Response to George Allen's comments

(Responses *italicized and blue*.)

Overall:

Schwenk et al. presents a method to assign steady-state flow directions to channel links of delta and braided rivers with complex river morphologies. This information is useful for a range of biogeochemical and hydrological flux processes. The manuscript is generally well written and I find the manuscript easy to understand. I think it was a high-quality study and I appreciate that the code and datasets were made freely available. The one major problem I found with this study is that, it might be submitted to the wrong journal. It is a methods study, essentially explaining the RivGraph Python package (https://github.com/jonschwenk/RivGraph), which may not be suitable for ESD, at least according to the aims and scope of the journal. In my opinion this study is probably most appropriate for a journal like IEEE Geoscience and Remote Sensing Letters. However, that being said, if the Editor wishes to continue with the review process, I think Schwenk et al. is a nice contribution.

We thank George for the positive and constructive comments. Maarten Kleinhans (other reviewer) had the same concern: that this paper is too methods-y for Earth Surface Dynamics (ESD). We submitted this work to ESD because we believed that the ESD community is most likely to contain researchers who would find both the method and its implications interesting, although we acknowledge it is an atypical ESD submission. Accordingly, we corresponded with the Associate Editor both before and after reviews to ensure its suitability here and received a supportive response. We have also made additional efforts to appeal to a broader readership by contextualizing this research through comparisons with other work (L50-54, L60-62, L102-105), explaining the theoretical basis for DPAs when applicable (L109-119, L151-154, L170-173, L201-205, L218-223, L246-247, L253-255), and adding discussion about the implications of our results to process-form relationships across CNs (L21, L64-66, L70-71, L413-414, L437-444, L494-497).

We would also like to emphasize that this is not an overview of the RivGraph package. This manuscript represents only one piece of RivGraph, albeit one of the more complicated pieces. We intended this paper to be independent of RivGraph as the recipes described herein are only examples of possible recipes—optimized for the deltas and braided rivers we took as example cases, but hopefully generally applicable.

Major comments:

1. I think the manuscript would benefit from a paragraph in the introduction discussing other channel vectorization algorithms to provide additional context and motivation (e.g. RivaMap, RivWidth, RivWidthCloud, MERIT Hydro, etc.).

We appreciate the suggestion. As mentioned above, this paper is not intended to be a presentation of RivGraph, and therefore we prefer not to discuss in-detail the many tools available for pre-processing data. However, we do agree that it would be beneficial to the reader to be aware of available tools for preparing their channel network, so we have added some text mentioning these. We note, however, that (as far as we know), none of these tools "do it all," including resolving centerlines, network structure (links, nodes, and connectivity), and morphologic properties (width, length). L102-105.

2. Include in the intro and/or abstract that RivGraph determines the steady-state, or mean long-term flow direction. Deltaic systems are often bidirectional flow and this point was only acknowledged in pass in the Conclusions.
We have addressed this by adding a paragraph to the introduction. L36-42.

3. I think a very large potential improvement of this approach would be automatic identification of inlets and outlets and this should either be implemented into the code or acknowledged in the "Improvements and Speed" section.
This is a great suggestion and one we have spent considerable time achieving. For the methods described herein, however, we did not want to confound the thrust of this paper—automatically setting channel directions—with other pre-processing steps one may use to implement these techniques. For this same reason we have neglected to include methods of generating channel masks. However, the RivGraph package does indeed automatically find inlet and outlet nodes for braided rivers. For deltas, the lack of a linear overall flow direction (compared with a valley direction of braided rivers) and the wide range of possible delta configurations renders an automatic solution intractable. However, outlet nodes for deltas are automatically determined by providing a shoreline.

4. It appears that lakes and other non-channelized water bodies are not included in the Delta river masks. Were these removed? These features can be some of the most difficult to skeletonize and I am curious how the authors handled these features.
The methods of mask generation are described in the cited Piliouras and Rowland work. Separate masks were created for the channels and lakes, such that we could isolate the different types of water bodies for various metric calculations and inter-delta comparisons. This was done by classifying water bodies by size, such that the largest connected water body represents the channel network. We then resegmented the channel network by shape to remove lakes that were structurally connected to the channels. The non-channelized water bodies were therefore not 'removed' from the analysis in the present manuscript, but rather the masks presented here represent only the channel class. Future work includes plans to add structurally connected lakes to the network/topology, but as you point out, these features are difficult to skeletonize, and that is beyond the scope of this paper.

5. While the authors may have captured all the major sources of errors for their sample data set, applying these algorithms worldwide will likely cause a number of currently unidentified errors to be identified. I recommend noting this point somewhere in the manuscript main text (e.g. end of section)
We have added text explicitly mentioning that global application might require some recipe modification in the final paragraph of the conclusions. L502-503.

Minor comments:
1. Add how the authors identified inlets and outlets. Was this done manually? Could it be automated?
Please see the response to Major comment 3.

2. Figures:
a. Panels should be in the same order as they are referred to in the main text.
This has been fixed.
b. Figures would benefit from having labeled panels (e.g. "a","b","c", etc..).

Done.

c. Figures are sometimes mislabeled (e.g. Figure 5 is referred to as Figure 6 several times).

Corrected.

d. Figures with maps: Add North arrow(s) to maps that are not oriented North as up.

Added to Figure 1. We did not add arrows to Figure 4, as it is not intended to be a "map" figure, but rather display results.

3. If the authors wish to add an additional end-member sample, the braided section of the Congo River has a very distinct planform geomorphology and could be an additional case to test RivGraph. This idea is just a gentle suggestion, not a demand.

Thank you for the suggestion! Due to the relative similarity of large braided river morphology, we only analyzed two braided river CNs, but we expect very similar results for the Congo River as those achieved for the Brahmaputra and Indus Rivers.

L76: Islands of size 20 pixels or less were removed (filled) from all channel networks. Please justify this action.

We have added a line stating that this island-filling procedure is not strictly necessary, but reduces some of the noise by eliminating smaller channels that are relatively unimportant to the network topology. L90-91.

L87: Replace "GISs" with "GIS software packages"

We ended up removing the reference to GIS in favor of citing specific community-developed tools. L102-105.

L269: "the shortest link is selected as the one to be fixed (flipped), as DPAs are generally more certain about longer links." I probably don't completely understand but why not just flip the link with the lowest direction certainty?

The problem is that while we have certainty estimates for various DPAs, we don't know their degree of certainty relative to each other. Because of this, we can't simply select the highest-certainty prediction. Instead, we resort to link length, as we have empirically observed that the final flow direction prediction is generally more reliable for longer links. Note that this only refers to (a very small number of) cases where there are multiple options for flipping links to fix an interior source/sink.

L283: change "informations" to "information"

Done.

L421: "one second" Is this on one core or is this code paralllelized?

Added some text to clarify this is unparallelized and run on a typical desktop computer. As a side note, the recipes are generally not (easily) parallelizable because they require sequential setting of flow directions. However, the computational expense is minor enough that parallelization would not be needed, even if more complex recipes were created. L471.

Response to Maarten Kleinhans's review

Responses *italicized and blue*.

Note: some comments have been rearranged in order to condense responses.

This manuscript presents algorithms to improve river and delta channel networks derived from water/land-binarised and skeletonised imagery, specifically to assign flow directions to the links between channel nodes. The set of algorithms is tested against expert judgement and found to be accurate. As such the manuscript adds to a growing family of channel network production tools needed for graph-related and other network analysis tools.

The paper is mainly method-oriented and presents no tests or explanations why certain algorithms were added to, or left out of, the workflow. It almost reads as a cooking recipe that tells the reader to add an ingredient without explaining why, and without explaining what would happen without it or with an alternative ingredient. Perhaps this can be resolved by much better figures that explain what the method does and what comes out of it, placing much of the present technical material in a supplement and focusing the paper on the science rather than the method. While this can potentially be repaired, it requires doing such analyses and writing a paper about it, with much of the present manuscript in a supplement, suggesting rather major revisions. In general the paper is very difficult to read as anything other than a recipe for accomplishing something, and what I would expect for this journal is emphasis on that something.

The evaluation can also be part of [refocusing the manuscript], including explanatory reasons why those few links went wrong because that may tell us something very interesting about the method, why it works, and what basic understanding it embodies about fluvial systems. Or perhaps this manuscript is more suitable for another journal, if not a supplement to a paper about the science.

We thank Maarten for his review, and note that his concerns regarding the suitability of this paper for Earth Surface Dynamics (ESD) were shared by the other reviewer. This paper is atypical for ESD in that it is methods-oriented, but per our discussions with the handling editor is acceptable for publication here. However, we appreciate Maarten bringing to our attention the expectations of ESD readers and have attempted to give the paper more relevance among the ESD audience in three ways: First, we have added text to describe the reasoning behind each DPA while citing the relevant work supporting this reasoning (L109-119, L151-154, L170-173, L201-205, L218-223, L246-247, L253-255). Second, we have tried to contextualize this work better by better-defining the existing research landscape (L50-54, L60-62, L102-105). Third, we have tried to highlight the interesting process-form implications that arise from our study—particularly how the effectiveness of DPAs give us clues into the universality (or not) of particular morphodynamic features, and how the variability of the strength process-form relationships renders a morphology/topology-only approach difficult yet achievable (L21, L64-66, L70-71, L413-414, L437-444, L494-497).

There is some data analysis but very little discussion and comparison to work done in the literature. *We have added relevant comparisons as described in the above reply. However, we are unaware of a similar approach/method in the literature to compare directly against. We are aware of directions being set for each link of a CN, but only manually (Marra et. al., Tejedor et al., etc.). Our comparison of the* *recipes against the expert decisions is essentially testing against these previous works, though not on the same CNs.*

For example, section 3.1 is very hard to read and is probably much better understood when graphically presented in schematics.
We have reduced the clutter by eliminating redundant acronyms. We intend for readers to refer to Figure 2 while reading Section 3.1, as it does provide schematic illustrations of the DPAs.

Detailed comments

Text and figures are cluttered by abbreviations, many of which can be resolved. For instance DPA is unnecessary because the entire paper is about that thing so why not name the subalgorithms by the name that says what they do.
It says DPA everywhere so that is clearly redundant.
It is important to denote to readers that each DPA belongs to the class of all DPAs, hence the original notation. However, we have removed the "DPA" label from all text and figures and replaced it with the bold, italicized acronym for each DPA.

Figures are unclear and not so suitable in background choice for journal publication. This can easily be resolved.
Fig1: black background is beautiful for presentation but make more readable white background for the paper. This also applies to some other figures.
Fig4: again nice for presentation but as a figure it does not work. Why not make blue links for downstream and red for upstream with gray in between and white background.
We chose the black background for some figures for both practical and aesthetic reasons. The darkness of a black background provides a wider range of contrast that allows us to more clearly show lots of information (e.g. densely-packed flow directions for each link). We did try many color combinations on white backgrounds but found them all less clear than the black backgrounds. Please see Fig. 1 at the end of this document for some comparisons. We note that all figures in this manuscript comply with Earth Surface Dynamics figures guidelines (https://www.earth-surface-dynamics.net/for_authors/manuscript_preparation.html).

Polish needed: there are multiple grammatical and spelling errors and figure panels need letters for reference in captions.
We have re-edited and spellchecked the manuscript. Letters have been added to figure panels.

Fig2: if in DPA_mdc colours between equation and schematics are supposed to correspond then something is wrong.
This is now fixed.

Fig 3. The caption refers to the text for symbology, but readability would really improve if a figure explains that symbology. It says 'min dir change DPA_mdc' with different omega_ang in multiple places, but why and why these values cannot be understood from the figure.
Fig 3: likewise, this fig is very very hard to read with all the unexplained symbology. Perhaps put in the supplement and make a fig for the paper that explains rather than technically records what the recipes do. For braided rivers the cycles are not connected to the rest, and that bit is the same as in the deltas so for clarity merge the two.

Unfortunately, the definitions of the thresholds (ω) are too involved to include in this figure. However, we have added two sentences to the caption to help explain their purpose and where to find their definitions. We explain in the text and the caption that recipes are combinations of DPAs with the purpose of setting all links' directions. The explanation for what the components (DPAs) do is given in detail in Section 3.1.

