Global-scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density

A. Schneider, A. Jost, C. Coulon, M. Silvestre, Sylvain Théry, Agnès Ducharne

Global-scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density

A. Schneider[1] iD, A. Jost[1] iD, C. Coulon[1] iD, M. Silvestre[2] iD, S. Théry[2], and A. Ducharne[1] iD

[1]Sorbonne Universités, UPMC Univ Paris 06, CNRS, EPHE, UMR 7619 METIS, Paris, France, [2]Sorbonne Univ Paris 06, CNRS, FR 3020 FIRE (Fédération Île-de-France de Recherche sur l'Environnement), Paris, France

Abstract To improve the representation of surface and groundwater flows, global land surface models rely heavily on high-resolution digital elevation models (DEMs). River pixels are routinely defined as pixels with drainage areas that are greater than a critical drainage area (A_{cr}). This parameter is usually uniform across the globe, and the dependence of drainage density on many environmental factors is often overlooked. Using the 15″ HydroSHEDS DEM as an example, we propose the calibration of a spatially variable A_{cr} as a function of slope, lithology, and climate, to match drainage densities from reference river networks at a 1:50,000 scale in France and Australia. Two variable A_{cr} models with varying complexities were derived from the calibration, with satisfactory performances compared to the reference river networks. Intermittency assessment is also proposed. With these simple tools, river networks with natural heterogeneities at the 1:50,000 scale can be extracted from any DEM.

1. Introduction

Precise characterization of river geometry is crucial for many applications related to river hydraulics and has gained a lot from the advances in digital elevation models (DEMs) processing since *Hutchinson* [1989] first proposed the "stream burning" technique to correct the location of extracted streams. At large scales, due to the advent of very high resolution global DEMs such as HYDRO1k at 1 km [*Verdin and Greenlee*, 1998] and HydroSHEDS at 30″ and 15″ [*Lehner et al.*, 2008], priority has been given to producing correct flow direction maps, including coarser resolutions used in global runoff routing models [*Graham et al.*, 1999; *Fekete et al.*, 2001; *Döll and Lehner*, 2002; *Wu et al.*, 2012].

Proper characterization of stream length from DEMs has received less attention despite its broad influence on water sciences (e.g., on erosion and sediment transport [*Moore and Burch*, 1986]; riverine water quality, which is strongly controlled by the upstream residence time [*Billen et al.*, 2009]; and the buffering effect of groundwater on extreme flows, which partly depends on the distance from the recharge zones to the rivers [*Brutsaert and Lopez*, 1998]). A major issue is identifying the smallest streams, called the "aqua incognita" by *Bishop et al.* [2008]. This is true in the field because of stream intermittency, artificial ditches, or hydraulic obstructions in flat areas. Additional problems arise when implementing traditional methods to map the "blue lines," usually from airborne or satellite imagery, because of insufficient resolution, vegetation masking, or scarce field data [*Tarboton et al.*, 1991; *Lehner and Döll*, 2004; *Benstead and Leigh*, 2012; *Persendt and Gomez*, 2016]. It is believed that river networks extracted from very high resolution DEMs are more comprehensive, assuming that they are properly calibrated [*Benstead and Leigh*, 2012; *Stein et al.*, 2014].

Many methods exist to extract river networks from DEMs, mostly from a geomorphological reference [*Tarboton et al.*, 1992; *Montgomery and Foufoula-Georgiou*, 1993; *Heine et al.*, 2004; *Pelletier*, 2013]. However, the most widely used method at small and regional scales, by far, simply relies on flow direction and a critical drainage area (A_{cr}) that represents the minimum upstream drainage area required to initiate a river [*O'Callaghan and Mark*, 1984]. In this classical framework, extracted rivers are shorter if the A_{cr} is larger, which results in a smaller drainage density (δ), defined as the total river length inside a watershed divided by its area [*Horton*, 1932].

Drainage density, δ, gives a macroscale measure of stream length and is often used to quantitatively evaluate watershed properties [*Strahler*, 1957; *Vogt et al.*, 2007; *Dingman*, 2015]. In natural river networks, δ is spatially variable [*Tucker et al.*, 2001; *Vogt et al.*, 2007], since river initiation depends on climate, slope, lithology, soil

properties, and vegetation cover [*Tarboton et al.*, 1992; *Montgomery and Foufoula-Georgiou*, 1993; *Vogt et al.*, 2003; *Colombo et al.*, 2007; *Sangireddy et al.*, 2016]. However, given the lack of sufficient information to constrain A_{cr}, it is common to use a single value for river network extraction in a given domain, resulting in a rather uniform δ given the long-established link between A_{cr} and δ [*Tarboton et al.*, 1992].

In global hydrographic data sets such as HYDRO1k, HydroSHEDS, and AQUAMAPS (a river network product derived from 15″ HydroSHEDS) [*Food and Agriculture Organization (FAO)*, 2014], the A_{cr} used to extract river networks are uniform across continents, and their values are arbitrarily chosen to limit the number of small streams for large-scale applications: 1000 km^2 in HYDRO1k and the global version of AQUAMAPS, 100 km^2 for the continental version of AQUAMAPS, and approximately 25 km^2 (at the equator) in HydroSHEDS (100 15″ pixels). These discrepancies highlight the fact that each of the corresponding networks is a "hypothetical" river network, constrained in its extent by the chosen A_{cr}. The above values all result in δ under 0.2 km^{-1}, which is much lower than the values obtained from detailed observations in small watersheds (40 up to 1000 km^2), which show drainage densities between 0.5 and 1.6 km^{-1} [*Horton*, 1945; *Brutsaert and Nieber*, 1977; *Zecharias and Brutsaert*, 1988; *Brutsaert and Lopez*, 1998]. This led *Raymond et al.* [2013] to define their own global river network based on the 15″ HydroSHEDS and a smaller A_{cr} of 10 km^2 for estimating carbon sinks and sources for inland waters (resulting in global mean δ of 0.28 km^{-1}).

