Multiple Equidistant Belt Technique for Width Function Estimation through A Two-Segmented-Distance Strategy

Pengfei Wu[1], Jintao Liu[1], Meiyan Feng[1], and Hu Liu[2]

[1]Hohai University
[2]Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences

February 20, 2023

Abstract

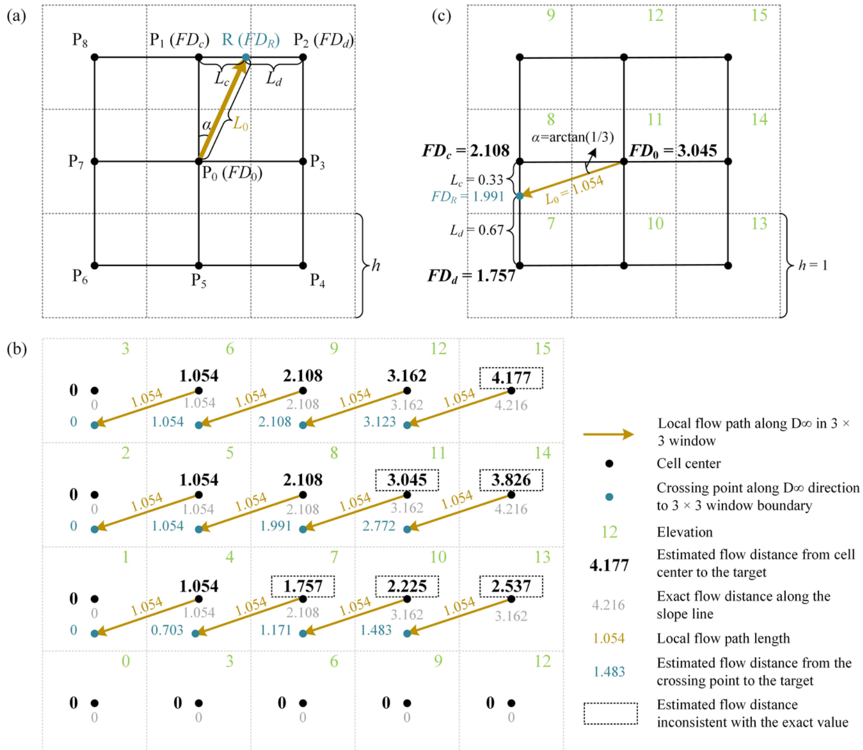
The arbitrary adoption of cell center to represent the whole cell is a compromise to the grid structure of the digital elevation models (DEMs), which greatly limits the accuracy of estimating flow distance and width functions. This study uses the triangulation with linear interpolation (TLI) method to approximate the missing flow distance values within a cell except for the cell center. A new flow distance algorithm (D[?]-TLI) is proposed to improve the flow distance estimation by using a two-segment-distance strategy. The first segment distance from a cell center to a crossing point at the local 3 × 3 window boundary is modeled by the D[?] method. The second segment distance souring from the crossing point is estimated by the TLI using the flow distance values assigned for the two closest downstream cell centers, while these values have been assigned by iterating from lowest to highest cells. Then, using the continuous flow distance field approximated over a cell region, this cell can be divided into multiple equidistant belts (MEB) to estimate the width function. Four numerical terrains and two real-world terrains are used for assessments. The results demonstrate that D[?]-TLI outperforms nine existing flow distance algorithms over any numerical terrains, and it is overall optimal for real-world terrains. Meanwhile, MEB extracts the width function which is less affected by unreasonable artificial fluctuation than the previous method. Hence, MEB combined with D[?]-TLI can obtain a high-accuracy estimation of hydro-geomorphological attributes that may be conducive to the application of hydrologic modeling.