Fig6: Nice results, but write out meaning of legend so it becomes readable
The DPA abbreviations have been fully written in the revised caption.

40 missing the most important problem here: bed slopes are nearly as much upward in downstream direction because of shoals and bifurcations, which requires a very different method to get the networks (Kleinhans et al. 2017, Van Dijk et al. 2019).
We were motivated by techniques that are globally-applicable, and high-resolution bathymetry is not widely available for most CNs (especially large ones). We have added text that clarifies our motivation and mentions the difficulties cited in this comment.

46,51 then why is there this remark in the online supplement readme that "Important: The Colville, Kolyma, Lena, Mackenzie, Yenisei, and Yukon channel network masks are not included in these Supplementary Data, as they were painstakingly created by Anastasia Piliouras"? What was so much work about it?
The devil's in the details. While it is now quite easy to create or obtain a binary channel mask, the quality of the mask can vary substantially, and the desired quality depends on the use. The adjective "painstakingly" was included to indicate that attention to detail was paramount in these masks' generation, and that it included a significant amount of fine-scale corrections. We have replaced this adjective with more precise terminology in the Supplementary readme.

136 why this weight? this needs arguments and support. In Marra et al (cited in the paper) we tested and discuss a number of possibilities in view of fluvial morphodynamic functioning.
Marra et al. used width, 1/length, and width/length as possible weights for computing topologic metrics. Here, we found that width alone was sufficient to define "main channels" and did not test other metric. This choice is now further explained in L201-205.

The link to the data https://doi.org/10.15485/1505624 leads to the repository but gives a blank page as result.
We are not sure why a blank page resulted—we are able to download the data from the provided DOI as of 9/24/2019.

References

van Dijk,W.M., Hiatt, M. R., van der Werf, J. J., & Kleinhans, M. G. (2019). Effects of shoal margin collapses on the morphodynamics of a sandy estuary. Journal of Geophysical Research: Earth Surface, 124. https://doi.org/10.1029/2018JF004763

Note: this paper comes from a community where the authors are in alphabetical order. Willem Sonke is the lead author and did this work as part of his PhD thesis.
Kleinhans, M., M. van Kreveld, Tim Ophelders, W. Sonke (lead author), B. Speckmann (PI), and K. Verbeek (2017). Computing Representative Networks for Braided Rivers. 33rd International Symposium

on Computational Geometry (SoCG 2017), pp. 48:1-48:15. Editors: Boris Aronov and Matthew J. Katz. http://dx.doi.org/10.4230/LIPIcs.SoCG.2017.48

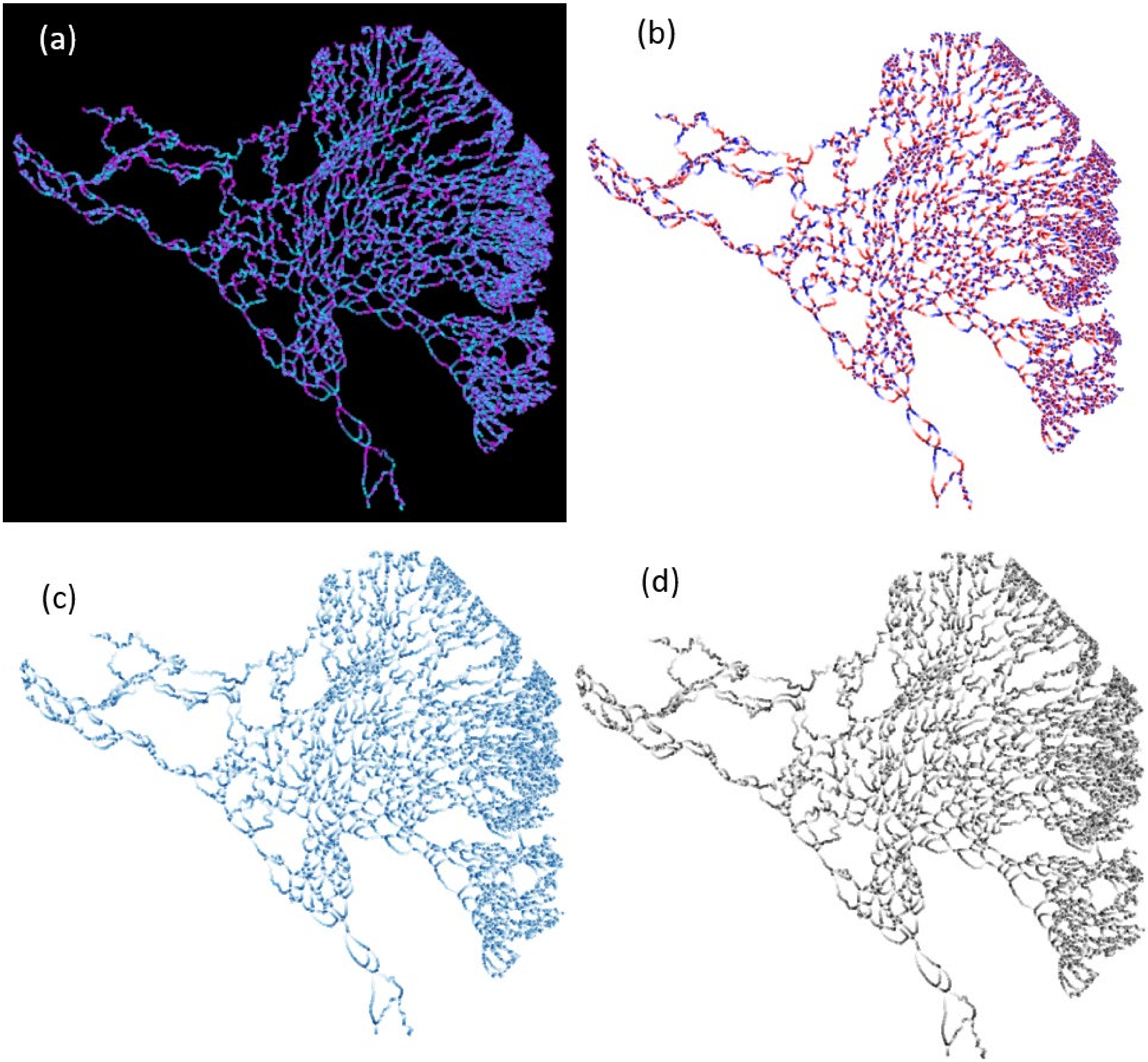

Figure 1. Comparison of color schemes for the Lena Delta.

The red-blue approach suggested (Fig. 1b) is very difficult to interpret. Additionally, the grey between the red-blue transition becomes nearly invisible on a white background. We tested a variety of other colormaps, but they were all difficult to interpret on a white background. To avoid the busy-ness of dense, multiple colors, we also tried single-color colormaps (Fig. 1, c and d). These suffer two problems: (1) again, the brighter portion of the colormap blends into the white background, rendering channels *invisible and (2) human vision tends to lump the darker colors together, making it difficult to readily identify the upstream portions of individual links. In contrast, (a) has both elements of (1) differing contrast (bright to dark) that allow full visibility along each link's length, and (2) different color which makes clear the upstream-downstream differentiation.*

This principle applies to the other figures as well. For example, in Figure 5, we use nine entries in the legend that require visual distinction. The use of contrast differences provided by the black background helps make these differences clear, striking, and obvious. We agree that this technique is likely not necessary for Figure 1, but here we choose it for its aesthetic and are unaware of principles that dictate this choice unsuitable.

Determining flow directions in river channel networks using planform morphology and topology

Jon Schwenk[1*], Anastasia Piliouras[1], Joel C. Rowland[1]

[1]Los Alamos National Laboratory, Earth and Environmental Sciences Division

Correspondence to: Jon Schwenk (jschwenk@lanl.gov)

Abstract. The abundance of global, remotely-sensed surface water observations has accelerated efforts toward characterizing and modeling how water moves across the Earth's surface through complex channel networks. In particular, deltas and braided river channel networks may contain thousands of links that route water, sediment, and nutrients across landscapes. In order to model flows through channel networks and characterize network structure, the direction of flow for each link within the network must be known. In this work, we propose a rapid, automatic, and objective method to identify flow directions for all links of a channel network using only remotely-sensed imagery and knowledge of the network's inlet and outlet locations. We designed a suite of direction-predicting algorithms (DPAs), each of which exploits a particular morphologic characteristic of the channel network to provide a prediction of a link's flow direction. DPAs were chained together to create "recipes", or algorithms that set all the flow directions of a channel network. Separate recipes were built for deltas and braided rivers and applied to seven delta and two braided river channel networks. Across all nine channel networks, the recipes' predicted flow directions agreed with expert judgement for 97% of all tested links, and most disagreements were attributed to unusual channel network topologies that can easily be accounted for by pre-seeding critical links with known flow directions. Our results highlight the (non-)universality of process-form relationships across deltas and braided rivers.

1. Introduction

River channel networks (CNs) sustain communities and ecosystems across the globe by delivering and distributing fluxes of water, sediment, and nutrients. Under a changing climate and widespread anthropogenic influences, modeling the transport of riverine fluxes has become vital for predicting changes in flooding hazards (Hirabayashi et al., 2013; Milly et al., 2002), habitat availability (Erős et al., 2011; Gilvear et al., 2013), contaminant transport, and water resources. CN structure affects both spatial and temporal patterns of riverine fluxes that control changes in habitat availability (Benda et al., 2004; Grant et al., 2007), flooding and floodplain nourishment (Edmonds et al., 2011), and biogeochemical cycling (Czuba et al., 2018; Hiatt et al., 2018). Flow directionality, defined as the direction of flow within each channel of a network, is critically important for vector-based models that route fluxes through CNs and are built atop a graphical representation of the CN (Czuba and Foufoula-Georgiou, 2014, 2015; Lehner and Grill, 2013). Additionally, recent research seeking to characterize deltas and braided rivers based on network structure relies on CN metrics that require knowledge of flow directions for each link (Marra et al., 2014; Tejedor et al., 2015a, 2015b, 2017).

In reality, the flow direction of river discharge may not be steady through time or may be multiple directions simultaneously. Such bi-directional flows may result from large, irreversible perturbations to the channel network (e.g. Shugar et al., 2017), fluid density differences within the channel (e.g. Garcia et al., 2006), or most commonly tidal influence (Fagherazzi et al., 2004). In these cases, local velocity measurements are usually needed to reliably ascertain the flow direction at a given time and location. Although delta CNs often feature some tidally-influenced bidirectional channels, we focus only on the long-time, steady-state flow direction of discharge as it moves from the delta apex to its shoreline.

For watershed-scale (and larger) modeling of river tributary networks, flow directionality can often be ascertained from knowledge of the CN structure and/or a digital elevation model (DEM) (Czuba and Foufoula-Georgiou, 2014; Dottori et al., 2016; Lehner et al., 2008). However, for dense CNs like those of a delta or braided river, a DEM may be unavailable or too coarse to characterize the flow direction of each link within the CN. Even where a DEM is available, the low slopes characterizing most deltas require high vertical precision for reliable estimates of flow directions. Additionally, both deltas and braided river CNs may be dense with short links that require high spatial resolution elevation data to capture the elevation difference across their lengths. Even when a high-resolution DEM is available, the presence of shoals and bifurcations in multi-threaded CNs can result in flows that travel upslope, requiring sophisticated techniques to resolve flow directions (van Dijk et al., 2019; Kleinhans et al., 2017). These challenges render popular DEM-based hydrologic processing algorithms (Schwanghart and Scherler, 2014; Tarboton, 1997) and related products (Lehner et al., 2008; Yamazaki et al., 2019) ineffective. A method for estimating flow directions of links in a CN without auxiliary data would overcome these shortcomings.

With the burgeoning availability of global remotely-sensed surface water products (Allen and Pavelsky, 2018; Pekel et al., 2016; Yamazaki et al., 2015), mapping CN morphologies has become almost trivial. However, the ease of identifying CNs is accompanied by a need for tools that can automatically abstract, model, and analyze CN imagery. Classically, river boundaries, channel networks, and flow directions were simply resolved by hand (Bevis, 2015; Leopold and Wolman, 1957; Marra et al., 2014; Tejedor et al., 2015a), a time-consuming process subject to the operator's judgement. In this work, we present a flexible framework for automatically estimating flow directions in all links of a delta or braided river CN objectively and rapidly that requires only the CN's planform morphology and knowledge of its inlet and outlet locations. While this work focuses primarily on the techniques developed for setting flow directions, analysis of the most effective algorithms also provides clues toward understanding how a dominant flow direction is expressed through a CN's morphologic and topologic CN characteristics.