Similarly, several studies have analyzed the spatial dependence of δ on environmental parameters at the continental scale [*Colombo et al.*, 2007; *Vogt et al.*, 2007; *Luo et al.*, 2016]. *Vogt et al.* [2007] evaluated drainage density over Europe (CCM2 data set, second version of *Colombo et al.* [2007] work), with a mean value of 0.31 km^{-1} based on a 100 m resolution DEM and 10 different A_{cr} values, ranging from 0.72 to 12.80 km^2, defined from the association of landscape classes to subjective scores of valley dissection potential. *Luo et al.* [2016] used a geomorphological detection method and a 30 m resolution DEM to extract the land dissection density (considered equivalent to drainage density) over the United States, with values ranging from 0 to 5 km^{-1}, and a clear spatial dependence on climate, lithology and several terrain-based attributes.

Based on this analysis, our goal is to propose spatially variable A_{cr} values for a simple extraction of river networks from large-scale DEMs, as well as a first-order assessment of intermittent streams, using the 15″ HydroSHEDS data as an example. The heterogeneities of A_{cr} and the resulting drainage density, δ, are linked to environmental parameters (i.e., slope, lithology, and climate) to match drainage densities from reference river networks at the 1:50,000 scale in France and Australia. An independent evaluation conducted against reference river networks from the United States and Brazil is discussed before generalizing the method to extract 15″ river networks consistent with 1:50,000 blue lines across the continents.

2. Data Description

We used ArcGIS (version 10.3.1) tools to process several hydrologically corrected layers from the 15″ (approximately 500 m at the equator) resolution HydroSHEDS database [*Lehner et al.*, 2008] including flow directions and flow accumulation for river network extraction, and elevation to calculate the local slope, using the neighborhood method [*Burrough and McDonell*, 1998] corrected for latitudinal distortions. Global lithology data are from *Hartmann and Moosdorf* [2012], with an average scale of 1:3,750,000 and 12 classes mapped in Figure S4 in the supporting information (Pi = Intermediate plutonic rocks; Mt = Metamorphic rocks; Va = Acid volcanic rocks; Ss = Siliciclastic sedimentary rocks; Vi = Intermediate volcanic rocks; Pb = Basic plutonic rocks; Pa = Acid plutonic rocks; Vb = Basic volcanic rocks; Py = Pyroclastics; Sm = Mixed sedimentary rocks; Su = Unconsolidated sediments; and Sc = Carbonate sedimentary rocks). Climate is represented by the multiannual mean of total precipitation (1980–2009) raster at 0.5° resolution from the WFDEI Water and Global Change Forcing Data methodology applied to ERA-Interim data set, including correction by the GPCC (Global Precipitation Climatology Centre) data [*Weedon et al.*, 2014].

The reference river networks were acquired in vector format from four national databases, from Australia [*OSDM*, 2015] (Geofabric) and France [*IGN*, 2014] (CARTHAGE) at a 1:50,000 scale, from the United States [*USGS*, 2015] (NHD) at a 1:24,000 scale, from Brazil [Instituto Brasileiro de Geografia e Estatística (*IBGE*), 2015] at variable scales between 1:25,000 and 1:250,000. They were chosen for their quality and reported scale of approximately 1:50,000, despite some inconsistencies, as analyzed in the supporting information Text S1. For additional comparison with a "hypothetical" global river network, we used the global
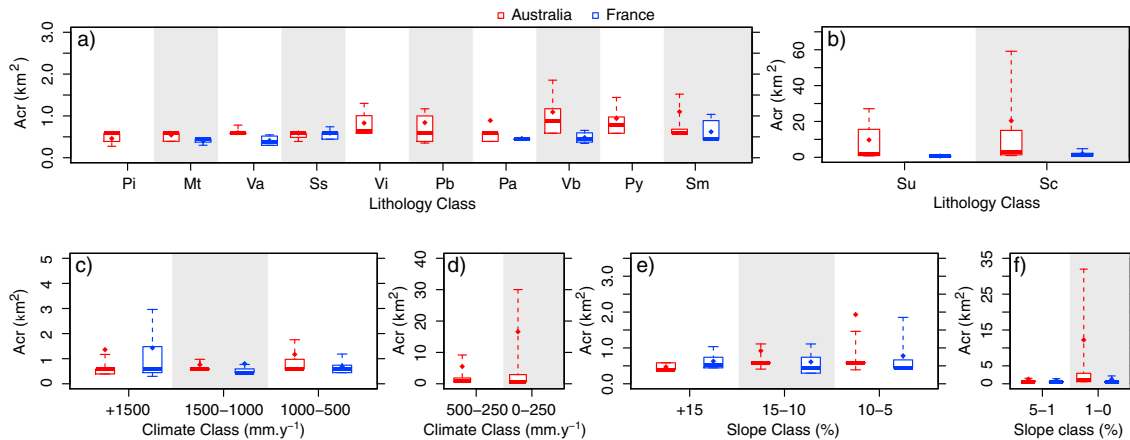
Figure 1. A_{cr} calibration results in Australia (red) and France (blue) per environmental class. (a and b) The lithology, (c and d) climate, and (e and f) slope classes. The bottom and top of the box plots represent the first and third quartiles, the middle bar gives the median, and dots indicate the mean values. Whiskers show the 10th and 90th percentiles.

