AN AUTOMATED BOTTOM UP HYDROLOGIC CODING SYSTEM FOR DENDRITIC RIVER SYSTEM

Zhaoxin Dai, *, Chengming Li, Pengda Wu, Yong Yin

Chinese Academy of Surveying and Mapping, Beijing, China. - daizx@lreis.ac.cn, cmli@casm.ac.cn, wupd@casm.ac.cn

KEY WORDS: Dendritic River System, Stream Ordering, Stroke Connections, Tracing from Estuary, Automated, Hydrologic Coding

ABSTRACT:

The hydrologic coding of a dendritic river system can be used to indicate the stream order and spatial-structure features of a river system. However, most conventional hydrologic coding systems use a top down (i.e., from the source to the estuary) approach, which does not accurately reflect the hierarchies of a dendritic river system. Based on Gravelius' definition of stream order, we have proposed an automated bottom up (i.e., traced from the estuary to the source) coding system for dendritic river systems. Firstly, directed topology trees are constructed to organize the data of a river network, followed by the calculation of stroke connections to construct river entities and identify the main stem. Secondly, the river system is then progressively coded from the estuary up to the source according to the definition of stream order in our system. Finally, actual topographic map data of a river system in some regions of Hubei Province are used to comparatively validate our hydrologic coding system against Horton's coding system. The experimental results demonstrate that our coding system is very effective in highlighting hierarchies formed by catchment relationships. Our method is also found to be suitable for the selection of streams in topologically shallow dendritic river systems with uniformly distributed tributaries.

1. INTRODUCTION

Because dendritic river systems are characterized by distinct hierarchical and fractal features (He, 2004), hydrological stream coding is an effective approach for revealing the spatial-structure features of these river systems and the order of each stream. Therefore, the hydrologic coding of dendritic river systems has always been a focal point for map generalization studies (Zhang et al., 2007; Gülgen, 2017).

Horton (1945) proposed the classic Horton coding system based on the natural patterns in which water flows develop under the effects of gravity. In this coding system, branchless tributaries at the end of dendritic river systems are defined as first-order streams, and streams in the next level of the hierarchy are defined as second-order streams, the trunk of a river system is, thus, defined as the highest order stream (i.e., the nth-order stream). The Horton coding system distinguishes the order of each stream and provides an intuitive reflection of the subtree depth of the stream, which makes it a powerful tool in quantitative analyses on the structural features of river networks. Strahler et al. (1957) and Scheidegger et al. (1961) were the first to use Horton's stream ordering system for hydrographic analysis, while Moharir et al. (2014) and Harish et al. (2016) combined Horton's stream ordering system with remote sensing images to analyze the topographic features of watersheds. In addition, since Horton's coding system can be used to distinguish stream order, Sen et al. (2015) used this coding system for stream selection in the generalization of river systems.

Strahler (1957) further developed Horton's theories to establish the Strahler coding system, which is based on river reaches. This coding system is similar to Horton's system, terminal and branchless reaches of a river system are defined as first-order streams, while reaches that are formed by the convergence of two or more first-order reaches are denoted as second-order streams. This process is repeated until all of the reaches of a river system have been ordered. Strahler's coding system reflects on the size and morphological features of substreams and is useful for highlighting the dendritic features of a river network. Khatun et al. (2018) classified streams using the Strahler system and utilized this coding system in combination with remote sensing images to analyze the river networks of a certain region. Stanislawski et al. (2011) used the Strahler coding system as a criterion to identify the importance of reach during stream generalization and selection. Since reach-based coding system is a better fit for organizing river networks with map data, Shreve (1975), Scheidegger (1965), Horsfield and Cumming (1968) proposed their own hydrologic coding systems based on Strahler's coding system. These hydrologic coding systems reflect upon features such as the number of tributaries and differences in the density of a river system.