The remainder of the paper is structured as follows: Section 2 describes the datasets used to create channel network topologies. Section 3 describes the algorithms we designed to set flow directions for all links of a CN. Section 4 assesses the accuracy of our approach, highlights where our method might fail, and discusses how particular characteristics of a river or delta's network relates to uncertainty in directionality. Improvements to reduce errors in setting link directionalities are also discussed.

2. Masks and Networks

We tested our method on a variety of channel networks (CNs) in order to sample a wide range of configurations and scales (Fig. 1). In particular, we selected CNs where network outlets are clustered along disparate regions of the shoreline (Niger, Yukon, Colville), where many channels flow roughly perpendicular to the apparent general flow direction (Lena, Mackenzie, Brahmaputra, Indus), where channel widths span a wide range (Kolyma, Yenisei), and where channels are heavily tidally-influenced (Niger). Only two braided river CNs were selected because braided river CNs exhibit less macro-morphologic variability than delta CNs, and the total number of braided river links we analyzed surpassed that of the deltas. The algorithms presented herein require three independent but related data: 1) binary image of the channel network, 2) vector representation (including connectivity) of the channel network, and 3) locations of inlet and outlet nodes.

The binary image of a CN, or a "mask", is simply a raster wherein "on" pixels belong to the network (Fig. 1, blue). In general, our masks include pixels identified as water or connected-to-water, unvegetated bars. Channel masks for all deltas except the Niger were created from Landsat imagery classified using eCognition software (see Piliouras and Rowland, *in revision*). The Niger CN mask was created from the Global Surface Water monthly-integrated maps, also based on Landsat imagery (Pekel et al., 2016). Both the Brahmaputra and Indus River masks were taken from the Global River Width from Landsat mask of Earth's rivers at mean annual discharge (Allen and Pavelsky, 2018). Islands of size 20 pixels or less were removed (filled) from all channel networks. This infilling, though not strictly necessary, eliminated smaller channels that play relatively unimportant roles in the network structure. Georeferenced .tif files of the Niger, Brahmaputra, and Indus CNs are provided as Supplementary Info; other CNs are downloadable from Piliouras and Rowland, (2019).

The topology of each channel network was resolved from its mask into its constituent links and nodes (nodes shown in Fig. 1) using the Python package RivGraph (Schwenk et al., 2018). Given an input CN mask, RivGraph vectorizes the skeletonized (Zhang and Suen, 1984) mask into links and nodes and stores their connectivities. RivGraph also appends links' morphologic properties including centerline coordinates, channel width at each coordinate, average channel width, and length. RivGraph ensures that all connectivities present in the original masks are preserved in the vector representation. Finally, input and output nodes of each channel network are identified either manually or by RivGraph. Shapefiles of links and nodes and their associated properties for each CN are provided as Supplementary Info. While we used RivGraph to vectorize the network, a number of other tools are available for channel centerline extractions: RivMAP (Schwenk et al., 2017), RivWidth (Pavelsky and Smith, 2008), Rivamap (Isikdogan et al., 2017), and RivWidthCloud (Yang et al., 2019) are among these. However, these approaches would require manual construction of the network's connectivity.

3. Setting Channel Flow Directions

We found no single method sufficient to accurately set all links' flow directions across the variety of tested CNs. We therefore developed a number of sub-algorithms to predict link directionality, deemed here as direction-predicting algorithms (DPAs). Each DPA falls into one of three classes: Exact, Exploitative, or Heuristic. Exact DPAs are those that enforce continuity by ensuring that flow at any point within the CN has a path to an outlet. Sources and sinks are not allowed within CNs, except for pre-identified inlets and outlets. Exploitative DPAs are those that exploit known relationships between particular morphologic or topologic features and dominant flow directions but may not hold for all links within a CN. Finally, a Heuristic DPA is one that assumes an often-intuitive rule but has no strong or formal theoretical basis. Heuristic DPAs were developed through a combination of trail-and-error and qualitative observations of many CNs. With the exception of Exact DPAs, each DPA has an associated uncertainty that is quantified uniquely depending on the particular DPA. Each uncertainty quantity may be thresholded ($\omega_{subscript}$), where *subscript* denotes the relevant uncertainty quantity. These measures of uncertainty are vital for determining flow directions of links, for example where DPAs disagree. The rationales and implementations for each DPA, along with the definitions of uncertainty, are described in Section 3.1.

By chaining together DPAs, "recipes" for setting all flow directions of a CN may be designed. Due to the qualitatively different natures of relatively confined and elongated braided river CNs compared with distributed, multi-directional delta CNs, we developed two separate recipes for fully setting CN directionality of deltas (DR) and braided rivers (BR). Similarly, delta- and braided river-specific DPAs were developed to exploit the qualitative differences between delta and braided river CNs. Section 3.3 describes how DPAs were assembled to create the BR and DR.

3.1. Direction-Predicting Algorithms (DPAs)

3.1.1 Exact DPAs (no uncertainty)

IO: inlets and outlets. Links attached to inlets and outlets are predicted such that flow travels away from inlet nodes and towards outlet nodes.

PAR: parallel links, Fig 2a. Parallel links occur when two links begin and end at the same node. To avoid creating a cycle within the graph, all parallel links must flow the same direction. As a consequence, if the direction of one of a group of parallel links is known, the others are predicted the same direction.

CON: continuity, Fig 2b. Enforcing continuity ensures that no sources or sinks appear within the network other than the inlet and outlet nodes. Continuity is enforced at each node by first identifying nodes where only one connected link's direction is unknown. If the remaining group of known links are all either entering or departing the node, the unknown link is predicted to an orientation opposite the group.

BDG: bridge links, Fig. 2c. Bridge links are those for which all flow must travel through to reach an outlet. Removal of a bridge link from a CN breaks the connectivity of the CN, forming two disconnected CNs. Bridge links are identified in a CN graph via NetworkX's (Hagberg et al., 2008) *bridges* function and temporarily removed, creating two subnetworks. Each subnetwork is searched for the presence of inlet and outlet nodes. If either of the subnetworks has either only inlets or only outlets, the flow direction for the bridge link can be predicted as either away from the subnetwork containing the inlets or toward the subnetwork containing the outlets. In some cases, both subnetworks may contain both inlets and outlets; the bridge link direction is thus not predictable.

3.1.2 Exploitative DPAs

MDC: minimize direction change, Fig 2d. ***MDC*** is based on the principle that branching angles are more likely to be acute, as observed in both inland (Devauchelle et al., 2012; de Serres and DeRoy, 1990) and deltaic (Coffey and Shaw, 2017) CNs. We extend these observations by hypothesizing that the change in flow directions should be minimized at each node. Candidate links for ***MDC*** are identified as unknown links connected to at least one known link. At each end node of a candidate link there may be one or more links flowing into or out of the node. Each of these links, along with the candidate link, is represented by a unit vector whose direction is defined by its endpoint locations (l_u for the unknown candidate link). If multiple links flow into (or out of) the node, their unit vectors are averaged to provide a single direction vector (l_i and l_o for into- and out of-node, respectively). The goal is to determine which of l_i or l_o is most parallel to l_u; thus angles are computed between l_i , l_o and and l_{u0}, l_{u1}, where l_{u0} represents the original position of the unknown link, and l_{u1} represents its 180° rotation about the node. The minimum of all angles is computed via Eq. 1:

$$\alpha_{min} = \min(\alpha_{u_o,l_i}, \alpha_{u_o,l_o}, \alpha_{u_i,l_i}, \alpha_{u_i,l_o}) , \tag{1}$$

where the subscripts denote the vectors defining the angle. If $\alpha_{min}= \alpha_{u_o,l_o}$ or α_{u_i,l_i}, the unknown link is set to flow out of the node, else into it. Where possible, this procedure is repeated for both end nodes of l_u, and α_{min} becomes the minimum of both nodes' minima. The magnitude of α_{min} provides a measure of certainty of the prediction; α_{min} closest to 0 represent links whose flow directions are more aligned with at least one of the known connected links. A threshold (ω_{ang}) may thus be set on α_{min} to specify the maximum level of direction change allowed before setting the unknown link's direction.

SDEM: synthetic DEM (deltas only), Fig. 2e. As discussed in Section 1, DEMs may provide valuable information toward discerning flow directions, but are often too coarse for use with low-sloped delta CNs. ***SDEM*** invokes our conceptualization of long-time, steady-state flow that moves from the apex of a delta to its outlets to construct a synthetic DEM. This procedure creates inlet and outlet DEMs separately, designed such that elevations are higher near inlets and lower near outlets. The final synthetic DEM is simply the sum of the inlet and outlet DEMS.

For the outlet DEM, an image of the same size and resolution of the input mask is created and filled with ones. To estimate the delta's shoreline, the convex hull of the outlet nodes is computed, and the edge of the convex hull connecting the two most distant outlet nodes is removed to provide an ordered set of input nodes. Line segments between each input node are linearly interpolated at 0.1 pixel intervals, and this interpolated shoreline is "burned" into the image of ones by lowering their elevations to zero. A distance transform (Jones et al., 2001) of the image returns an image where each pixel's value represents its distance to the nearest shoreline. This image ($I_{DEM,o}$) is normalized on the interval [0, 1] according to

$$I_{DEM,o} = \frac{I_{DEM,o} - \min(I_{DEM,o})}{\max(I_{DEM,o}) - \min(I_{DEM,o})}. \quad (2)$$

The inlet DEM ($I_{DEM,i}$) is constructed similarly, but with some exceptions. Only inlet nodes whose associated channel widths are at least 75% of the widest inlet channel are considered. Before normalization (Eq. 2), $I_{DEM,i}$ is inverted via

$$I_{DEM,i} = \max(I_{DEM,i}) - I_{DEM,i} \quad (3)$$

to ensure that elevations near the inlets are raised rather than lowered. The final synthetic DEM is simply the sum of $I_{DEM,o}$ and $I_{DEM,i}$. The synthetic DEM for the Mackenzie Delta is shown in Fig. 2e; only one of the inlet nodes contributed to its $I_{DEM,i}$.

The slope of each link may be computed by drawing elevation values from I_{DEM}, and a prediction of a link's flow direction can be made. Channels often flow perpendicular to the general flow direction dictated by I_{DEM}, so predictions made by ***SDEM*** may be poor. However, the magnitudes of a link's slope and its length serve as measures of certainty; links may be thresholded by length (ω_{len}), slope (ω_{slope}), or both to ensure that ***SDEM*** only sets the longest, steepest links.

MC: main channels, Fig 2f. Typically, but with exceptions, the main channels (i.e. those that transport the largest discharge) of a braided river or delta CN are the widest of the CN. This concept originates in well-studied, quasi-universal hydraulic geometry relationships of the form $W=aQ^b$ for width (W), discharge (Q), and fitted parameters (a,b) (Leopold and Maddock, 1953; Parker et al., 2007). We impose two additional constraints to this relationship to define main channels: they must begin at inlets and end at outlets, and they tend to follow the most direct path. Under these conditions, each outlet has a corresponding "shortest and widest" path from each inlet. This path is found by creating a weighted graph of the CN, where weights are defined according to

$$wt_i = l_i \cdot \left(\max(w) - w_i\right) \quad (4)$$

for the i^{th} link with length l_i and width w_i. This weighting scheme results in larger weights for longer and narrower channels. The shortest paths of the weighted graph are computed from each inlet to each outlet using Djiktstra's Method implemented in NetworkX. The direction of each link along each path may then be predicted according to the ordered list of nodes returned. A CN may contain multiple main channels; if a link's direction has already been predicted by a main channel, it is not re-predicted by other main channels that share it. Therefore, in rare cases where two main channels might predict opposite flow directions for a link, the link is predicted by only the flow direction of the first. No uncertainty measurements are made for ***MC***.