AQUAMAPS [*FAO*, 2014] network, further called Food and Agriculture Organization (FAO), based on the 15″ HydroSHEDS data and a constant A_{cr} of 100 km^2, which also provides an intermittency assessment (details in Text S1).

As detailed in Text S2, all drainage density analyses and calculations were made using a 7.5′ grid (grid cells approximately 150 km^2 at equator) over the continents. Drainage densities (km^{-1}) were calculated as the total river length (km) inside a 7.5′ grid cell, divided by the corresponding land area (km^2). Each 7.5′ grid cell was characterized by a single value for δ and each of the selected environmental parameters (dominant class for lithology, areal average of total precipitation (mm yr^{-1}), and slope (%), each reclassified into five classes) (Table S1 and Figures 1c, 1d, and S6 for precipitation and Figures 1e, 1f, and S7 for the slope).

3. A_{cr} Calibration and Model Selection in France and Australia

The spatial intersection of the selected environmental parameters (12 lithology classes, 5 precipitation classes, and 5 slope classes) defines 300 environmental classes. In each of them, A_{cr} was calibrated by minimizing the bias between the corresponding drainage density (δ) and δ_{Ref} (drainage density with a reference river network). This minimization was performed independently in each country, by testing 300 extracted river networks, defined by a wide range of A_{cr} (from 0.3 to 200 km^2, see Text S2). To avoid calibration errors due to significant heterogeneities inside the 7.5′ grid cells, calibration was only performed in grid cells where the dominant lithology class covered more than 70% of the grid cell (representing 81% of the continents).

To permit independent validation, calibration was restricted to France and Australia, in which the reference river networks share the same 1:50,000 scale and show few inconsistencies. Together, these two countries also encompass all the precipitation, lithology, and slope classes of the currently available HydroSHEDS domain (56°S to 60°N). The resulting calibrated values in the two countries are given in Data Set S1.

We found much smaller A_{cr} values than the values used to define stream networks in the global-scale databases (i.e., HYDRO1k, AQUAMAPS, and HydroSHEDS, with A_{cr} values between 25 and 1000 km^2), and a similar dependency on lithology, climate, and slope in both countries (Figure 1). In agreement with previous studies that were reviewed in the Introduction, the calibrated A_{cr} increases and δ_{Ref} decreases when precipitation decreases (arid and semiarid climates, Figure 1d) for permeable rocks (unconsolidated sediments and carbonated rocks, Figure 1b) and when the slope decreases (consistent with stronger erosive power in steeper watersheds). The calibrated A_{cr} also tends to be larger in Australia (especially for classes with the highest values which permit the largest variability), which can be attributed largely to arid and semiarid climates (65% of Australia), where rivers are rare and often intermittent (the latter amounting to 69% of total stream length in FAO and to 98% in Geofabric, Table S2).
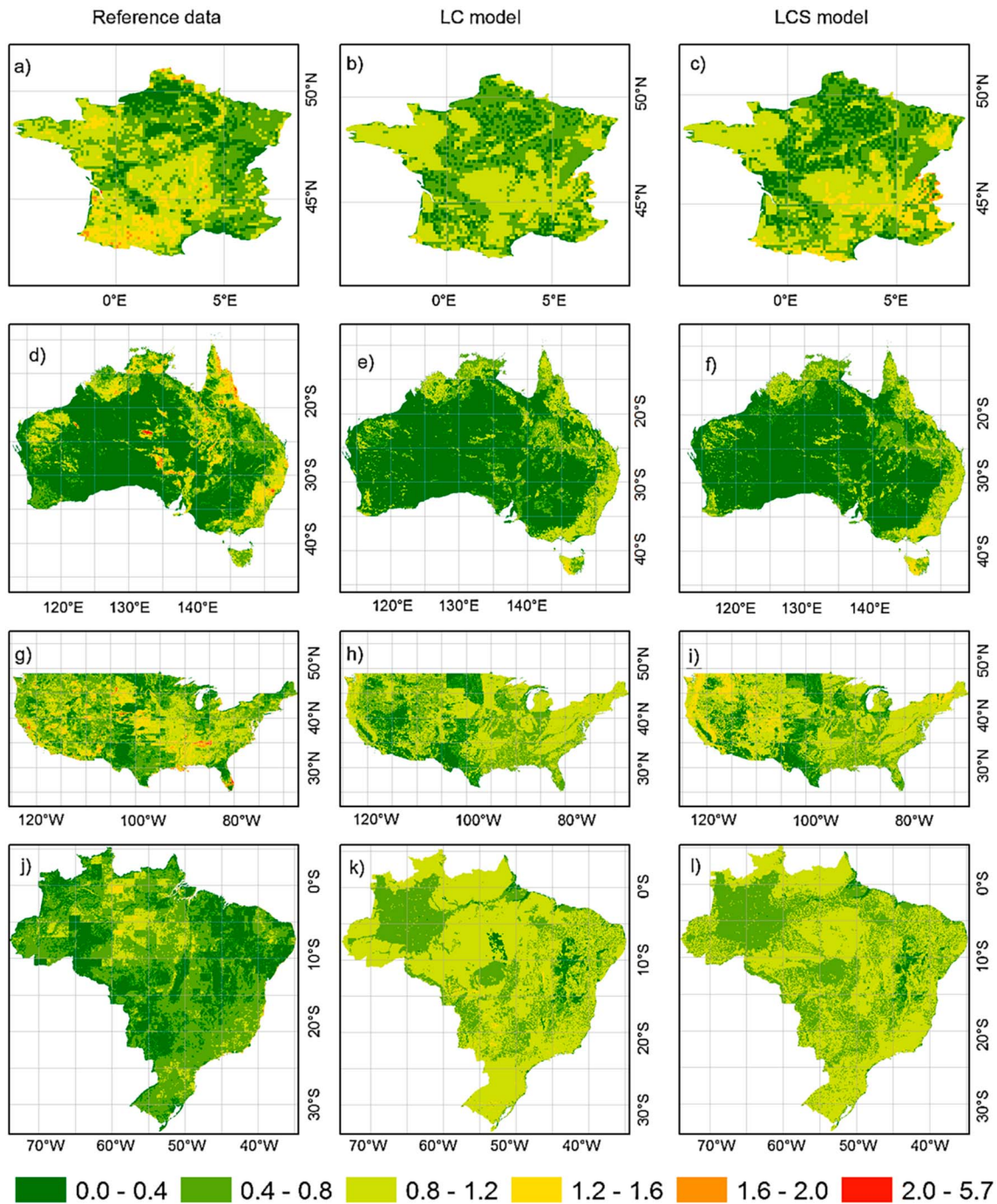
Figure 2. Maps of drainage density (km^{-1}) for the reference river networks and the evaluated models in (a–c) France, (d–f) Australia, (g–i) U.S., and (j–l) Brazil: Reference data (Figures 2a, 2d, 2g, and 2j); LC (Figures 2b, 2e, 2h, and 2k); and LCS (Figures 2c, 2f, 2i, and 2l).