The aforementioned coding systems are widely applied in hydrographic analyses and river generalizations. However, these systems generally use a top down approach, where streams are ordered from the end of a river system (i.e., river source) towards the main stem (i.e., estuary). When these coding systems are used to process large dendritic river systems, the resulting stream codes are not very useful for highlighting the hierarchies of a river system, which are formed by catchment relationships. Gravelius (1914) proposed a classic definition for the stream order based on catchment relationships. Based on this definition, an automated bottom up hydrologic coding system was proposed for dendritic river systems. In the method, river entities are treated as elements, and the stream hierarchy is accounted for in the coding of the stream order.

---

* Zhaoxin Dai

2. RELATED WORKS

2.1 Current stream coding methods

(1) Horton's stream coding system
The objective of Horton's coding system is to categorize river entities that have similar shapes and structural features in different watersheds into the same order. To this end, branchless terminal tributaries in a dendritic river system are assigned an order of 1 in the Horton stream system. Based on the directions of each stream, streams that only contain first-order streams are defined as second-order streams, while streams that contain first- and second-order streams are defined as third-order streams. In this way, the main stem of a river system, which has the greatest number of tributaries, is defined as the highest order stream (nth-order), as shown in Fig. 1(a).

(2) Strahler's stream coding system
In map data, rivers are usually stored as a collection of reach elements. Strahler stream ordering may be performed on a dendritic river system using the procedure described below, once the directions of the streams have been determined:

Step 1: A tree structure that contains the topological information of the nodes and arcs of a river system is constructed.

Step 2: Based on the directions of the river system, the indegree and outdegree of each node are calculated. Nodes that have an outdegree of 1 and indegree of 0 are identified as source nodes, while nodes that have an indegree of 1 and outdegree of 0 are identified as estuary nodes.

Step 3: Arcs that are connected to source nodes are defined as first-order reaches, while reaches formed by the convergence of two or more first-order reaches have a stream order of 2. In this way, any reach formed by two or more nth-order reaches has a stream order of $n + 1$.

Step 4: All of the reaches are traversed to repeat Step 3 until all of the streams have been coded. The Strahler codes of a river system are then obtained, as shown in Fig. 1(b).
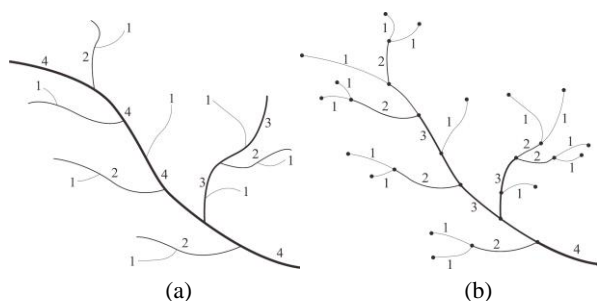

Figure 1. Two classic stream coding systems for dendritic river systems: (a) Horton coding and (b) Strahler coding

2.2 Inadequacies of the current methods

Fig. 1(a) shows that Horton's coding system (which is based on river entities) is an excellent method for distinguishing the stream order and describing the subtree depth of a stream. However, since Horton's coding system uses a top down approach, tributary entities that have the same Horton order in large river systems may not actually have the same stream order. In Fig. 2, the AB (green), CD (yellow) and EF (red) streams all

have a Horton code of 1 despite having different stream orders in reality. The stream AB is connected to the main stem and is, therefore, a first-order stream, stream CD is connected to a first-order tributary and is, therefore, a second-order stream, stream EF is connected to a second-order tributary and is, therefore, a third-order stream.

Therefore, river entities that have the same Horton stream order cannot be compared to each other, and the orders assigned by this system do not accurately reflect the stream hierarchies formed by the catchment relationships. Similarly, Strahler's coding system does not produce an accurate reflection of the hierarchies between river entities. Fig. 1(b) shows that the reach-based Strahler coding system mainly focuses on river depth and shape features rather than the contextual relationships of the streams. Thus, it is impossible to highlight the connectivity and hierarchical nature of river entities using this stream coding system.
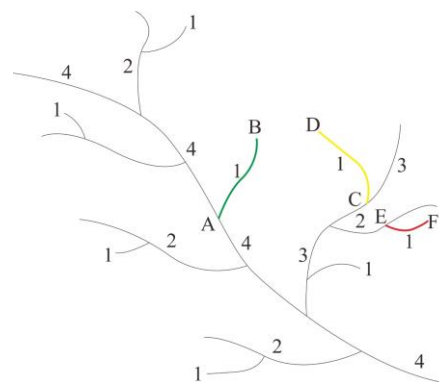