VD: valleyline distance (braided rivers only), Fig 2g. Rivers typically flow through corridors of some sort, referred to here as valleys. Valleys feature the lowest elevations in a landscape and contain river floodplains. Multi-scale analyses of river valleys indicate that significant information about the local (link) scale is shared with the valley scale (Gutierrez and Abad, 2014; Vermeulen et al., 2016). ***VD*** attempts to impart flow direction information from the river valley-scale to the link-scale.

A river corridor centerline is created by filling the holes in the CN mask, skeletonizing it, and smoothing. A mesh is generated over the CN by drawing perpendicular line segments along the centerline. This mesh-generation technique was introduced by (Schwenk et al., 2017) and adapted to a Python implementation here. Knowledge of the inlet and outlet nodes' locations allows an ordering of the polygons and perpendiculars comprising the mesh.

A prediction for each link is made by finding the two perpendiculars that encompass the link's endpoints (dotted white lines, Fig. 2g). The link's upstream node is predicted as the one closer to the upstream perpendicular. Similarly to deltas, channels of a braided river may flow approximately perpendicularly to the centerline, resulting in an uncertain prediction. To account for the certainty of ***VD***, the number of perpendiculars required to encompass a link (N_{perps}) is also computed. Links passing through more perpendiculars carry a greater prediction certainty, and a threshold (ω_{perps}) may be applied to predict only the most certain links.

VA: centerline angle (braided rivers only), Fig 2g. The logic behind ***VA*** exactly follows that of ***VD***, but instead of considering downvalley distance, we consider the local angle of the valley centerline. Flow direction can be predicted by comparing a link's angles with the nearby centerline angle. The endpoints of the link are mapped to the nearest perpendicular, and the centerline angle between these two perpendiculars (a_{cl}) is computed (Fig. 2g). The angles of the link computed from the vector defined by its endpoints (a_0) and its 180-degree rotated version (a_1) are also computed. The link's direction is predicted as the orientation whose angle is closest to a_{cl}. The difference between a_{cl} and the closer of a_0 and a_1 provides a measure of certainty of ***VA***, with smaller differences corresponding to higher certainties. This difference may be thresholded (ω_{cla}) to specify the level of parallelism between the link and the valley centerline required to make a prediction.

3.1.3 Heuristic DPAs

SP: shortest path. ***SP*** stemmed from the observation that in most cases where flow direction is unknown, the true flow path corresponded to the shortest distance to its outlet. The ***SP*** implementation is identical to ***MC*** except the links are unweighted. In cases where the shortest path between inlets and outlets results in opposite predictions of flow direction for a link, the mode is selected as the prediction. ***SP*** may fail when macro-morphology of the CN, e.g. the change in Brahmaputra's valley direction from west to south (Fig. 1), imposes a low-frequency direction change.

PMC: parallel to main channel, Fig 2h. Similarly to how ***VA*** and ***VD*** transfer information from the valleyline to predict individual links, the links of main channels contain information of local flow directions that may be exploited to predict nearby, approximately-parallel links whose flow directions are unknown. For each link that is not part of a main channel, the nearest (Euclidean distance) main channel node is found. Each of the endpoint nodes of the unknown link is mapped to their nearest main channel nodes (for example, d_1 and d_2 in Fig. 2h). If the endpoint nodes map to the same main channel node, no prediction can be made for the link. In all other cases, a prediction can be made that aligns the flow direction of the unknown link with the main channel nodes to which its end nodes were mapped. The strength of this prediction (ω_{nodes}) is captured by the difference of mapped-to node positions along the main channel. In Fig. 2h, for example, this number is one (ω_{nodes}=MC_2 - MC_1).

MMA: multiple methods agree. If DPAs disagree about the flow direction of a link, ***MMA*** simply chooses the most common prediction. A minimum number of agreeing DPAs may be enforced (ω_{agree}) to ensure greater certainty of predictions made by ***MMA***.

3.2. Recipes for Deltas and Braided Rivers

DPAs provide a number of tools for predicting flow directions, and they may be assembled into "recipes" designed to set flow directions for all links in a CN. Morphologic variability across our study deltas and braided rivers prevented the design of a "one size fits all" recipe, so we designed both a delta recipe (DR) and a braided river recipe (BR). The arrangement of and thresholds applied to the DPAs used to construct each recipe are detailed in Fig. 3; here, the guiding design principles are discussed.

DPAs provide predictions of some link directions, and each prediction has an associated uncertainty. Only ***IO***, ***CON***, ***PAR***, and ***BDG*** are fully deterministic (i.e. unreliant on thresholding), while all other DPAs provide predictions based on some degree of thresholding. Because some DPAs are only effective when some links' directions are already known (i.e. ***MDC*** and ***MAA***), a recipe must be designed that sets links iteratively, rather than all-at-once.

Setting links iteratively is disadvantageous because an improperly-set link's direction may "infect" nearby links (i.e. cause them to be improperly set), and the infection may spread throughout the network. However, an iterative approach also allows links to be set from most-certain to least, minimizing the likelihood of an infection.

Most DPAs provide a metric of uncertainty in addition to their prediction. By applying thresholds (ω) to these metrics, directions may be set for only a DPA's most certain links rather than applying the DPA to all links. For example, for ***SDEM***, longer and steeper links are more certain, so the first call to ***SDEM*** in the DR is only applied to links that are in the upper-25th and upper-50th percentiles for length and slope, respectively (i.e. ω_{len}=25% and ω_{slope}=50%). Thresholds for DPAs also include ω_{ang} (***MDC***), ω_{nodes} (***PMC***), ω_{agree} (***MAA***), ω_{n_perps} (***VD***), and ω_{cl_ang} (***VA***). For the angle-based thresholds, smaller values correspond to higher certainty, and conversely for the non-angle-based thresholds. The meaning of these thresholds was described in detail in Section 3.1.

The most certain non-Exact DPAs are those containing information of the general flow direction--***SDEM*** for deltas and ***VD*** and ***VA*** for braided rivers. These are applied second, following the Exact DPAs. Continuity (***CON***) is not explicitly shown in the recipes (Fig. 3), but whenever a link's direction is set, all its connected links are attempted to be set by ***CON***. Each time ***MDC*** is applied, the threshold ω_{ang} is applied in equally-spaced intervals of 10 to ensure most certain links are set first. For example, ω_{ang}=1.0 would apply ***MDC*** with ω_{ang}= (0.1, 0.2, … 1.0). It is possible that the BR fails to set all links' directionality; however, we found through visual inspection that flow directions of these unset links were ambiguous, and their flow directions are thus set randomly. Similar links exist in delta CNs, but ***SDEM*** is used to set their directions in the DR. Attempts to fix internal sources/sinks and cycles are made at the ends of both the DR and BR.

3.3. Cycles and Continuity

After all link directions of a CN have been set, the resulting graph may contain interior sources or sinks and/or cycles. A cycle is a set of directed links and nodes for which a node is reachable from itself. While it is possible that a real CN may truly contain a cycle, our conceptualization of a CN as delivering all fluxes from its apex(s) to its outlet(s) precludes their existence in our graphs. Thus cycles identified in a CN indicate a set of links for which at least one link flows in a direction opposite of what is desired; in other words, cycles identify links that should be corrected. Cycles are identified via the NetworkX method *simple_cycles()*. Sources and sinks are identified by ensuring that for all interior nodes (i.e. not inlets or outlets), at least one link departs the node and one link enters it.

If an interior source or sink is present in a CN, a "fix" is attempted. Its goals are to flip the directionality of a single link so that a) the source/sink is no longer present, b) the flipped link does not create another source or sink, and c) the flipped link does not create a cycle. To fix the source/sink, each link connected to the source/sink node is flipped and continuity is re-evaluated. If the link violates continuity post-flip, it is discarded from consideration. For each of the flipped links that did not violate continuity, if flipping its direction creates a cycle, it is also discarded. If more than one links meet these criterion, the shortest link is selected as the one to be fixed (flipped), as DPAs are generally more certain about longer links.

Cycles may be more complicated to fix automatically because there is no upper bound on the number of links they may contain. In practice, cycles typically contained fewer than ~10 links, so an automated cycle fix was implemented. This procedure simply unsets all the directions of links in a cycle, with the exceptions of directions that were set via ***IO***, ***MC***, ***BDG***, ***SDEM***, ***VD***, or ***VA***. The unset links are then reset according only to ***MDC***, beginning with the most-certain angles (lowest ω_{ang}) and longest links. After all links have been reset, a check ensures the cycle has been resolved. If the cycle persists, the same procedure is repeated except the directions of the cycle links plus all links connected to the cycle are initially un-set. If the cycle still remains unfixed, links are returned to their original directions and the cycle is noted for manual inspection.

3.4. Validating Flow Directions

In the absence of data for all links of all CNs that would allow a deterministic evaluation of each link's flow direction, we created a validation database of link directions set according to the judgements of a delta and a braided river expert. For each CN, at least 10% of the total number of links were randomly selected, and their directions were determined manually by the experts using only the same information available to the recipes, i.e. the channel network mask and its graph. Each of the selected link ids were stored along with the experts' best judgement of the corresponding upstream node id. We note that the recipes were developed prior to the development of this validation database. Each disagreement between the expert and the recipe-predicted link direction was investigated, and we also counted the number of expert-errors either due to mistaken data entry or obviously incorrect judgement; expert errors were less than 4% across all individual CNs with an average of 1.7% for all sampled links (Table 1).

4. Results and Discussion

4.1. Overall Accuracy of the Recipes

Overall, we found 97.0% and 98.2% agreement between expert judgement and links set according to the DR and BR, respectively. Henceforth, we consider expert judgement to be "the truth" and refer to disagreements as errors, although the expert judgements were also subject to mistakes (Section 3.4, Table 1). No errors were found within four of the seven delta CNs, with the Niger CN having the highest error rate (9.5%) followed by the Mackenzie (5.3%) and Lena (3.4%) CNs. The BR performed similarly for both braided river CNs, with errors of 2.3% and 2.2% for the Brahmaputra and Indus CNs, respectively. No CNs contained internal sinks or sources, but 4/9 CNs did contain cycles. Of these, only a single cycle was not automatically resolvable for the Lena and Brahmaputra CNs.

4.2. Erroneous links

Each of the 42 identified links that were erroneously set by our recipes was inspected to identify where and how DPAs are likely to fail. Due to the iterative nature of the recipes, erroneous links set early in a recipe are more likely to infect neighboring links, and we found that erroneous links were rarely isolated but occurred in clusters. Because of this, evaluating the accuracy of a particular DPA requires deeper investigation than simply counting the number of erroneous links set by that particular DPA. For example, if ***MC*** erroneously sets a link, ***MDC*** may use the local flow direction of the mis-set link to erroneously set further links.

The following subsections describe sources of errors, including the most common error (Section 4.3.1) and morphologic properties of the Lena (Section 4.3.2) and Niger (Section 4.3.3) CNs that were problematic for our recipes. These subsections explain 29/42 of the identified erroneous links. Of the 13 remaining errors, ***PMC*** was responsible for two, ***MAA*** was responsible for one, and ***MDC*** was responsible for 10. We note that not all erroneous links were identified in the CNs as we only tested >=10% of the total links in each CN. However, because errors tended to occur in clusters and links were randomly sampled for testing it is likely that we captured all the major sources of errors.

4.2.1. Ambiguous links

Sixteen of the 42 link direction errors were attributable to ambiguous links for which morphology alone cannot provide certainty of flow direction. Generally, ambiguous links flow perpendicularly to the local (or overall) flow direction (Figs. 5a-d). Flow directions through ambiguous links can reasonably be argued to go both directions, and in many cases bi-directional flow may be reality (e.g. Fig. 5a). In these cases, ***MDC*** cannot be applied with certainty due to the high junction angles; nor can ***SDEM*** be applied with certainty because ambiguous links are typically short and not parallel to the main flow direction. ***MAA***, which employs shortest-path methods, sets many of these links, but we found shortest path to be unreliable for CNs with large-scale morphologic variability, e.g. the ~90° bend in the Brahmaputra CN. In the case of Fig. 5c, neither ***VA*** nor ***VD*** could set the erroneous links because of their perpendicular orientation with respect to the centerline. Figure 5d shows an unusual ambiguous link created by the formation of an oxbow lake; the expert judgement was based on the flow direction before the oxbow lake was cut off from the main channel, but the modern topology suggests flow could travel in the opposite direction. We were unable to design DPAs that set ambiguous links with certainty; however, ambiguous links were the last ones (i.e. least certain) to be set by our recipes, which limited the influence that their potentially-erroneous flow directions propagated to other links in the network. Although not strictly true, ambiguous links typically play unimportant roles in overall CN routing.