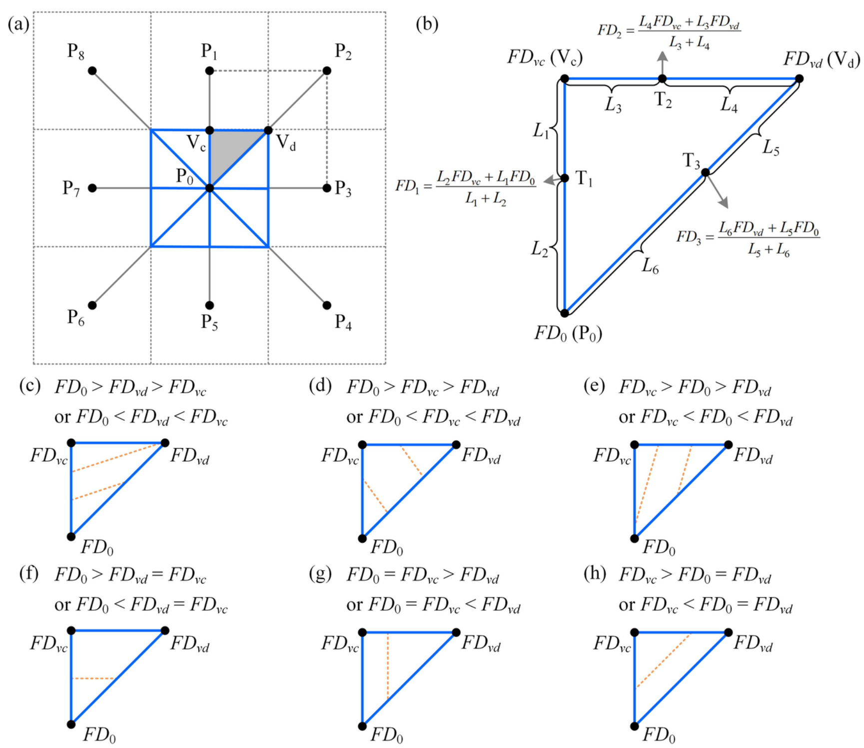
Hosted file

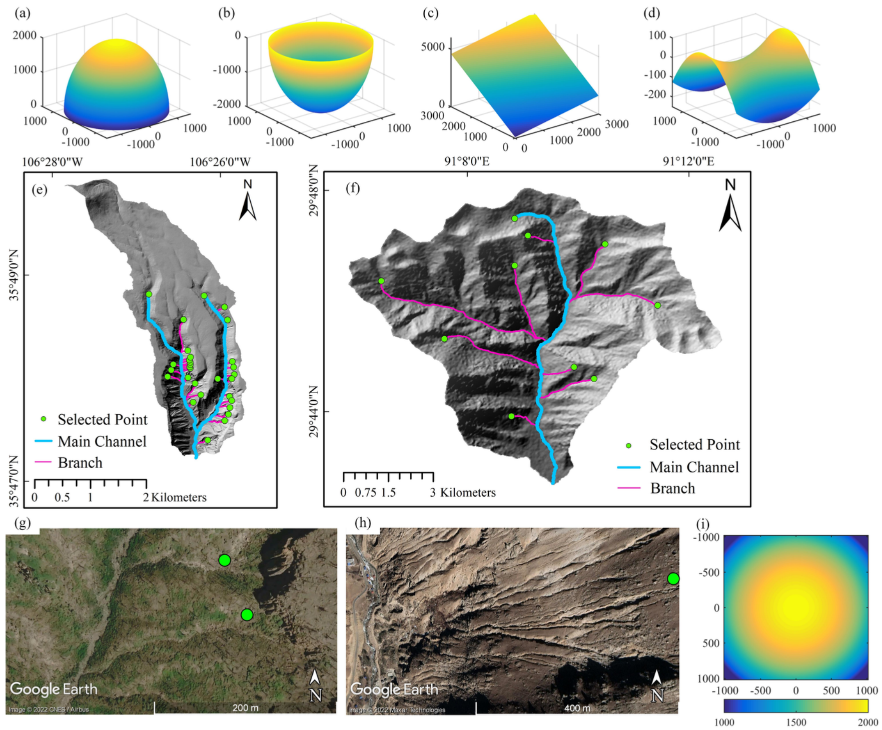
955816_0_art_file_10686706_rppv9y.docx available at https://authorea.com/users/546750/articles/625179-multiple-equidistant-belt-technique-for-width-function-estimation-through-a-two-segmented-distance-strategy