Figure 2. Inadequacies of current stream coding methods

3. BOTTOM UP CODING SYSTEM FOR DENDRITIC RIVER SYSTEMS

Gravelius (1914) proposed a classic definition for stream order, in which the main stem is defined as a first-order stream, while the tributaries connected to the main stem are second-order streams. In this way, branchless tributaries at the end of a river system are defined as the highest order streams. This definition of stream order clearly defines the hierarchies formed by catchment relationships in a river system. On this basis, a bottom up coding system is proposed for dendritic river systems that consists of three parts: (1) the construction of directed topology trees (DTTs) for a river system, (2) the construction of strokes to facilitate the connection of various reaches in a river system and the identification of main stems, and (3) the coding of a dendritic river system by tracing streams from estuary.

3.1 The construction of directed topology trees for river systems

Topological structure diagrams for dendritic river systems that contain flow directions are also known as DTTs (Wu et al., 2013), which are collections of nodes and arcs that record information such as the degree, out-degree and in-degree of each node (Zhang et al., 2004). The direction of each arc (edge) is defined as the direction from the source node to the terminal node. In addition, DTTs also contain semantic (e.g., name and type) and geometric (e.g., length and width) stream information.

3.2 The construction of stroke connections and main stem identification

The recognition of main stems is an important step in the coding of river systems. Some scholars have proposed that river reaches could be connected to form river entities either by length maximization, straight-line approximation, or a combination of both. However, these methods do not account for the semantics or directions of a stream. Thomson et al. (Thomson & Brooks, 2000; Thomson & Brooks, 2001) proposed that the principle of connectivity in Gestalt psychology could be used to construct strokes for river systems based on the semantics, geometries, topologies and directions of their streams. Excellent results have been obtained using this approach. In our method, the stroke connections of a dendritic river system are iteratively calculated using DTTs based on the principles of semantic consistency, directional consistency and length prioritization (i.e., longer is better). The main stem is then identified on this basis. The procedures for the calculation of stroke connections and main stem identification are described below, with Fig. 3 being used as an illustrative example:

Step 1: In most cases, the downstream reaches of a dendritic river system only have one estuary. Therefore, we have chosen the estuary as the origin (Point O) from which the stroke connections are traced. The arc associated with the estuary is the tracing arc (Arc OP), and the other node of this arc (Point P) is the tracing node.

Step 2: The arcs associated with P are incorporated into the set of candidate arcs for stroke connection R {PS, PT}. The angles of these arcs with respect to OP (i.e., { ∠ OPS, ∠ OPT}, respectively) are also calculated.

Step 3: The arc in set R that forms a stroke with OP is selected according to the principles of semantic consistency, directional consistency and length prioritization. In this case, S was selected to form OS, and S is now the tracing node.

Step 4: Stroke connections are continuously calculated and traced according to the methods of Steps 2 and 3 until the stream has been traced up to its source. The resulting stroke is then the main stem of the estuary.

Step 5: All of the stroke connections to the arc associated with the river source are then calculated to obtain the tributaries of the main stem.

Step 6: All of the stroke connections associated with the aforementioned tributaries are then calculated until all of the arcs in the DTT have been calculated. The calculation of stroke connections terminates at this point.
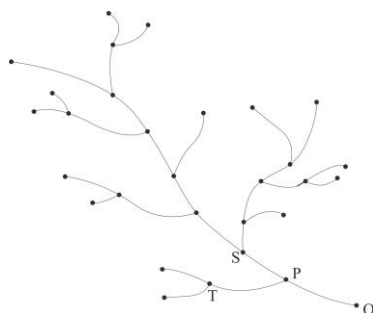

Figure 3. The construction of stroke connections

3.3 Bottom up coding method for dendritic river systems

In this work, streams are coded in a bottom up fashion (i.e., traced from the estuary) based on the definition of stream order given by Gravelius (1914). The main stem of a estuary is defined as a first-order stream, while the streams connected to the main stem are defined as tributaries and are, thus, second-order streams. All secondary tributaries associated with an nth-order stream, therefore, have an order of n + 1. The stream coding process is repeated until all of the streams associated with the source have been coded.