4.2.2. The Niger CN

At 9.5%, the Niger CN contained the highest fraction of erroneous links (Table 1). However, we found that all four erroneous links shared the same source of error. The Niger delta features a number of tidal channels that are typically wider at their outlets and eventually fade away toward the delta's apex. Some of these tidal channels are connected to the CN, while others terminate on the delta plain without a surface connection (Fig. 5e). The erroneous links of the Niger were feeder links from the main CN to a tidal channel that, while connected, likely receives very little flow from main subnetwork. In other words, fluxes to the outlet of this tidal channel should originate at the tidal channel inlets, but these inlets were not considered to be inlets of the CN. Their absence resulted necessarily in a main channel from the CN inlet to the tidal channel outlet, which in turn forced flows right-ward towards the tidal channel and resulted in erroneous links. We verified that placing a single inlet at the source of the tidal channel resolved these erroneous links, resulting in a 0% error for the tested links of the Niger CN.

4.2.3. The Lena CN

The Lena CN had a total of 15 identified erroneous links and an unresolved cycle. Nine of these links and the cycle are attributed to the Lena CN's unusual structure. The Lena CN features two clusters of outlets; a long, continuous shoreline on its upper-right side contains the majority of the outlets, but a separate subnetwork delivers fluxes to the left side (directions with respect to orientation in Figs. 5f-h). Fluxes entering its inlet node are either immediately routed to the left subnetwork or flow upwards to a pseudo-apex (Fig. 5f). While the majority of flow through the pseudo-apex heads toward the right shoreline, some is routed through smaller channels to the left shoreline. Recall that ***MC*** finds the shortest, widest path from inlets to outlets as a main channel. The main channel from the inlet to the outlet denoted in Fig. 5f is incorrect, as flow to that outlet node should travel through the pseudo-apex. However, because of the narrowness of the channels connecting the pseudo-apex to the outlet, the shortest, widest path bypasses the pseudo-apex and flow to the outlet approaches from the wrong side. ***MC*** thus erroneously sets a number of links, including one particularly critical link (Fig. 5g). This link is critical because it bridges two subnetworks; incorrectly setting its flow direction prevents flow from the upper subnetwork from reaching the leftmost outlets. As ***MC*** is applied early in the DR, its incorrect direction more readily infects nearby links as evidenced by the numerous erroneous links surrounding it. Its incorrect direction also created the unresolvable cycle in the Lena; this cycle was not present when we pre-set the critical link to the correct flow direction and then applied the DR.

Four of the Lena CN's 15 identified erroneous links were attributable to the difficulty of creating a representative synthetic DEM. Elevations in the synthetic DEM are proportional to their distance from the outlet nodes and inversely proportional to the distance from the inlet nodes; this scheme created a ridge in the synthetic DEM (Fig. 5h) that divided the inlet subnetwork from the rest of the delta, effectively forcing the links of the inlet subnetwork to flow uphill and resulting in the four erroneous links. Because elevations near the inlet are raised, the slopes of the inlet subnetwork links were relatively smaller, allowing the DR to pass over them (i.e. not set their directions) in the early stages of the recipe. Other DPAs were thus employed to set the vast majority of the inlet subnetwork links, preventing the ridge from adversely affecting their directions. Interestingly, the end of the ridge coincides with the location of the pseudo-apex due to the radial layout of the Lena's outlets.

4.3. Effectiveness of DPAs

Channel networks can exhibit a wide range of morphologic variability, but also contain consistent features that may be exploited by various DPAs. In order to understand which features are more universally consistent, we measured the effectiveness of a DPA by the fraction of a CN's links it sets (Fig. 6). Morphologically, delta and braided river CNs have three key differences exploitable to predict flow directions through their links. First, deltas typically have more outlets than braided rivers. ***IO*** reflects this difference; 5% of delta CN links were set by ***IO*** compared with <1% for braided rivers (Fig. 6). ***SDEM*** also takes advantage of the additional outlets of delta CNs to construct the synthetic DEM, accounting for setting 10% of delta CN links' directions. ***MC*** best exploits the delta CNs' relative numerous outlets as it finds a "main channel" from each inlet to each outlet (25% set, Fig. 6b). However, we found ***MC*** unsuccessful for braided river CNs because link widths were too similar to confidently define a main channel, highlighting the second morphologic difference between delta and braided river CNs: the tendency for delta main channels to be wider relative to the full width distribution, and thus more certainly identified. The average coefficient of variation of link widths reflects this; 1.05 for delta CNs compared with 0.83 for braided river CNs. Finally, the third exploitable morphologic difference between braided rivers and deltas, evident from Fig. 1, is the relatively elongated and confined domain occupied by braided rivers compared with the radial shape of most deltas. The confinement of braided rivers to a relatively narrow band permits a meaningful centerline to be resolved, which we exploited through ***VD***, ***VA***, and ***VAD***. These three DPAs accounted for setting 54% of link directions in the braided river CNs (Fig. 6b).

Across all CNs, more than half of all links were set by ***MDC*** (33%) and ***CON*** (27%) (Fig. 6b). Thus 60% of all links were set with only local flow direction information, highlighting the importance of the accuracy of other DPAs. The basis for developing ***MDC*** lies in theoretical and empirical observations that indicate that channel bifurcation angles tend to deviate an average of ~36° with respect to the upstream channel direction for both river (Devauchelle et al., 2012) and delta (Coffey and Shaw, 2017) CNs. However, the distribution of bifurcation angles may be quite broad and outliers are not uncommon. ***MDC*** was thus applied iteratively and thresholded to be applied only to the links nearest the mean of the bifurcation angle distribution. Despite this iterative approach, ***MDC*** was responsible for the greatest proportion of erroneous links, highlighting the spread around the cited ~36° average. Nevertheless, the similarity of effectiveness of ***MDC*** across both delta and braided river CNs (Fig. 6b) suggests similar local processes at work to form and maintain channel bifurcations in both deltas and braided rivers. Although deltas and braided river CNs are shaped by shared fundamental processes of fluvial erosion and sediment transport, deltas are subject to additional processes including tides, waves, coastal currents, sea level rise, and subsidence that result in more topologic and morphologic complexity. This complexity was reflected in the construction of our recipes, which may be considered as a minimum number of rules required to accurately predict all flow directions in a CN. The DR required 15 DPA applications compared with the BR's seven.

DPAs showed fairly consistent effectiveness across the delta and braided river CNs (Fig. 6a). As expected, ***MC*** and ***IO*** were more effective for delta CNs with many outlets and fewer links (Colville, Kolyma, and Yukon). ***CON*** was relatively more effective and ***MDC*** less effective for the Niger CN than the other deltaic CNs, reflecting its smaller proportion of laterally-flowing (relative to general flow direction) links. The braided river CNs, on the other hand, feature numerous laterally-flowing links and the higher effectiveness of ***CON*** and ***MDC*** relative to the delta CNs reflects this morphologic difference in CNs. Interestingly, almost no bridge links were present in the braided river CNs, while 3% of delta CN links were set by ***BDG***..

4.4. Improvements and Speed

With an overall accuracy of 97.7%, our recipes can provide a suitable starting point for resolving flow directions in all links of delta and braided river CNs. However, some applications (e.g. flux routing) may require complete accuracy. In these cases, perhaps the simplest and most effective method to improve accuracy is to pre-seed the CN with known flow directions. This may be done prior to the initial application of a recipe or in an iterative fashion by identifying critical, erroneous links after the recipe's application. For example, when the correct flow direction was assigned to the critical link of the Lena (Fig. 5g) before applying the recipe, 20 erroneous links and the unresolvable cycle did not occur. Fully specifying all inlet and outlet nodes is also important to improve accuracy, as evidenced by the elimination of all erroneous links from the Niger by adding a single inlet. The flexibility of the recipes allows for easy implementation of other DPAs that can be designed to exploit other morphologic CN properties and improve overall recipe performance.

MDC was responsible for the greatest proportion of erroneous links. Our implementation considered only the link endpoint locations to compute flow direction vectors, but in the cases of longer links, this approach may fail. An alternative and potential improvement might consider only the local link directions (i.e. the pixels of the link closest to the node), although we found this approach challenging to implement. Often the near-node segments of links do not represent a link's actual direction, especially for narrower links connected to wider ones.

In our Python implementation, runtimes for the unparallelized recipes were on the order of one second on a typical desktop machine. However, depending on the size and resolution of the underlying mask, the image processing techniques to create the synthetic DEM for delta CNs and the centerline mesh for braided river CNs can require tens of seconds. These processes must only be run once, though, allowing for rapid development and testing of other DPAs and recipes.

5. Conclusions

This work presents a framework for building algorithmic recipes to automatically and objectively set the steady-state flow directions in all links of a channel network (CN) graph using only a binary mask of the channel network.

Twelve direction predicting algorithms (DPAs) were presented that exploit morphologic and topologic features of a CN to predict the direction of flow within links. By chaining DPAs together, we created recipes for delta CNs and braided river CNs that set all flow directions within the CN.

Knowing only the channel network mask and the locations of inlets and outlets, our recipes for setting link directions agreed with expert opinion for 97% (delta CNs) and 98% (braided river CNs) of links analyzed. Analysis of the links that disagreed showed that special attention must be taken to design recipes for CNs with unusual morphologic features. We also found that CNs may contain critical links that, if set incorrectly, may result in many other mis-set links and cycles in the CN. However, pre-seeding the CN with the correct directions of critical links effectively "cures" such infections.

Even across the wide range of delta morphologies we examined, only a handful of DPAs were required to set the vast majority of links of the CNs. Locally minimizing the change of flow direction between links and enforcing continuity were sufficient to set 60% of links' flow directions in both delta and braided river CNs. Most of the remaining 40% were set by incorporating information from the macroscale CN by identifying main channels, constructing a synthetic DEM (deltas), or leveraging an along-centerline mesh (braided rivers). The effectiveness of ***MDC*** for both deltas and braided rivers points toward the dominant expression of process-form relationships in fluvial systems under a range of environmental conditions. This expression was present but more obscure in delta CNs that are affected by tidal processes.

Although we analyzed large CNs whose masks originated from 30-m resolution Landsat imagery, our recipes are generally applicable to CNs of any scale. The accuracy of our delta recipe across a broad range of delta morphologies suggests a robustness to delta CN forms and suggests that our recipes are applicable to experimental and modeled CNs as well. However, globally, CNs exhibit a wider range of morphologies and topologies than we captured in our test set. If our recipes perform poorly on other CNs, their flexibility and adaptability allow for modification to rearrange the order of DPA application, change the DPA thresholds, or incorporate new DPAs. Relative to deltas, braided rivers exhibit less macro-morphologic variability, so we expect the braided river recipe to be more generally applicable. Our framework is also applicable to other networks and network-based models where directionality is crucial to understand transport, such as in the vascular systems of plants and animals, transportation systems, and utility grids, although application-specific DPAs may need to be developed for these systems.

Code and data availability

The algorithms and recipes detailed here are being implemented into RivGraph (Schwenk et al., 2018), a Python package for analyzing morphologies and topologies of channel networks. An unofficial release of this code can be found at https://github.com/jonschwenk/RivGraph. Georeferenced binary channel masks, distance transforms, link directions geotiffs, and shapefiles of each channel network's directed links and nodes are provided as Supplementary Data.

Author contribution

JS designed and implemented the method presented herein. AP was the delta expert who created the validation dataset. JS did the same for the braided rivers. JS, AP, and JR each contributed to preparing the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

Research presented in this article was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20170668PRD1. This research was also funded as part of the HiLAT project through the Department of Energy, Office of Science, Biological and Environmental Research Program's Regional and Global Model Analysis program. Additionally, this work was partly supported by the DOE Office of Science, BER under the Subsurface and Biogeochemical Research Program Early Career Award to JC Rowland.