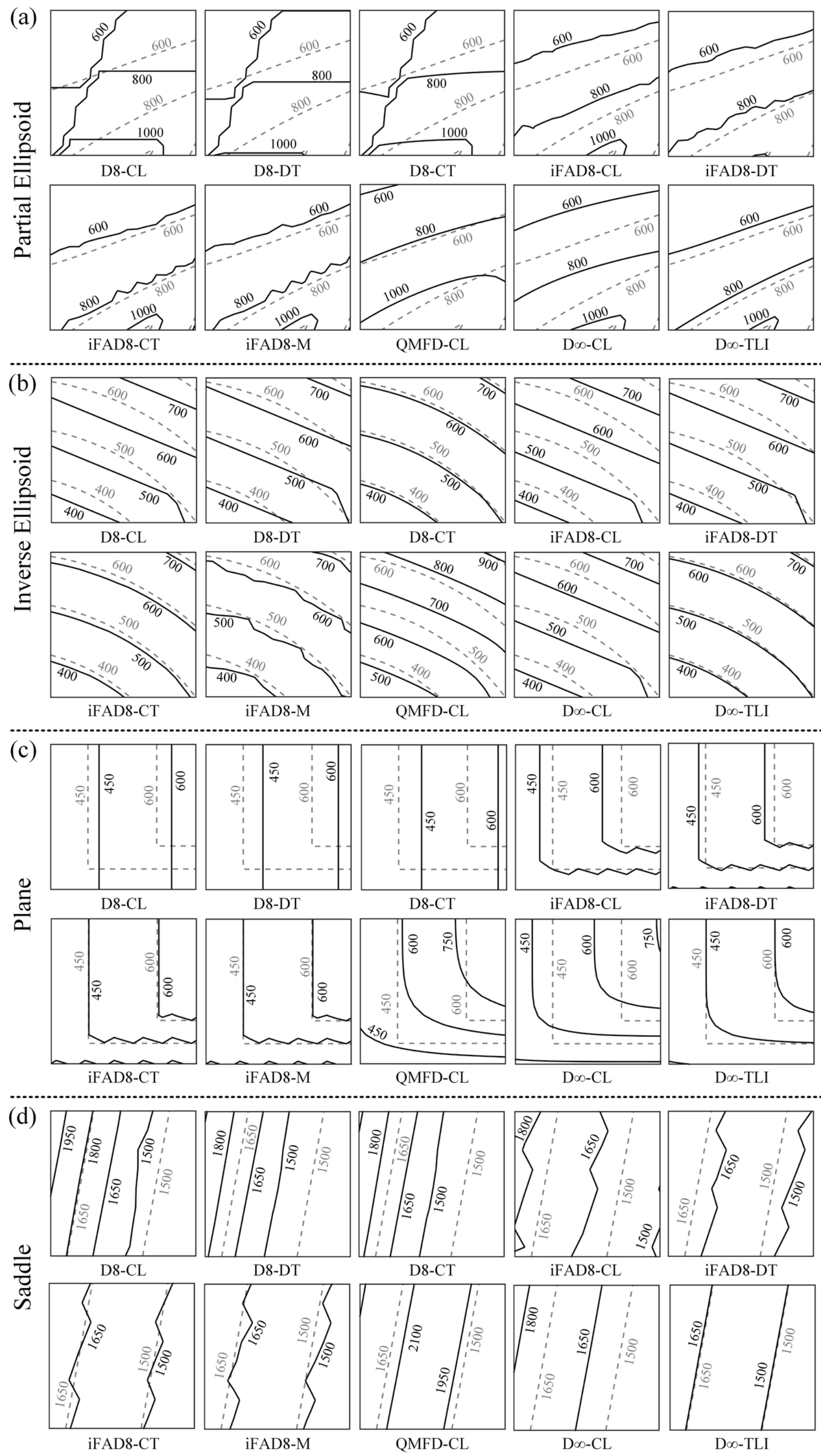
Hosted file

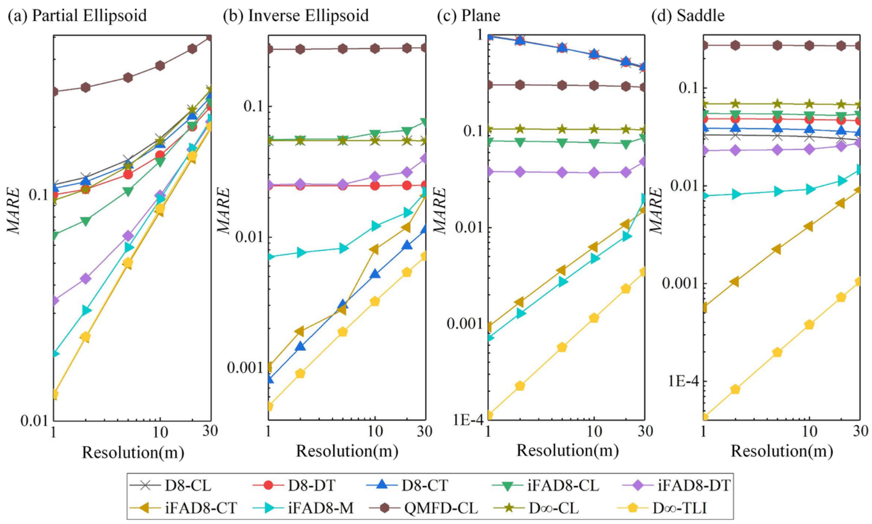
955816_0_supp_10686741_rpptt0.docx available at https://authorea.com/users/546750/articles/625179-multiple-equidistant-belt-technique-for-width-function-estimation-through-a-two-segmented-distance-strategy