In the second step, we developed statistical models to produce a set of A_{cr} values for the different environmental classes, to be used to extract river networks worldwide. The rationale was to use the rules emerging from Figure 1 regarding the effects of lithology, slope, and climate to create consistency between the two countries and define one single A_{cr} value for a given environmental class. In doing so, we ensured that we accounted for the effects of latitude on pixel area (Text S2), and we tried to group some environmental classes to limit the number of different A_{cr} values.

Table 1. Statistical Parameters for the Comparison of the Calculated and Reference Drainage Densities δ per Country[a]

Region	Parameter	FAO	LC	LCS	Ref.	Others
AU	Mean	0.08	0.42	0.45	0.43	-
	SD	0.07	0.32	0.32	0.43	-
	Bias	−0.36	−0.01	0.02	0.00	-
	Corr.	−0.01	0.57	0.59	1.00	-
	RMSE	0.56	0.36	0.34	0.00	-
FR	Mean	0.08	0.68	0.71	0.75	0.34
	SD	0.08	0.24	0.30	0.36	0.18
	Bias	−0.66	−0.07	−0.04	0.00	−0.41
	Corr.	0.01	0.30	0.32	1.00	0.19
	RMSE	0.81	0.40	0.39	0.00	0.56
US	Mean	0.07	0.72	0.78	0.75	1.30
	SD	0.07	0.28	0.30	0.37	0.84
	Bias	−0.68	−0.03	0.03	0.00	0.55
	Corr.	0.09	0.09	0.11	1.00	0.24
	RMSE	0.77	0.47	0.47	0.00	1.00
BR	Mean	0.08	0.93	0.92	0.52	-
	SD	0.07	0.16	0.16	0.29	-
	Bias	−0.44	0.41	0.40	0.00	-
	Corr.	0.12	0.14	0.14	1.00	-
	RMSE	0.49	0.49	0.48	0.00	-
Europe	Mean	0.08	0.59	0.63	-	0.31
	SD	0.08	0.29	0.32	-	0.15
Global	Mean	0.08	0.69	0.74	-	-
	SD	0.08	0.32	0.34	-	-
	% Intermittent	34	27	29	-	-

[a]Ref. = δ from river network reference data (AU = Geofabric and FR = CARTHAGE used for calibration and US = NHD and BR = IBGE [2015] used for validation); Others = δ from Luo et al. [2016] over the U.S. and from CCM2 [Vogt et al., 2007] over France and Europe; Mean = mean δ (km^{-1}); SD = standard deviation of δ (km^{-1}); Bias = bias between model and Ref. (km^{-1}); Corr. = correlation coefficient between model and Ref. (dimensionless); RMSE = root-mean-square error between model and Ref. (km^{-1}). % Intermittent corresponds to the ratio of total length of intermittent rivers to the total river length.

Two models of different complexity were eventually constructed to evaluate the effects of slope on river network extraction by considering only lithology and climate parameters (LC) and adding the slope (LCS). The corresponding A_{cr} values are given in the supporting information Tables S3 and S4. The LC model comprises 11 different A_{cr} values varying from 0.3 km^2 to 97 km^2 for 35 environmental classes, derived from 7 lithology groups (Mt + Pi + Pa + Py, Pb, Sc, Su, Sm, Ss + Va, and Vb + Vi) and 5 climate classes (Table S1). The LCS model has 29 different A_{cr} values varying from 0.3 to 193 km^2 corresponding to 120 environmental classes that combine 6 lithology classes (Mt + Pi + Pa + Py, Pb, Sc, Su, Ss, and Sm + Va + Vb + Vi), 5 climate classes, and 4 slope classes (under 1%, between 1 and 5%, between 5 and 10%, and over 10%).