3.4 Procedures of automated coding system for dendritic river systems

The procedures of automated coding system for dendritic river systems are described below, with the river system shown in Fig. 4 being used as an illustrative example:

Step 1: Stroke connections are calculated from the estuary, O, until the main stem of the river system has been calculated. The main stem is assigned a code (order) of 1, as shown in Fig. 4(a).

Step 2: All of the stroke connections associated with the arc of the main stem are calculated to obtain the tributaries of the main stem. These tributaries are assigned a code of 2, as shown in Fig. 4(b).

Step 3: All of the stroke connections to the tributary arcs are calculated and coded. The final results of the stream ordering process are shown in Fig. 4(c).
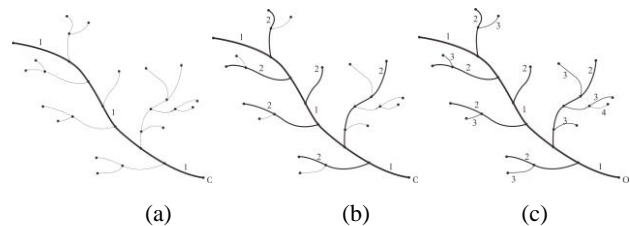

| (a) | (b) | (c) |

Figure 4. Procedures of automated coding system for dendritic river systems: (a) coding the main stem, (b) coding of tributaries and (c) final results of stream coding

4. EXPERIMENTS AND ANALYSIS

4.1 Experimental data and environment

The bottom up hydrologic coding system was embedded into the WJ-III mapping workstation developed by the Chinese Academy of Surveying and Mapping. The new stream coding system was experimentally validated via a comparison with the classic river entity-based stream coding system (i.e., Horton's hydrologic coding system). The experimental data correspond to part of a dendritic river system in a city in southern China. The original scale of these data is 1:200,000, and there are a total of 80 rivers in this dataset. The hardware environment for this experiment runs on a Microsoft Windows 7 64-bit operating system equipped with an Intel Core i7-3770 running at 3.2 GHz, with 16 GB of RAM and a 1,024 GB solid-state drive.

4.2 Comparative analysis on the features of each coding system

Fig. 5 illustrates the results of hierarchically coding the experimental river system using (a) Horton's coding system and (b) the new coding system. Here, the ranges of stream orders given by the new method and Horton's method are identical (i.e., a maximum of 6). However, the Horton system uses a top down approach (from the source to the estuary), whereas our system uses a bottom up approach (from the estuary to the source). Therefore, the stream orders given by our method are the exact opposite of those given by Horton's system. For example, if a stream was assigned a stream order of 1 by our system, the corresponding order given by Horton's system would be the maximum stream order (6 in this case). Horton's coding system can be used to directly indicate the subtree depth. For example, a stream that has a Horton stream order of 3 must contain a bifurcation of first-order and second-order tributaries, and the subtree depth of this stream is 3. Our coding system does not directly indicate the subtree depth of a stream, but it does provide a better indication of the stream hierarchies formed by catchment relationships.
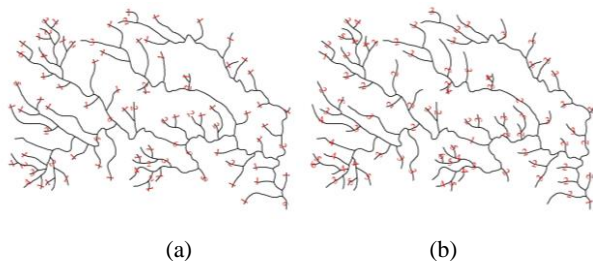

| (a) | (b) |

Figure 5. Comparison of the stream coding systems: (a) Horton's coding system and (b) Our coding system

The number of streams corresponding to each stream order, as provided by the aforementioned hydrologic coding systems, is shown in Table 1.