Figures

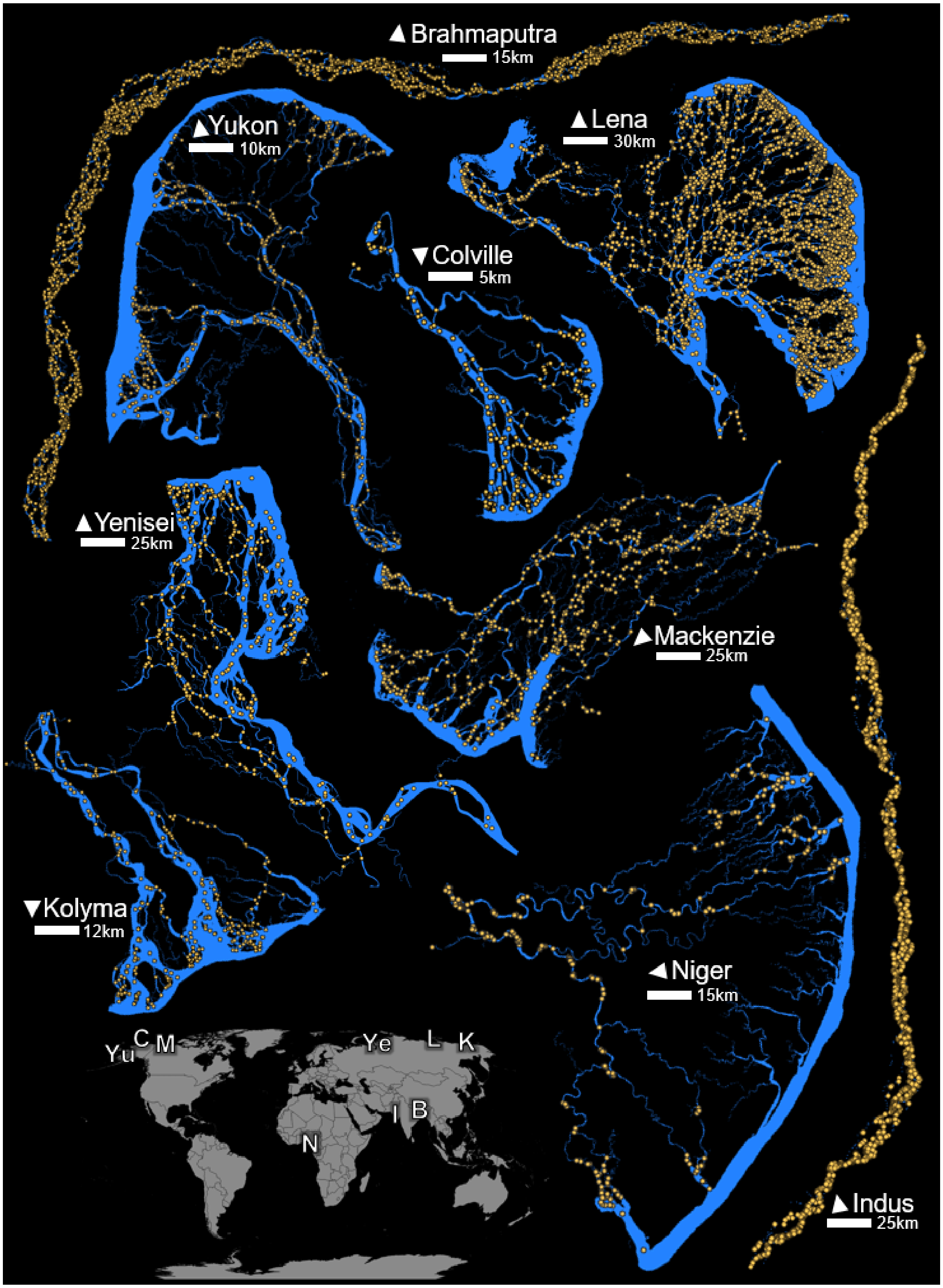

Figure 1. Study channel networks. Channel masks are blue and nodes of the extracted networks are orange. Delta channel masks also include a portion of the ocean along the delta front. For clarity, links are not displayed but may be assumed between each pair of adjacent nodes. Locations of the (Yu)kon, (C)olville, (M)ackenzie, (Ye)nisei, (L)ena, (K)olyma, and (N)iger Deltas and the (B)rahmaputra and (I)ndus Rivers are shown on the map of continents. Arrows for each point north.

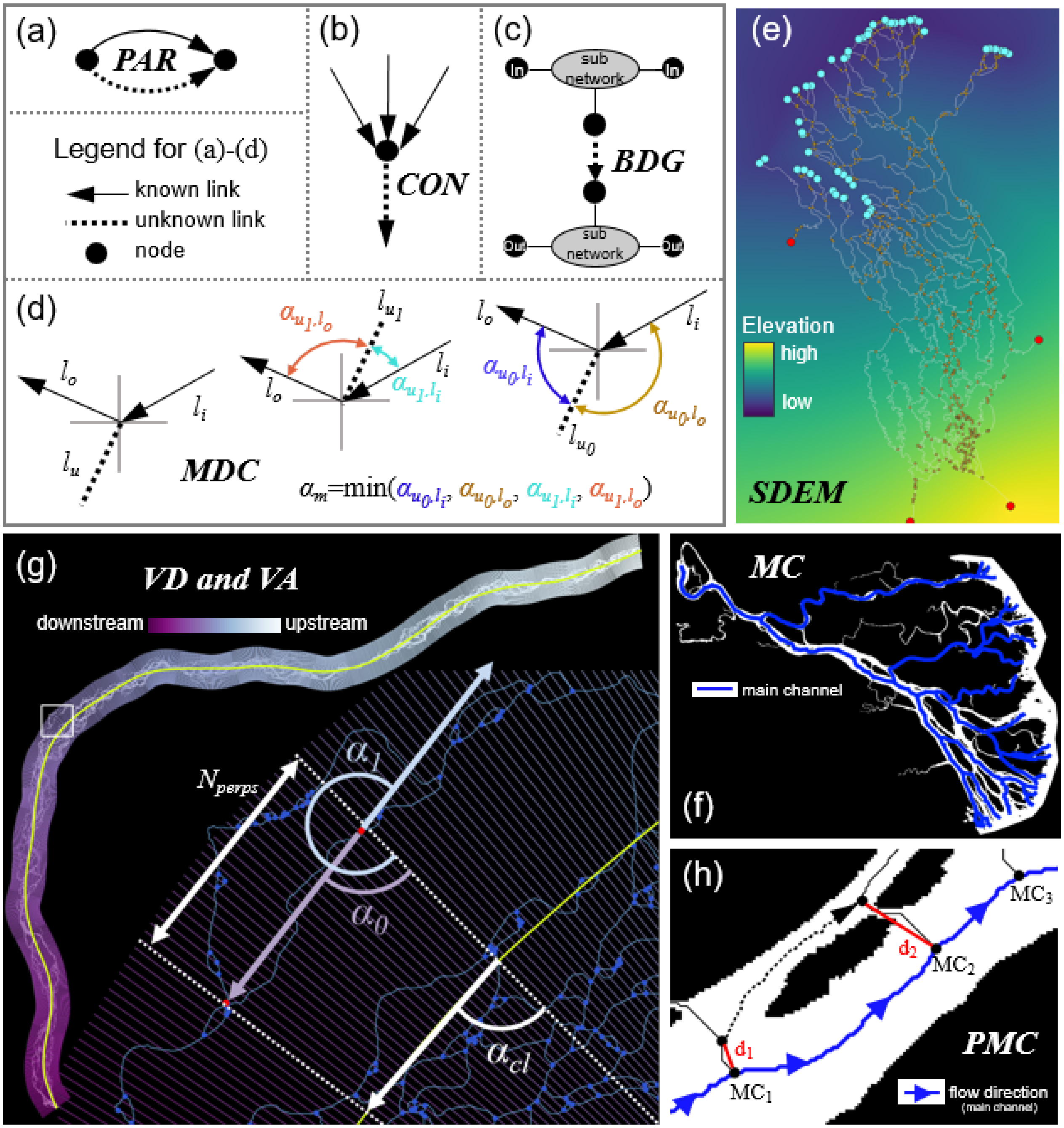

Figure 2. Diagrams showing the direction-predicting algorithms (DPAs). Symbology is further explained in Section 3.1. (a) Predicting an unknown parallel link. (b) An example of applying *CON* to determine the unknown link. (c) Predicting a bridge link with *BDG*. (d) Using the minimum direction change *MDC* for predicting the unknown link. (*e*) A synthetic DEM (*SDEM*) for the Mackenzie Delta is shown with blue outlets and red inlets. (f) A centerline mesh is shown for the Brahamputra River with a yellow centerline to demonstrate *VD* and *VA*. The box denotes the bounds of the zoom-view. The unknown link's endpoints are marked by red points. Dashed lines follow the mesh perpendiculars. (g) Main channels found by *MC* for the Colville Delta are denoted by blue lines. (h) An example of main channel parallels (*PMC*). Distances d_1 and d_2 are defined in Section 3.1; MC_n refers to the n^{th} main channel node, ordered from upstream to downstream. The dashed link's flow direction is unknown.

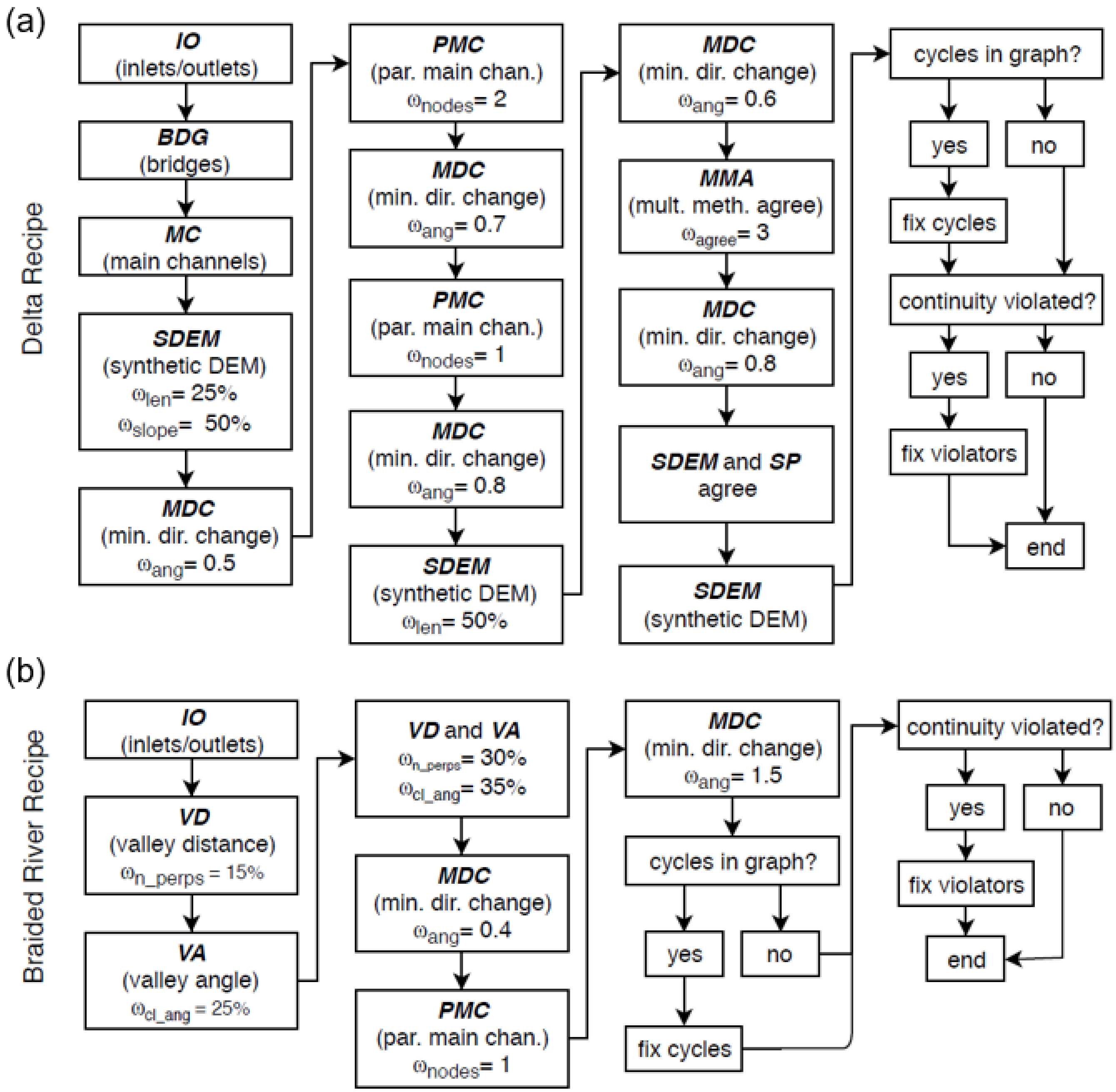

Figure 3. Recipes for setting link flow directions by chaining together DPAs. (a) Delta recipe. (b) Braided river recipe. Continuity (*CON*) is not explicitly represented in the diagrams, but is applied locally after any link's direction is set. Thresholds (ω) are implemented to ensure that only the most certain links are set by each DPA and are defined in Section 3.1. Each threshold has a different meaning that corresponds to the particular DPA.

