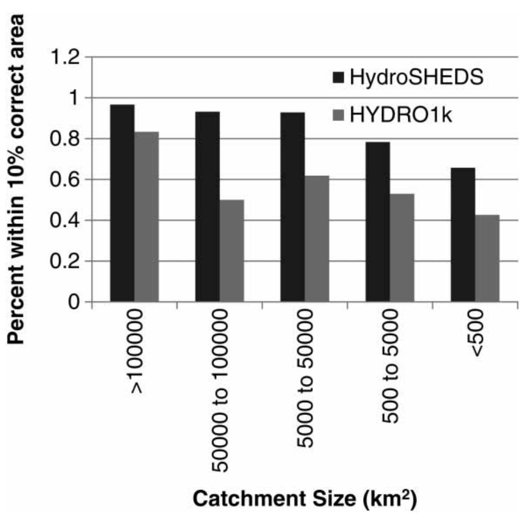

Figure 4 | Percentage of catchments of various scales, for which the upstream area could be estimated to within ±10% of the published area.

many catchments with overestimated area than underestimated).

Figure 4 shows the percentage of stations for which the upstream catchment area could be estimated to within 10% of the published catchment area. The threshold of 10% was chosen to be large enough to allow for errors in the published catchment area or station location but small enough to reflect a minimum accuracy required of hydrological modellers. In this study it can be seen for HydroSHEDS that more than 90% of the catchments >5,000 km^2 could be estimated to within 10% of the published catchment area. The threshold of 10% error in the catchment area was also used by Lehner (2012) for categorising the accuracy of delineated catchments from HydroSHEDS as low. In their catchment delineation of over 7,500 catchments worldwide, they found that only about 84% of the catchments could be delineated within this error threshold. This may be due to including catchments covering a wider range of hydrographic conditions worldwide. Nevertheless, in both these studies there is still some uncertainty regarding the truthing data and this should be taken into account when considering these results.

It can also be seen in Figure 3(b) that the number of incorrectly estimated catchment areas decreases significantly for catchments greater than approximately 5,000 km^2. It is therefore suggested that this be a limiting resolution for which catchments should be delimited using HydroSHEDS, where no local data for correction are available. Even given the uncertainties discussed above, this threshold is suggested because whether or not the errors come from the truthing data or the RRN, it is essential that the user is aware of the greater probability of making errors when working with continental and regional scales databases for making smaller scale studies.

For HYDRO1k, it is more difficult to suggest such a limit. Even for catchments of scale 50,000–100,000 km^2, only half of the catchment areas could be estimated within acceptable limits. Nevertheless, the HYDRO1k database has the advantage of covering latitudes above 60 degrees north and may in fact be more accurate in other regions than the European continent, as the underlying DEM, GTOPO30, is based on different local and regional mapping products (USGS EROS Data Center 1996).

The analysis presented here has large significance when a RRN is used to define catchments for which discharge or other variables are to be predicted for ungauged basins (defined here as points for which the upstream drainage basin shape and area is unknown) or for multi-catchment