Stream coding method	Number and percentage of streams for each stream order						
	Order	1	2	3	4	5	6
Horton's coding system	Code	6	5	4	3	2	1
	Number	1	2	3	4	15	55
	Percent	1.25	2.5	4	5	19	69
Our coding system	Code	1	2	3	4	5	6
	Number	1	22	34	14	7	2
	Percent	1.25	27	43	17	8.75	2.5

Table 1 Comparison between two hydrologic coding systems in terms of the number of streams corresponding to each stream order

A histogram was plotted using the data in Table 1 to provide an intuitive description of the distribution of stream numbers for each stream order (Fig. 6).
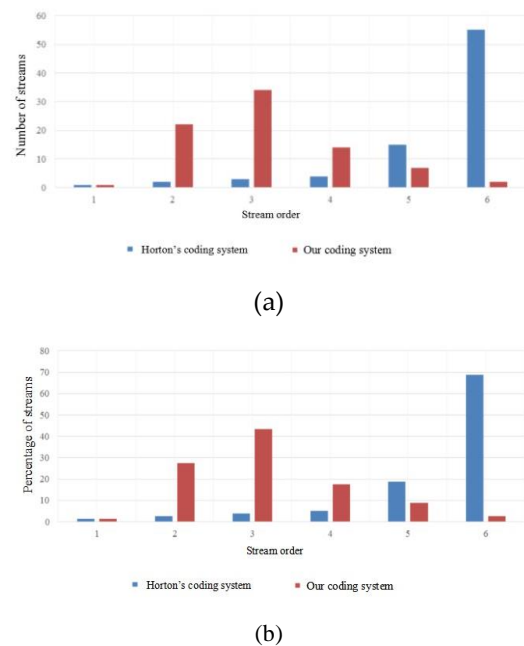

(a)

(b)

Figure 6. Comparison between Horton's coding system and our coding system in terms of the number of streams corresponding to each stream order: (a) a histogram comparing two hydrologic coding systems in terms of the number of streams corresponding to each stream order and (b) a histogram comparing two hydrologic coding systems in terms of the percentage of streams corresponding to each stream order

Although both stream coding systems result in the identification of one main stem, there are significant differences in the number of streams corresponding to each stream order, as shown in Table 1 and Fig. 6. In the Horton system, the number of streams decreases with stream order; first-order tributaries around the boundaries of the river system (i.e., the locations of river sources) account for 66.27% of all streams, which exceeds the sum of streams from all other stream orders. The distribution of streams corresponding to our coding system exhibits a bell curve (i.e., lower and higher order streams are fewer in number, whereas medium-order streams are greater in number), which is caused by differences in the identification of source tributaries (i.e., suspended arcs). In Horton's system, all source tributaries are defined as first-order streams; in our system, a source tributary could be assigned a higher order if it is connected to other higher order streams, which results in the assignment of a variety of stream orders (by our coding system) as source tributaries.

Therefore, our hydrologic coding system results in a larger proportion of higher order streams (streams with lower code numbers in our scheme) in the river system as a whole. However, when Horton's coding system was used to code the streams of the river system, lower order streams (streams with lower numbers in Horton's scheme) accounted for a much larger proportion of the streams in the river system.

4.3 Analyzing the efficacy of each hydrologic coding system in river selection and generalization

The stream code (order) is usually an important factor for determining streams that are selected during the generalization of dendritic river systems. Stream selections were performed using the classic stream selection method based on the stream

codes given by our system and Horton's system to compare the effectiveness of these hydrologic coding systems in stream selection. The procedure for the classic stream selection method is as follows (Stanislawski et al., 2011): (1) the number of streams that are selected with this process is first determined using the square root model (Töpfer, 1963; Töpfer, 1966; Regnauld, 2001). (2) The streams are coded to distinguish the stream order, and then stream selection is performed to decrease the tributaries and preserve main stems. Streams of the same order are distinguished according to length; longer streams are preserved, while shorter streams are removed. (3) The streams that are preserved are selected, one order after another, starting from the main stem and extending to the tributaries. This process continues until a sufficient number of streams (as defined by the square root model) have been selected.