We also used the Australian case to propose a first-order method to identify intermittent rivers explained by the aridity of climates. The principle is that rivers initiated in regions where precipitation is smaller than a calibrated threshold are defined as intermittent until they reach a pixel where precipitation exceeds this value. Once a stream becomes perennial, all downstream river pixels remain as such (see the Nile in Figure 3b), so this strategy cannot reproduce cases of disrupted connectivity. As detailed in Text S2, the precipitation threshold was calibrated to get the best overlap with areas of high intermittency in FAO, which defines 69% of intermittent rivers in Australia (% of total length). This defined a threshold of 500 mm yr^{-1}, i.e., the classical upper bound of semiarid climates, according to which 42% of the rivers from the LC or LCS models in Australia are intermittent. This percentage is smaller than in FAO because our models predict higher densities, so the total length of intermittent streams is much higher with our models, approximately 420,000 km in FAO; 1,483,000 km in LCS; and 3,242,000 km in Geofabric. We disregarded the latter data set as it classifies 98% of streams as intermittent, which exceeds the maximum of 90% used in Raymond et al. [2013]. Matching this 98% with our uniform precipitation threshold would also constrain the predicted perennial streams to very humid climates (annual mean precipitation > 1500 mm yr^{-1}).

4. Model Results and Discussion

The resulting drainage densities were first evaluated against the four national reference hydrographic data sets (Figure 2 and Table 1). The LC model captures the main features of the reference δ, in particular, the low values characterizing areas with carbonate rocks in France, and arid to semiarid climates in Australia. However, it underestimates the high δ values, mostly found in mountainous areas in both countries. These are better represented by LCS, which accounts for the increase of δ with slope, but this model still underestimates some very high values, such as in Australia, where siliciclastic rocks are present (Ss in Figure S5), and
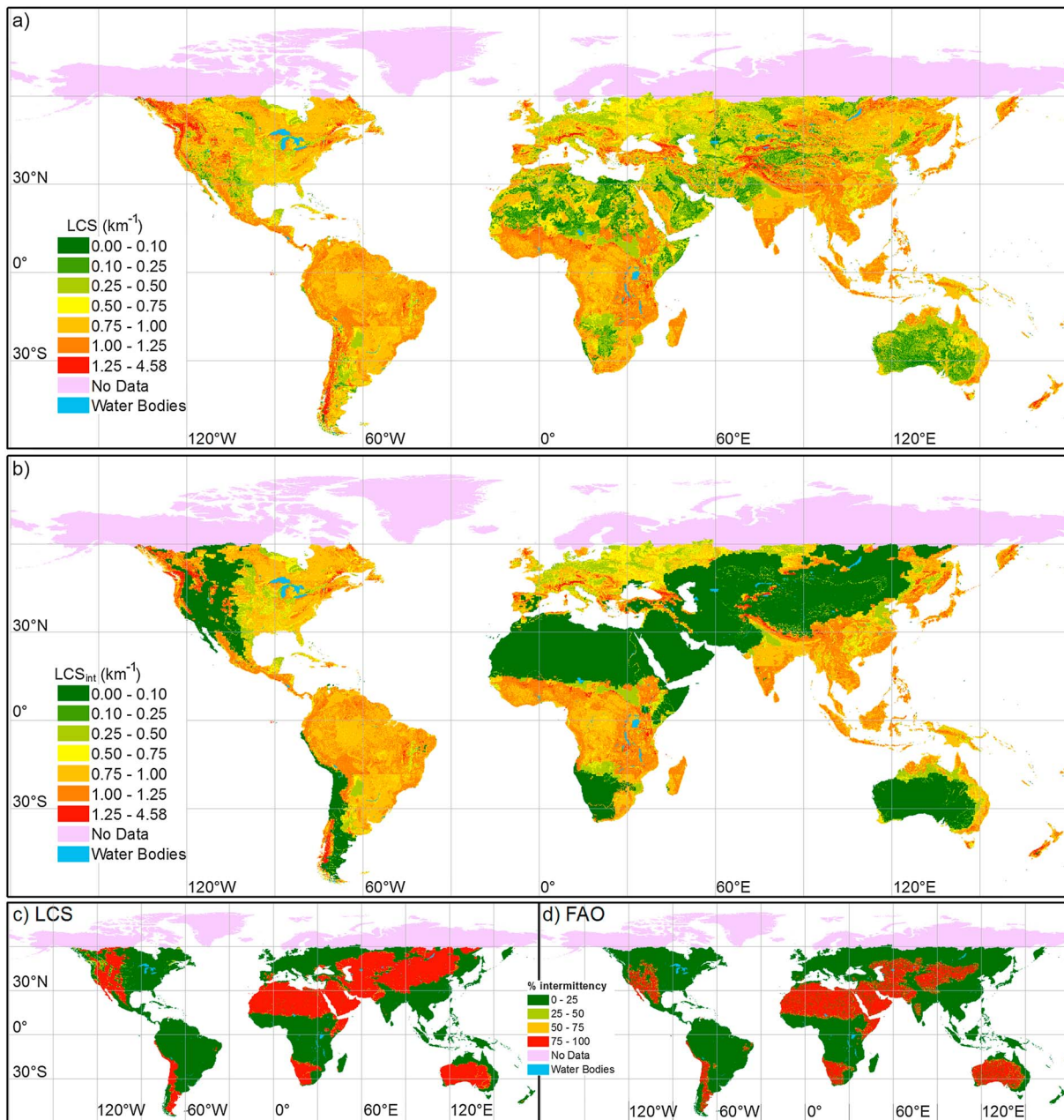
Figure 3. Global maps of drainage density (km^{-1}) for LCS model: (a) for full network including intermittent streams and (b) intermittent streams only. Percentage of intermittent streams: (c) LCS and (d) FAO. In each 7.5′ cell, the % of intermittent streams is the ratio of the intermittent stream length to total stream length. The violet color north of 60°N indicates where HydroSHEDS data are missing.

in the southwestern part of France (Landes), even though it is a rather flat area (Figure S7) with permeable rocks (Su, Sc, and Sm in Figure S5). The latter is likely due to artificial drainage densities in the reference data because of irrigation ditches in this area of intensive agriculture (Text S1).