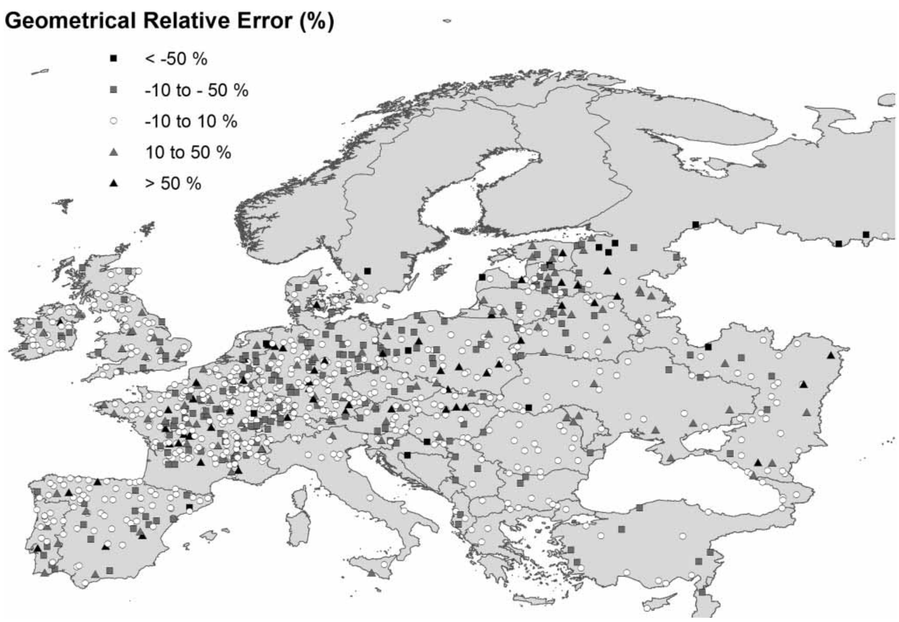

Figure 5 | Map showing location of gauging stations and the geometrical relative error of the catchment area for each station.

model domains where it may be impractical to check the routing of all points. It is suggested that the validation results presented here may be used to define the uncertainty in correctly defining a basin of a certain scale using a HydroSHEDS or HYDRO1K. For example, from Figure 4 it can be seen that about 65% of catchment areas <500 km^2 could be correctly estimated using HydroSHEDS. If the validation points used in this study are representative for all of Europe, it may be concluded that there is a 35% risk that a new catchment of this size could be incorrectly delineated. Of course, this uncertainty should also depend on how flat the basins are and the geology of the region (which governs the existence of small terrain features such as gorges). Nevertheless, it can be seen in Figure 5 that there is no obvious regional variation in estimated upstream area error. The quantitative results are valid for the studied domain, Europe, although the qualitative conclusions are of interest for all regions.

It should also be noted that this study does not evaluate how well a basin is delineated. Even when the calculated upstream area is correct, the actual location of that area may not be. This is easily checked for large basins for which basin delineations are readily available in global databases, but more difficult for small basins (where the errors have been shown to be larger) and in data poor regions where the RRN is taken to be correct. This has implications for distributed hydrological modelling where the correct distribution of landcover, soil types, elevations and precipitation are necessary to correctly simulate hydrological variables.

Finally, it is noted that errors may still remain in gauge location and published upstream catchment area which may affect these results. It is believed that the manual checks made ensure that the larger majority of this metadata is correct and that given the large amount of points against which the RRNs were tested, that the overall results are robust.

CONCLUSIONS

The accuracy of accumulated upstream areas in two widely used RRNs was validated. This was done by comparing calculated upstream areas with known upstream areas for many gauging stations at varying scales. The results show that the newer HydroSHEDS database is more accurate than HYDRO1k. For HydroSHEDS accuracy increased with increasing catchment size and it is suggested that estimated catchment areas are sufficiently correct for catchments >5,000 km^2. A similar threshold could not be suggested for HYDRO1k. Nevertheless, HYDRO1k is still useful due to its truly global coverage, particularly for latitudes above 60 degrees north.

ACKNOWLEDGEMENTS

The authors would like to thank the following funding agencies for supporting studies which resulted in this analysis: the European Union's FP7 (GEOLAND2, ECLISE, and OPERR) and the BONUS program (ECOSUPPORT project).

REFERENCES

Arnell, N. W. 1999 A simple water balance model for the simulation of streamflow over a large geographic domain. *J. Hydrol.* **217**, 314–335.

BHDC 2009 *Baltex Hydrological Database*. http://www.smhi.se/sgn0102/bhdc/ (accessed 10 February 2009).

Crouzet, P. & Simonazzi, W. 2008 *Building the EEA European Catchment and Rivers Network System (ECRINS) from CCM v2.1. Part 1: Setting and Implementing Rules and Producing the Watersheds Layer*. European Environmental Agency, Copenhagen, Denmark. 68 pp.

Defense Mapping Agency 1992 *Digital Chart of the World*. Government Printing Office, Washington, DC.

Döll, P. & Lehner, B. 2002 Validation of a new global 30-min drainage direction map. *J. Hydrol.* **258**, 21–213.