Figs. 7(a), 8(a), and 9(a) are the results of the selections performed in three representative regions of the experimental data. Fig. 7(a) shows an ordinary river substream (i.e., all of the tributaries connected to the main stem have the same topologically depth). Fig. 8(a) shows a topologically shallow substream with uniformly distributed tributaries (i.e., the tributaries of the main stem are relatively shallow in terms of topologically depth, and their depths do not vary significantly). Fig. 9(a) shows a topologically deep substream (i.e., the tributaries connected to the main stem have very large differences in topological depth). These three regions were generalized to a scale of 1:500,000, and the selections obtained using the two aforementioned stream coding systems were comparatively analyzed.
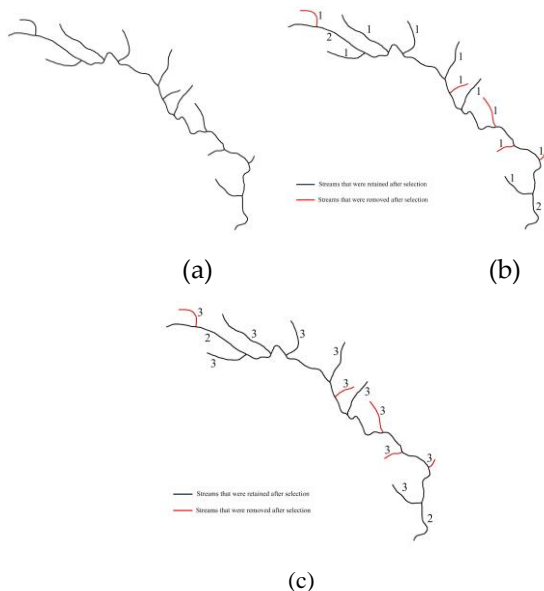

(a)   (b)

(c)

Figure 7. Comparison of selections in an ordinary river substream: (a) original map, (b) selections based on Horton's coding system and (c) selections based on our coding system

There are 12 streams in the substream shown in Fig. 7. Based on the square root model, 7 streams need to be preserved, while 5 streams have to be removed. The results of the selection processes are shown in Figs. 7(b) and 7(c). The hierarchies and depths of each tributary are almost the same as those in an ordinary river substream, and the same orders are assigned to each tributary by both of the aforementioned stream ordering systems. Thus, these tributaries are comparable to each other. In other words, both stream ordering systems resulted in the same stream selections in this scenario.
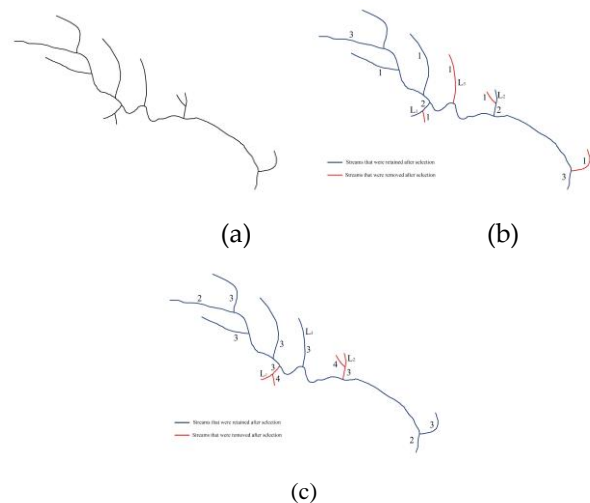

(a)   (b)

(c)

Figure 8. Comparison of stream selections for a topologically shallow substream with uniformly distributed tributaries: (a) original map, (b) selections based on Horton's coding system and (c) selections based on our coding system