Similar behaviors are found in the U.S. and Brazil, where the inclusion of a slope constraint in the LCS model also leads to slightly better results than LC, but both models exhibit poor correlations with the reference data sets, as quantified in Table 1. Nevertheless, the excessive underestimation of δ by FAO is markedly improved by the variable A_{cr} models, and the general spatial patterns are well depicted. In the U.S., France, and Australia, the biases approach zero (less than 5% absolute value). They remain negative over France, which is mainly attributed to the overestimation of the highest δ_{Ref} compared to a "pristine" case because of

human-made networks, as discussed previously. In Brazil, in contrast, the proposed method leads to high positive biases; this likely results from the reference river network being based on multiple scaled data, up to 1:250,000, which alters the natural variability of δ (Figure 2) and leads to underestimation of δ_{Ref} compared to what would prevail at the 1:50,000 scale used for the A_{cr} calibration in France and Australia. The inconsistencies of the reference river networks, discussed in Text S1 for both Brazil and the U.S., largely explain the poor spatial fit to δ_{Ref} as revealed by the correlation coefficients and RMSE.

These inconsistencies prevented *Luo et al.* [2016] from analyzing the relationships between drainage density from NHD and possible explanatory factors and led them to construct their own drainage density map. Table 1 shows that this latter map overestimates the mean δ compared to NHD and our two estimates, which suggests that the 30 m DEM of *Luo et al.* [2016] corresponds to a finer scale than both 1:24,000 and 1:50,000. As expected, the correlation coefficients increased (doubled, as Corr$_{(LCS,Luo)}$ = 0.24) when comparing δ of LC and LCS model to δ_{Luo} rather than δ_{Ref} (see also the maps of δ_{Ref}, δ_{Luo} and δ_{LC} in Figures S11 and S13). However, this correlation remains low, and the weak performance of the proposed models in the U.S. could be due to the quality of the lithological map (Figure S6), which explains the main discontinuities along the Great Plains from North Dakota to Texas and around Lake Michigan (anticlockwise from Wisconsin to Michigan). It is worth noting that the geological map of *Schruben et al.* [1994] does not exhibit these discontinuities.

We also compared the proposed δ to values extracted from *Vogt et al.* [2007], which exhibited weak spatial variations throughout Europe, strongly controlled by topography (Figure S12) with values below 0.4 km^{-1} in most lowland areas, which is below the range of 0.5 to 1.6 km^{-1} from the small-scale studies reported in section 1. As a result, δ_{Ref} is much closer to δ from LC and LCS than to δ_{CCM2} over France (Table 1), both in terms of spatial match and bias.

Figure 3a shows the global drainage density map from the LCS model, which exhibits complex patterns arising from the combined controls by lithology, climate, and slope (see also LC and FAO in Figures S13 and S14). The perennial streams show the same density patterns (Figure 3b) but are absent from arid and semiarid areas (39% of global land based on our precipitation data set), except for rivers initiated as perennial which conserve this feature when crossing arid regions, as the Nile, for instance. By design, the complementary intermittent streams are concentrated in arid and semiarid areas both in FAO and LCS (Figures 3c and 3d), with a rather similar proportion of the full network (around one third, Table 1), and a correlation coefficient of 0.65 over land. The main difference is found north of 45°N (North America and eastern Asia), with a larger fraction of intermittent streams according to LCS than FAO. This is consistent with the use of an aridity index combining both precipitation and evapotranspiration to discriminate intermittent streams in FAO/AQUAMAPS. Note that the total length of intermittent streams is much higher with our models than with FAO (Table S2), primarily due to differences in drainage density.

5. Conclusions

This study presents a method to obtain multiple critical drainage areas (A_{cr}) that are spatially dependent on lithology, climate, and slope. This new method addresses an important component of the large-scale river delineation process, in combination with proper DEM hydrologic conditioning, obtained here from HydroSHEDS. The A_{cr} values were calibrated against the national hydrographic data at a 1:50,000 scale over France and Australia, resulting in two models of increasing complexity: LC (using seven lithology and five climate classes), and LCS (using six lithology, five climate, and four slope classes). This work is based on the 15″ hydrologically conditioned version of HydroSHEDS, but the proposed A_{cr} values are a priori suitable to constrain river network extractions from any DEM with similar or higher resolutions.

Both models show fair performance compared to the reference river networks, with better agreement in the countries used for calibration. The inclusion of slope in the model improves the performance criteria in the evaluated countries, but the effect is modest. Combined with an intermittency assessment solely based on mean precipitation, the proposed variable A_{cr} models give similar locations and percentages of intermittent streams as FAO/AQUAMAPS, but with higher and more spatially variable drainage densities. The limitations of this first-order classification underline the need for better description of the multiple controls of intermittency.