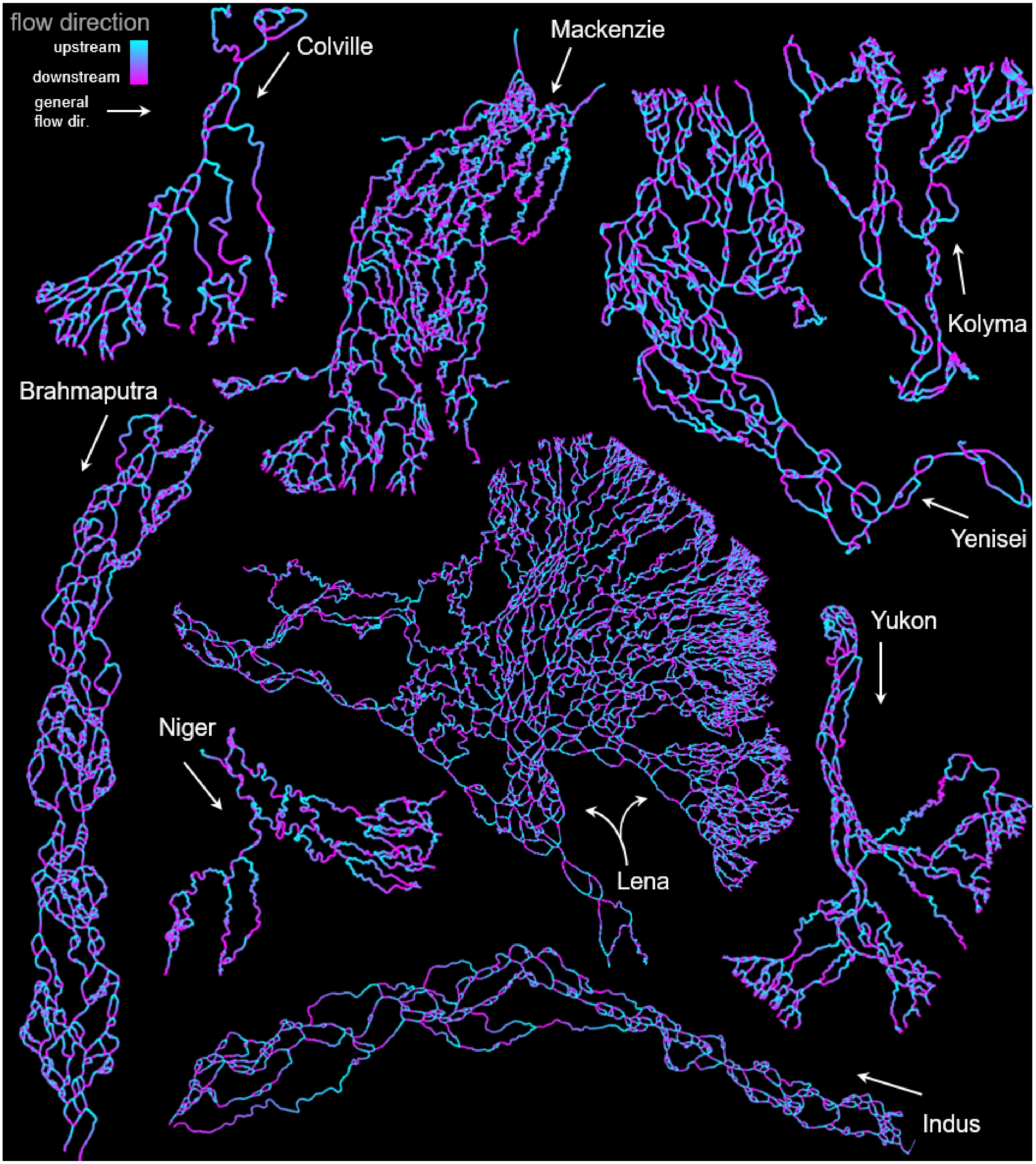

Figure 4. Flow directions for each channel network. The Brahmaputra and Indus Rivers are cropped for improved visibility. White arrows denote the general flow direction for each CN.

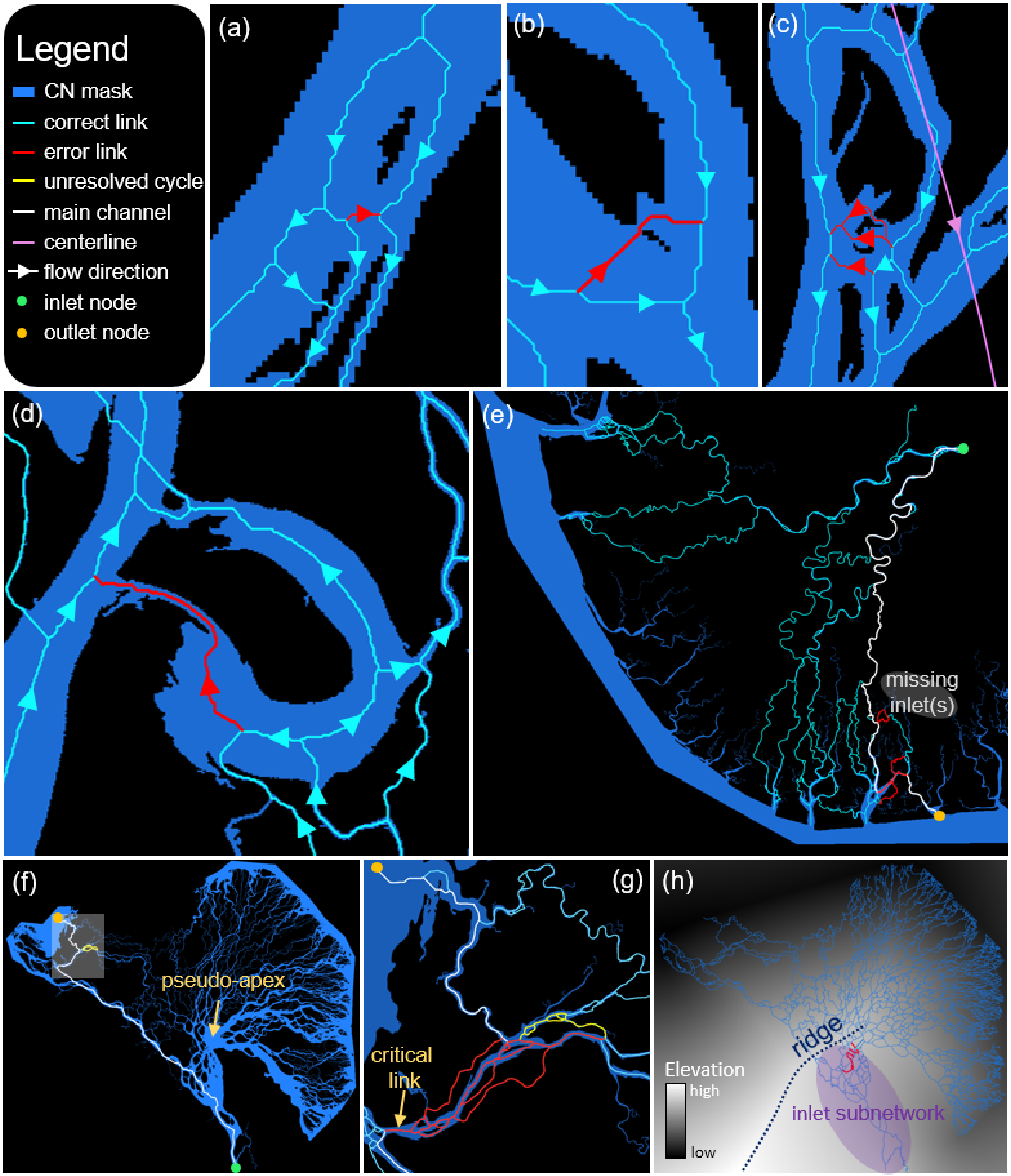

Figure 5. Errors of the recipes. (a-d) Ambiguous links that were erroneously set for the Mackenzie (a and d), Indus (b) and Brahmaputra (c) CNs. (e) The Niger CN features tidal channels whose inlets were not considered. A main channel is shown to an outlet node that should be fed by the missing inlets. (f) A problematic main channel (white) is shown for the Lena CN. A zoom view of the shaded area is shown in (g). (h) Synthetic DEM for the Lena CN with a ridge of the synthetic DEM marked.

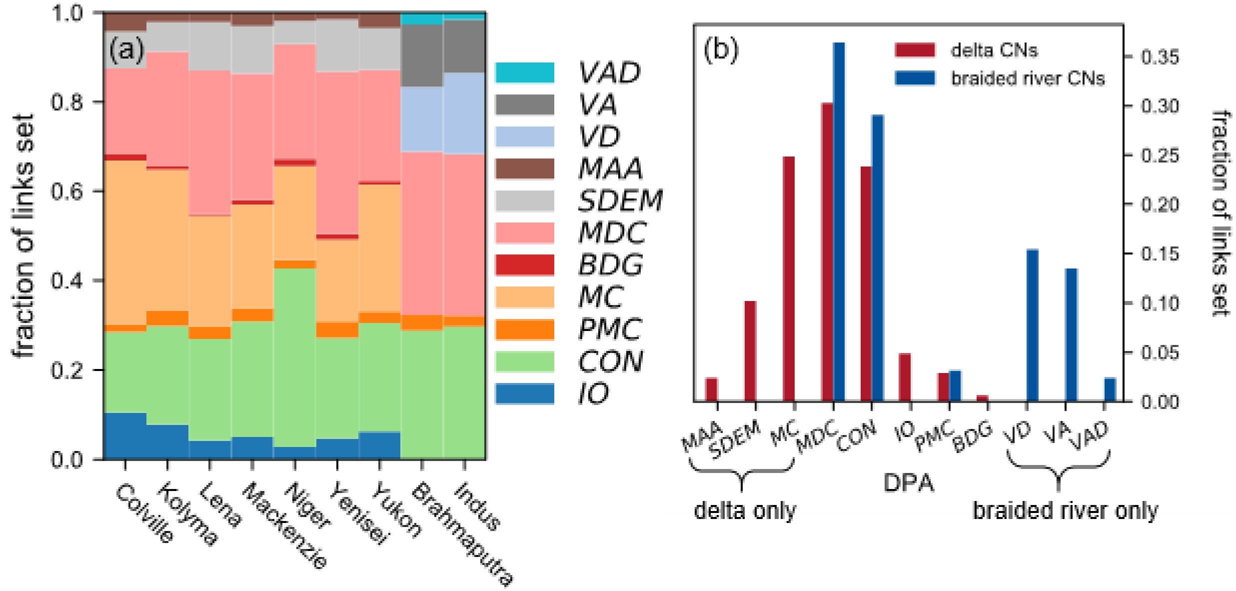

Figure 6. DPA effectiveness. (a) Fraction of links set by each DPA for each study CN. DPAs are defined as VAD-valley angle and distance, VA-valley angle, VD-valley distance, MAA-multiple DPAs agree, SDEM-synthetic DEM, MDC-minimize direction change, BDG-bridge links, MC-main channels, PMC-parallels to main channels, CON-continuity, and IO-inlet/outlet links. (b) Fraction of each DPA for all CNs; red (and blue) bars sum to one.