Donnelly, C., Dahné, J., Rosberg, J., Strömqvist, J., Yang, W. & Arheimer, B. 2010 High-resolution, large-scale hydrological modelling tools for Europe. *IAHS Publ.* **340**, 553–561.

Donnelly, C., Strömqvist, J. & Arheimer, B. 2011 Modelling climate change effects on nutrient discharges from the Baltic Sea catchment: processes and results. Water quality: current trends and expected climate change impacts. *Proceedings of symposium H04, IUGG2011*, Melbourne, Australia, July 2011. IAHS Publ. 348, 2011.

ESRI 1992 *ArcWorld 1:3 Mio. Continental Coverage*. Environmental Systems Research Institute, Redlands, CA. Data available on CD.

Global Runoff Data Centre 2009a *European Water Archiver (EWA)/Global Runoff Data Centre*. Federal Institute of Hydrology, Koblenz, Germany.

Global Runoff Data Centre 2009b *River Discharge Data*. Federal Institute of Hydrology, Koblenz, Germany.

Lehner, B. 2012 Derivation of watershed boundaries for GRDC gauging stations based on the HydroSHEDS drainage network. GRDC Report Series, Report 41, 12 pp. Available at http://www.bafg.de/GRDC/EN/02_srvcs/22_gslrs/222_WSB/methodology_Lehner.html (accessed 29 April 2013).

Lehner, B. & Döll, P. 2004 Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* **296**, 1–22.

Lehner, B., Verdin, K. & Jarvis, A. 2006 *HydroSHEDS Technical Documentation Version 1.0*. World Wildlife Fund US, Washington, DC. Available at http://hydrosheds.cr.usgs.gov.

Li, L., Hong, Y., Adler, R., Policelli, F., Habib, S., Irwn, D., Korme, T. & Okello, L. 2009 Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa. *Nat. Hazards.* **50**, 109–123.

Maidment, D. R. 1996 GIS and hydrological modeling: an assessment of progress. *Presented at the Third International Conference on GIS and Environmental Modeling*, Santa Fe, New Mexico, 20–25 January 1996.

NASA/JPL 2005 *SRTM Topography* (SRTM documentation; ftp://e0srp01u.ecs.nasa.gov/srtm/version2/Documentation/SRTM_Topo.pdf), NASA, FL. 8pp.

Renssen, H. & Knoop, J. M. 2000 A global river routing network for use in hydrological modeling. *J. Hydrol.* **230**, 230–243.

USGS EROS Data Center 1996 *GTOPO30 Documentation*. US Geological Survey EROS Data Center. Available at http://www1.gsi.go.jp/geowww/globalmap-gsi/gtopo30/README.html.

Verdin, K. L. & Greenlee, S. K. 1996 *HYDRO1k Documentation*. US Geological Survey EROS Data Center. Available at http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/HYDRO1k.

Wesseling, C. G., van Deursen, W. P. A. & de Wit, M. 1997 Large scale catchment delineation: a case study for the river Rhine Basin. In: *Third Joint European Conference & Exhibition on Geographical Information* (S. Hodgson, M. Rumor & J. J. Harts, eds). Vienna, Austria, pp. 487–496.

Widén-Nilsson, E., Halldin, S. & Xu, C.-Y. 2007 Global water balance modelling with WASMOD-M: parameter estimation and regionalisation. *J. Hydrol.* **340**, 105–118.

Xie, H., Nkonya, E. & Wielgosz, B. 2010 Evaluation of the SWAT Model in Hydrological Modelling of a Large Watershed in Nigeria. *Proceedings Water Resource Management Africa 2010*. ACTA Press, Calgary, Canada.

Yamakazi, D., Oki, T. & Kanae, S. 2009 Deriving a global river network map at flexible resolutions from a fine-resolution flow direction map with explicit representation of topographical characteristics in sub-grid scale. *Hydrol. Earth Syst. Sci. Discuss.* **6**, 5019–5046.

First received 17 October 2011; accepted in revised form 2 August 2012. Available online 27 October 2012