The two proposed models predict global mean drainage density to reach approximately $0.70\,\mathrm{km}^{-1}$, with a precision of approximately 5% compared to the reference data. Drainage density and scale are tightly linked [*Tarboton et al.*, 1992], so the proposed value is consistent with the 1:50,000 scale that prevailed for A_{cr} calibration. It is also higher than the mean densities derived from classical single A_{cr} river networks, which are thus shorter and should be used with caution for fine-scale applications, as previously reported by *Raymond et al.* [2013].

As previously discussed, the quality of A_{cr} calibration and validation in our methodology strongly depends on the reference data. The main uncertainties seem to come from unnatural channels and scale inconsistencies. The quality of the input DEM and flow accumulations is also important, although this is difficult to evaluate. As shown, environmental input parameters can induce uncertainties, particularly with respect to the lithological map. As a result, any improved lithology, either at the global scale or over a specific region, could assist with the estimation of drainage density, provided that calibration is updated accordingly. Eventually, another way to improve drainage density estimates would be to include more control factors at the calibration step, e.g., more complex geomorphologic information or vegetation parameters [*Colombo et al.*, 2007; *Luo et al.*, 2016], which were not addressed here for the sake of simplicity. Another approach is to use hyperresolution hydrological modeling to define the locations where streams initiate, as recommended by *Lehner and Grill* [2013] and recently achieved by *Maxwell et al.* [2015] over the United States based on the HydroSHEDS DEM. Land use information could also be used to generate artificialized river networks if the reference data for calibration includes information on anthropogenic pressures. However, when dealing with a surveyed river network with no such information, the drainage density difference with the LCS network may be used as a first-order indicator of anthropogenic impacts.

Taking the above caveats into consideration, river networks and corresponding drainage density maps that can be constructed from the proposed A_{cr} models have the advantage of describing the main heterogeneities of natural river networks with a uniform scale of 1:50,000 across all continents. This feature is important to support water management in regions with limited observations and to provide consistent information to large-scale models seeking higher resolution, which is an important evolution of both land surface and hydrological models [*Wood et al.*, 2011].

References

Benstead, J. P., and D. S. Leigh (2012), An expanded role for river networks, *Nat. Geosci.*, *5*, 678–679.

Billen, G., V. Thieu, J. Garnier, and M. Silvestre (2009), Modelling the N cascade in regional watersheds: The case study of the Seine, Somme and Scheldt rivers, *Agric. Ecosyst. Environ.*, *133*(3), 234–246.

Bishop, K., I. Buffam, M. Erlandsson, J. Fölster, H. Laudon, J. Seibert, and J. Temnerud (2008), Aqua incognita: The unknown headwaters, *Hydrol. Process.*, *22*, 1239–1242.

Brutsaert, W., and J. P. Lopez (1998), Basin-scale geohydrologic drought flow features of riparian aquifers in the southern Great Plains, *Water Resour. Res.*, *34*, 233–240, doi:10.1029/97WR03068.

Brutsaert, W., and J. L. Nieber (1977), Regionalized drought flow hydrographs from a mature glaciated plateau, *Water Resour. Res.*, *13*, 637–643, doi:10.1029/WR013i003p00637.

Burrough, P. A., and R. A. McDonell (1998), *Principles of Geographical Information Systems*, Oxford Univ. Press, New York.

Colombo, R., J. V. Vogt, P. Soille, M. L. Paracchini, and A. de Jager (2007), Deriving river networks and catchments at the European scale from medium resolution digital elevation data, *Catena*, *70*(3), 296–305.

Dingman, S. L. (2015), *Physical Hydrology*, Waveland Press, Long Grove, IL.

Döll, P., and B. Lehner (2002), Validation of a new global 30-min drainage direction map, *J. Hydrol.*, *258*, 214–231.

Food and Agriculture Organization (FAO) (2014), AQUAMAPS Global spatial database on water and agriculture, Food and Agriculture Organization of the United Nations. [Available at http://www.fao.org/nr/water/aquamaps/.]

Fekete, B. M., C. J. Vörösmarty, and R. Lammers (2001), Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error, *Water Resour. Res.*, *37*, 1955–1967, doi:10.1029/2001WR900024.

Graham, S. T., J. S. Famiglietti, and D. R. Maidment (1999), Five-minute, 1/2°, and 1° data sets of continental watersheds and river networks for use in regional and global hydrologic and climate system modeling studies, *Water Resour. Res.*, *35*, 583–587, doi:10.1029/1998WR900068.

Hartmann, J., and N. Moosdorf (2012), The new global lithological map database GLiM: A representation of rock properties at the Earth surface, *Geochem. Geophys. Geosyst.*, *13*, Q12004, doi:10.1029/2012GC004370.

Heine, R. A., C. L. Lant, and R. R. Sengupta (2004), Development and comparison of approaches for automated mapping of stream channel networks, *Ann. Assoc. Am. Geogr.*, *94*(3), 477–490.

Horton, R. E. (1932), Drainage basin characteristics, *Eos Trans. AGU*, *13*, 350–361.

Horton, R. E. (1945), Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology, *Geol. Soc. Am.*, *56*, 275–370.

Hutchinson, M. (1989), A new procedure for gridding elevation and stream line data with automatic removal of spurious pits, *J. Hydrol.*, *106*(3–4), 211–232.

Instituto Brasileiro de Geografia e Estatística (IBGE) (2015), Insituto Brasileiro de Geografia e Estatistica. [Available at ftp://geoftp.ibge.gov.br/.]

IGN (2014), Base de Données sur la CARtographie THématique des AGences de l'eau et du ministère chargé de Institut national de l'information géographique et forestière.