Tables

CN	links	cycles	cycles fixed	links compared, (%)	disagree, (%)	expert errors, (%)
Colville	256	0	0	30 (11.7)	0 (0.0)	0
Kolyma	421	0	0	49 (11.6)	0 (0.0)	0
Lena	4592	4	3	467 (10.2)	15 (3.2)	2 (0.4)
Mackenzie	1158	1	1	119 (10.3)	6 (5.3)	2 (1.7)
Niger	365	0	0	42 (11.5)	4 (9.5)	1 (2.4)
Yenisei	685	0	0	69 (10.1)	0 (0.0)	0
Yukon	750	1	1	80 (16.6)	0 (0.0)	2 (2.5)
Brahmaputra	6446	5	4	667 (10.3)	11 (1.6)	13 (1.9)

Indus	2103	0	0	308 (14.6)	6 (1.9)	11 (3.6)

Table 1. Channel network properties and errors of the recipes. Links compared % is fraction of total links for each CN. Disagree and expert errors %s are fractions of links compared.

References

Allen, G. H. and Pavelsky, T. M.: Global extent of rivers and streams, Science, 361(6402), 585–588, doi:10.1126/science.aat0636, 2018.

Benda, L., Poff, N. L., Miller, D., Dunne, T., Reeves, G., Pess, G. and Pollock, M.: The Network Dynamics Hypothesis: How Channel Networks Structure Riverine Habitats, BioScience, 54(5), 413–427, doi:10.1641/0006-3568(2004)054[0413:TNDHHC]2.0.CO;2, 2004.

Bevis, M.: Sediment budgets indicate Pleistocene base level fall drives erosion in Minnesota's greater Blue Earth River basin, Master of Science, University of Minnesota, Duluth, Minnesota., 2015.

Coffey, T. S. and Shaw, J. B.: Congruent Bifurcation Angles in River Delta and Tributary Channel Networks, Geophysical Research Letters, 44(22), 11,427-11,436, doi:10.1002/2017GL074873, 2017.

Czuba, J. A. and Foufoula-Georgiou, E.: A network-based framework for identifying potential synchronizations and amplifications of sediment delivery in river basins, Water Resources Research, 50(5), 3826–3851, doi:10.1002/2013WR014227, 2014.

Czuba, J. A. and Foufoula-Georgiou, E.: Dynamic connectivity in a fluvial network for identifying hotspots of geomorphic change, Water Resources Research, 51(3), 1401–1421, doi:10.1002/2014WR016139, 2015.

Czuba, J. A., Hansen, A. T., Foufoula-Georgiou, E. and Finlay, J. C.: Contextualizing Wetlands Within a River Network to Assess Nitrate Removal and Inform Watershed Management, Water Resources Research, 54(2), 1312–1337, doi:10.1002/2017WR021859, 2018.

Devauchelle, O., Petroff, A. P., Seybold, H. F. and Rothman, D. H.: Ramification of stream networks, PNAS, 109(51), 20832–20836, doi:10.1073/pnas.1215218109, 2012.

van Dijk, W. M., Hiatt, M. R., Werf, J. J. van der and Kleinhans, M. G.: Effects of Shoal Margin Collapses on the Morphodynamics of a Sandy Estuary, Journal of Geophysical Research: Earth Surface, 124(1), 195–215, doi:10.1029/2018JF004763, 2019.

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A. and Feyen, L.: Development and evaluation of a framework for global flood hazard mapping, Advances in Water Resources, 94, 87–102, doi:10.1016/j.advwatres.2016.05.002, 2016.

Edmonds, D. A., Paola, C., Hoyal, D. C. J. D. and Sheets, B. A.: Quantitative metrics that describe river deltas and their channel networks, Journal of Geophysical Research: Earth Surface, 116(F4), doi:10.1029/2010JF001955, 2011.

Erős, T., Schmera, D. and Schick, R. S.: Network thinking in riverscape conservation – A graph-based approach, Biological Conservation, 144(1), 184–192, doi:10.1016/j.biocon.2010.08.013, 2011.

Fagherazzi, S., Gabet, E. J. and Furbish, D. J.: The effect of bidirectional flow on tidal channel planforms, Earth Surface Processes and Landforms, 29(3), 295–309, doi:10.1002/esp.1016, 2004.

Garcia, C. M., Manriquez, C., Oberg, K. and Garcia, M. H.: Density currents in the Chicago River, Illinois, in River, Coastal and Estuarine Morphodynamics: RCEM (2005), pp. 191–202, Taylor & Francis., 2006.

Gilvear, D. J., Spray, C. J. and Casas-Mulet, R.: River rehabilitation for the delivery of multiple ecosystem services at the river network scale, Journal of Environmental Management, 126, 30–43, doi:10.1016/j.jenvman.2013.03.026, 2013.

Grant, E. H. C., Lowe, W. H. and Fagan, W. F.: Living in the branches: population dynamics and ecological processes in dendritic networks, Ecology Letters, 10(2), 165–175, doi:10.1111/j.1461-0248.2006.01007.x, 2007.

Gutierrez, R. R. and Abad, J. D.: On the analysis of the medium term planform dynamics of meandering rivers, Water Resources Research, 50(5), 3714–3733, doi:10.1002/2012WR013358, 2014.

Hagberg, A. A., Schult, D. A. and Swart, P. J.: Exploring Network Structure, Dynamics, and Function using NetworkX, in Proceedings of the 7th Python in Science Conference, p. 5., 2008.

Hiatt, M., Castañeda-Moya, E., Twilley, R., Hodges, B. R. and Passalacqua, P.: Channel-Island Connectivity Affects Water Exposure Time Distributions in a Coastal River Delta, Water Resources Research, 54(3), 2212–2232, doi:10.1002/2017WR021289, 2018.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S.: Global flood risk under climate change, Nature Climate Change, 3(9), 816–821, doi:10.1038/nclimate1911, 2013.

Isikdogan, F., Bovik, A. and Passalacqua, P.: RivaMap: An automated river analysis and mapping engine, Remote Sensing of Environment, 202(Big Remotely Sensed Data: tools, applications and experiences), 88–97, doi:10.1016/j.rse.2017.03.044, 2017.

Jones, E., Oliphant, T. and Peterson, P.: Scipy: Open source scientific tools for Python. [online] Available from: http://www.scipy.org/, 2001.

Kleinhans, M., Van Kreveld, M., Ophelders, T., Sonke, W., Speckmann, B. and Verbeek, K.: Computing Representative Networks for Braided Rivers, in 33rd International Symposium on Computational Geometry, pp. 1–16, Leibniz International Proceedings in Informatics., 2017.

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrological Processes, 27(15), 2171–2186, doi:10.1002/hyp.9740, 2013.

Lehner, B., Verdin, K. and Jarvis, A.: New Global Hydrography Derived From Spaceborne Elevation Data, Eos, Transactions American Geophysical Union, 89(10), 93–94, doi:10.1029/2008EO100001, 2008.

Leopold, L. B. and Maddock, T.: The hydraulic geometry of stream channels and some physiographic implications, U.S. Government Printing Office, Washington., 1953.

Leopold, L. B. and Wolman, M. G.: River Channel Patterns: Braided, Meandering, and Straight, U.S. Government Printing Office., 1957.

Marra, W. A., Kleinhans, M. G. and Addink, E. A.: Network concepts to describe channel importance and change in multichannel systems: test results for the Jamuna River, Bangladesh, Earth Surface Processes and Landforms, 39(6), 766–778, doi:10.1002/esp.3482, 2014.

Milly, P. C. D., Wetherald, R. T., Dunne, K. A. and Delworth, T. L.: Increasing risk of great floods in a changing climate, Nature, 415, 4, 2002.

Parker, G., Wilcock, P. R., Paola, C., Dietrich, W. E. and Pitlick, J.: Physical basis for quasi-universal relations describing bankfull hydraulic geometry of single-thread gravel bed rivers, Journal of Geophysical Research: Earth Surface, 112(F4), doi:10.1029/2006JF000549, 2007.

Pavelsky, T. M. and Smith, L. C.: RivWidth: A Software Tool for the Calculation of River Widths from Remotely Sensed Imagery, IEEE Geosci. Remote Sensing Lett., 5(1), 70–73, doi:10.1109/LGRS.2007.908305, 2008.

Pekel, J.-F., Cottam, A., Gorelick, N. and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540(7633), 418–422, doi:10.1038/nature20584, 2016.

Piliouras, A. and Rowland, J. C.: Morphologic variability of Arctic river deltas and the potential implications for fluxes to the coast, Journal of Geophysical Research: Earth Surface, in revision.

Piliouras, A. and Rowland, J. C.: Arctic delta Landsat image classifications, EES-DIVE Dataset, doi:10.15485/1505624, 2019.

Schwanghart, W. and Scherler, D.: Short Communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences, Earth Surf. Dynam., 2(1), 1–7, doi:10.5194/esurf-2-1-2014, 2014.

Schwenk, J., Khandelwal, A., Fratkin, M., Kumar, V. and Foufoula-Georgiou, E.: High spatiotemporal resolution of river planform dynamics from Landsat: The RivMAP toolbox and results from the Ucayali River, Earth and Space Science, 4(2), 46–75, doi:10.1002/2016EA000196, 2017.

Schwenk, J., Tejedor, A., Piliouras, A., Foufoula-Georgiou, E. and Rowland, J. C.: Automatic Extraction of Channel Network Topology (RivGraph), in AGU Fall Meeting Abstracts, Washington, D.C., 2018.

de Serres, B. and DeRoy, A.: Flow direction and branching geometry at junctions in dendritic river networks, The Professional Geographer, 42(2), 1990.

Shugar, D. H., Clague, J. J., Best, J. L., Schoof, C., Willis, M. J., Copland, L. and Roe, G. H.: River piracy and drainage basin reorganization led by climate-driven glacier retreat, Nature Geosci, 10(5), 370–375, doi:10.1038/ngeo2932, 2017.

Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, Water Resources Research, 33(2), 309–319, doi:10.1029/96WR03137, 1997.

Tejedor, A., Longjas, A., Zaliapin, I. and Foufoula-Georgiou, E.: Delta channel networks: 1. A graph-theoretic approach for studying connectivity and steady state transport on deltaic surfaces: Graph-theoretic approach for delta channel networks, Water Resources Research, 51(6), 3998–4018, doi:10.1002/2014WR016577, 2015a.

Tejedor, A., Longjas, A., Zaliapin, I. and Foufoula-Georgiou, E.: Delta channel networks: 2. Metrics of topologic and dynamic complexity for delta comparison, physical inference, and vulnerability assessment, Water Resources Research, 51(6), 4019–4045, doi:10.1002/2014WR016604, 2015b.

Tejedor, A., Longjas, A., Edmonds, D. A., Zaliapin, I., Georgiou, T. T., Rinaldo, A. and Foufoula-Georgiou, E.: Entropy and optimality in river deltas, PNAS, 114(44), 11651–11656, doi:10.1073/pnas.1708404114, 2017.

Vermeulen, B., Hoitink, A. J. F., Zolezzi, G., Abad, J. D. and Aalto, R.: Multiscale structure of meanders, Geophysical Research Letters, 43(7), 3288–3297, doi:10.1002/2016GL068238, 2016.

Yamazaki, D., Trigg, M. A. and Ikeshima, D.: Development of a global ~90m water body map using multi-temporal Landsat images, Remote Sensing of Environment, 171, 337–351, doi:10.1016/j.rse.2015.10.014, 2015.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H. and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resources Research, 55(6), 5053–5073, doi:10.1029/2019WR024873, 2019.

Yang, X., Pavelsky, T. M., Allen, G. H. and Donchyts, G.: RivWidthCloud: An Automated Google Earth Engine Algorithm for River Width Extraction From Remotely Sensed Imagery, IEEE Geosci. Remote Sensing Lett., 1–5, doi:10.1109/LGRS.2019.2920225, 2019.

Zhang, T. Y. and Suen, C. Y.: A fast parallel algorithm for thinning digital patterns, Communications of the ACM, 27(3), 236–239, doi:10.1145/357994.358023, 1984.