Fig. 8 shows that there are a total of 10 streams in the substream. According to the square root model, 6 streams are retained, while 4 streams need to be removed; the results of the stream selections are shown in Figs. 8(b) and 8(c). The use of Horton's coding system in a topologically shallow substream with uniformly distributed tributaries resulted in the preservation of the L1 and L2 streams and the removal of the L3 stream. The L1 and L2 streams have a depth of two but are relatively short, whereas L3 is relatively long but only has a depth of 1. In general, Horton's stream coding system has failed to preserve the spatial structure of the original map. Conversely, our system resulted in the deletion of L1 and L2 and the retention of L3. Our stream coding system has succeeded in preserving the spatial structure of the original map.
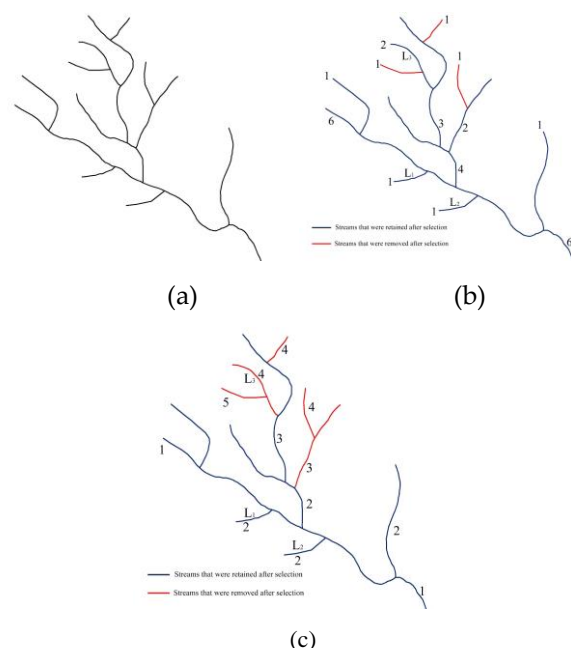

(a)   (b)

(c)

Figure 8. Comparison of stream selections in a topologically deep substream: (a) original map, (b) selections based on Horton's coding system and (c) selections based on our coding system

There are 12 streams in the substream shown in Fig. 9. Based on the square root model, 7 streams have to be retained, while 5 streams have to be removed. The results of the stream selections are shown in Figs. 9(b) and 9(c). The use of Horton's coding system in a topologically deep substream resulted in the deletion of L1 and L2 and the preservation of L3. L1 and L2 are connected to the main stem but are relatively shallow. The L3 stream is also connected to the main stem and is topologically deep. These selections have preserved the topological structure of the original map. Conversely, our stream ordering system resulted in the retention of L1 and L2 and the deletion of L3, which disrupted the topological structure of the original map.

5. CONCLUSIONS AND DISCUSSION

Conventional hydrologic coding methods for dendritic river systems are generally based on top down approaches. However, when these hydrologic coding methods are used to process the data of large river systems, the resulting stream orders do not accurately reflect the hierarchical features of the stream data. Therefore, we have proposed an improved bottom up stream coding system for dendritic river systems, with the hierarchical features of dendritic river systems being the core focus of our method. The following conclusions were drawn from the validation of our system via comparisons with the classic Horton stream coding system:

(1) Our method uses a bottom up approach (i.e., from the estuary to the source) for stream coding, whereas the Horton stream coding system uses a top down approach (i.e., from the source to the estuary). Consequently, the stream orders produced by our method are the exact opposite of those produced by Horton's method.

(2) When our system and Horton's system are used to code the streams in some sections of a dendritic river network, the resulting ranges of the stream orders are always identical.

(3) Horton's stream coding system can be used to directly indicate subtree depth; although our stream ordering system cannot directly indicate the subtree depth of a stream, it does provide a better indication of the stream hierarchies produced by catchment relationships.