Lehner, B., and P. Döll (2004), Development and validation of a global database of lakes, reservoirs and wetlands, *J. Hydrol.*, *296*(1–4), 1–22.

Lehner, B., and G. Grill (2013), Global river hydrography and network routing: Baseline data and new approaches to study the world\r qs large river systems, *Hydrol. Process.*, *27*(15), 2171–2186.

Lehner, B., K. Verdin, and A. Jarvis (2008), New global hydrography derived from spaceborne elevation data, *Eos Trans. AGU*, *89*(10), 93–94, doi:10.1029/2008EO100001.

Luo, W., J. Jasiewicz, T. Stepinski, J. Wang, C. Xu, and X. Cang (2016), Spatial association between dissection density and environmental factors over the entire conterminous United States, *Geophys. Res. Lett.*, *43*, 692–700, doi:10.1002/2015GL066941.

Maxwell, R., L. Condon, and S. Kollet (2015), A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3, *Geosci. Model Dev.*, *8*(3), 923–937.

Montgomery, D. R., and E. Foufoula-Georgiou (1993), Channel network source representation using digital elevation models, *Water Resour. Res.*, *29*, 3925–3934, doi:10.1029/93WR02463.

Moore, I. D., and G. J. Burch (1986), Physical basis of the length-slope factor in the universal soil loss equation, *Soil Sci. Soc. Am. J.*, *50*(5), 1294–1298.

O'Callaghan, J. F., and D. M. Mark (1984), The extraction of drainage networks from digital elevation data, *Comput. Vis. Graph. Image Process.*, *28*, 323–344.

OSDM (2015), Geofabric Surface Cartography—V.3.0, Office of Spatial Data Management, Commonwealth of Australia (Bureau of Meteorology). [Available at http://www.bom.gov.au/water/geofabric/download.shtml.]

Pelletier, J. D. (2013), A robust, two-parameter method for the extraction of drainage networks from high-resolution digital elevation models (DEMs): Evaluation using synthetic and real-world DEMs, *Water Resour. Res.*, *49*, 75–89, doi:10.1029/2012WR012452.

Persendt, F. C., and C. Gomez (2016), Assessment of drainage network extractions in a low-relief area of the Cuvelai Basin (Namibia) from multiple sources: LiDAR, topographic maps, and digital aerial orthophotographs, *Geomorphology*, *260*, 32–50, doi:10.1016/j.geomorph.2015.06.047.

Raymond, P. A., et al. (2013), Global carbon dioxide emissions from inland waters, *Nature*, *503*, 355–359, doi:10.1038/nature12760.

Sangireddy, H., R. A. Carothers, C. P. Stark, and P. Passalacqua (2016), Controls of climate, topography, vegetation, and lithology on drainage density extracted from high resolution topography data, *J. Hydrol.*, *537*, 271–282, doi:10.1016/j.jhydrol.2016.02.051.

Schruben, P. G. A., R. E. Bawiec, W. J. King, P. B. Beikman, and M. Helen (1994), Geology of the conterminous United States at 1: 2,500,000 scale —A digital representation of the 1974 PB King and HM Beikman map, Peter N. Schweitzer.

Stein, J. L., M. F. Hutchinson, and J. A. Stein (2014), A new stream and nested catchment framework for Australia, *Hydrol. Earth Syst. Sci.*, *18*, 1917–1933, doi:10.5194/hess-18-1917-2014.

Strahler, A. N. (1957), Quantitative analysis of watershed geomorphology, *Eos Trans. AGU*, *38*(6), 913–920.

Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe (1991), On the extraction of channel networks from digital elevation data, *Hydrol. Process.*, *5*, 81–100.

Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe (1992), A physical basis for drainage density, *Geomorphology*, *5*, 59–76.

Tucker, G. E., F. Catani, A. Rinaldo, and R. L. Bras (2001), Statistical analysis of drainage density from digital terrain data, *Geomorphology*, *36*, 187–202.

USGS (2015), National Hydrography Dataset at 1:24000, for all available states, United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS), the United States Geological Survey (USGS), and the Environmental Protection Agency (EPA). [Available at http://nhd.usgs.gov/data.html], accessed in December 2015.

Verdin, K., and S. Greenlee (1998), HYDRO1k documentation, Sioux Falls ND US Geol. Surv. EROS Data Cent. Httpedcdaac Usgs Govgtopo30hydroreadme Html.

Vogt, J., et al. (2007), A pan-European river and catchment database. European Commission, Joint Research Center, Inst. Environ. Sustain. JRC Luxemb., 120 pp.

Vogt, J. V., R. Colombo, and F. Bertolo (2003), Deriving drainage networks and catchment boundaries: A new methodology combining digital elevation data and environmental characteristics, *Geomorphology*, *53*(3), 281–298.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo (2014), The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resour. Res.*, *50*, 7505–7514, doi:10.1002/2014WR015638.

Wood, E. F., et al. (2011), Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resour. Res.*, *47*, W05301, doi:10.1029/2010WR010090.

Wu, H., J. S. Kimball, H. Li, M. Huang, L. R. Leung, and R. F. Adler (2012), A new global river network database for macroscale hydrologic modeling, *Water Resour. Res.*, *48*, W0970, doi:10.1029/2009WR008871.

Zecharias, Y. B., and W. Brutsaert (1988), Recession characteristics of groundwater outflow and base flow from mountainous watersheds, *Water Resour. Res.*, *24*, 1651–1658, doi:10.1029/WR024i010p01651.