(4) In stream selection and generalization, Horton's stream coding system is better suited for dendritic river systems that have topologically deep structures, whereas our stream coding system is better suited for topologically shallow dendritic river systems with a uniform distribution of tributaries. Both stream coding systems result in identical stream selections when used in ordinary river systems.

Our future research will focus on combining the advantages of these stream coding systems to construct a hybrid hydrologic coding system that is capable of indicating subtree depth while accounting for the stream hierarchies of dendritic river systems.

REFERENCES

He, Z. Principle and Method of Map Data Processing Model. Wuhan University Press, 2004. (in Chinese)

Zhang, L., Wang, G., Dai B., Li T., 2007. Classification and codification methods of stream network in a river basin, a review. Environmental Informatics Archives, 5, 364–372.

Jasiewicz, J., Metz, M.,2011. A new GRASS GIS toolkit for Hortonian analysis of drainage networks. Computers & Geosciences, 37(8): 1162-1173.

Harish, N., Kumar, P., Raja, M., et al., 2016. Remote sensing and GIS in the morphometric analysis of macro-watersheds for hydrological Scenario assessment and characterization-A study on Penna river sub-basin. SPSR Nellore district, India.

Moharir, K., Pande, C., 2014. Analysis of morphometric parameters using Remote-sensing and GIS techniques in the lonar nala in Akola district Maharashtra India. Int J Tech Res Eng, 1(10).

Sen, A., Gokgoz, T., 2015. An experimental approach for selection/elimination in stream network generalization using support vector machines. Geocarto International, 30(3): 311-329.

Khatun, F., Sharma, P., 2018. Strahler Order Classification and Analysis of Drainage Network by Satellite Image Processing. Advances in Communication, Devices and Networking. Springer, Singapore, 915-922.

Horton, R., 1945. Erosional development of streams and their drainage basins; hydro physical approach to quantitative morphology. Geological society of America bulletin, 56(3): 275-370.

Strahler, A., 1957. Quantitative analysis of watershed geomorphology. Eos, Transactions American Geophysical Union, 38(6): 913-920.

Shreve, R., 1966. Statistical law of stream numbers. The Journal of Geology, 74(1): 17-37.

Gravelius, H., 1914. Flusskunde, Goschen'sche Verlagshandlung. Berlin.

Thomson, R., Brooks, R., 2001. Exploiting perceptual grouping for map analysis, understanding and generalization: The Case of Road and River Networks. International Workshop on Graphics Recognition. Springer, Berlin, Heidelberg, 148-157.

Thomson, R., Brooks, R., 2000. Efficient generalization and abstraction of network data using perceptual grouping. Proceedings of the 5th International Conference on Geo-Computation, Chatham, 23-25.

Wu, J., Deng, M., Liu, H., 2013. An Integrated Model to Represent Topological Relation and Directional Relation Between Directed Line Objects. Geomatics and Information Science of Wuhan University, (11): 1358-1363. (in Chinese)

Zhang, Y., Li, L., Jin, Y., et al., 2004. Structured Design of Dendritic River Networks Based on Graph.Geomatics and

Information Science of Wuhan University, 29(6):537-539. (in Chinese)

Toepfer, F., 1963. Study of the application of the principle of square root in cartographic generalization. Beijing: Surveying and Mapping Press.

Töpfer, F., Pillewizer, W.,1966. The principles of selection, a means of cartographic generalization. Cartographic Journal, 1966.
REGNAULD N.,2001. Contextual Building Typification in Automated Map Generalization. Algorithmica, 30(2): 312-333.

Stanislawski, L., Savino, S., 2011. Pruning of hydrographic networks: a comparison of two approaches. Proceedings of the 14th ICA workshop on generalization and multiple representation, jointly organized with ISPRS Commission II/2 Working group on Multiscale Representation of Spatial Data, Paris.

Horsfield, K., Cumming G., 1968. Morphology of the bronchial tree in man. Journal of applied physiology, 24(3): 373-383.

Gülgen, F., 2017. A stream ordering approach based on network analysis operations. Geocarto international, 32(3): 322-333.

Revised August 2019