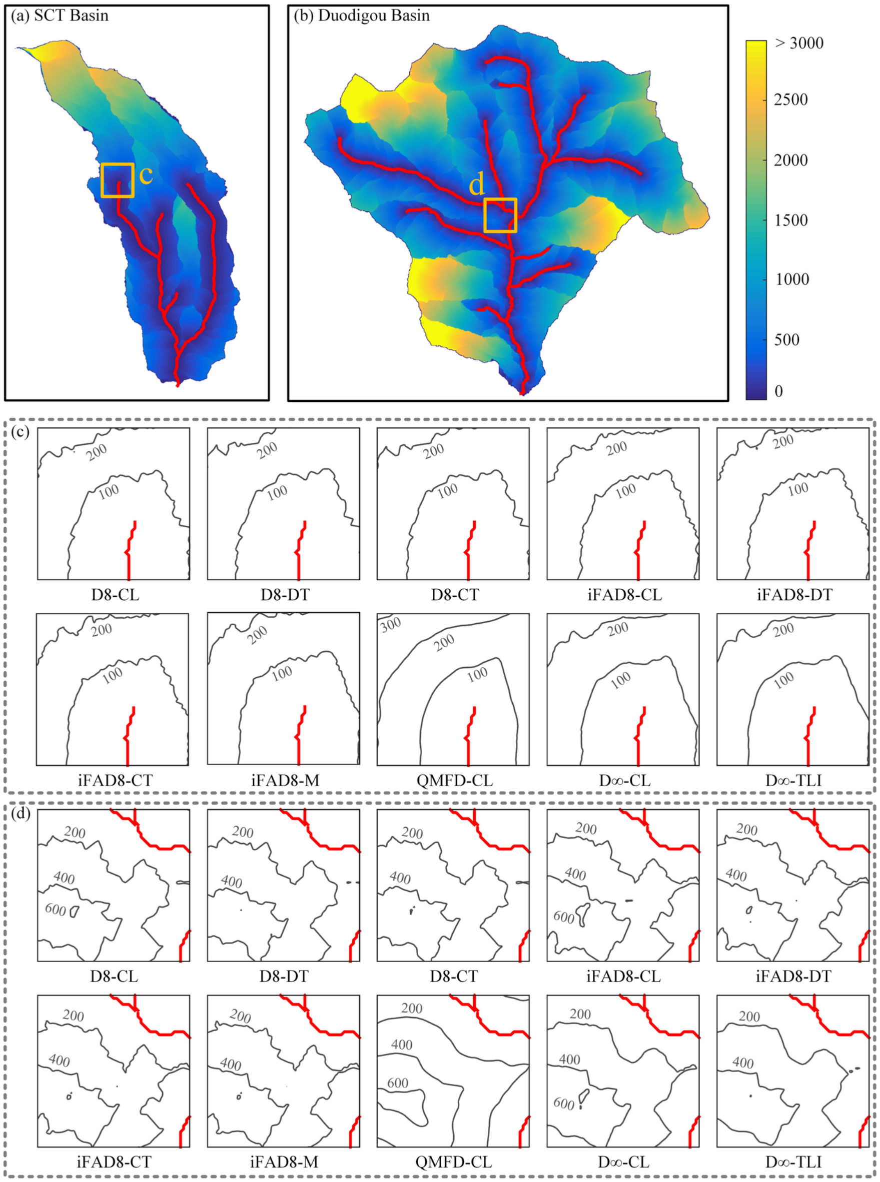

(a)
P8
P1 (FDc)
R (FDR)
P2 (FDd)
Lc
Ld
α
L0
P7
P0 (FD0)
P3
P6
P5
P4
h
(c)
FDc = 2.108
α=arctan(1/3)
FD0 = 3.045
Lc = 0.33
FDR = 1.991
L0 = 1.054
Ld = 0.67
FDd = 1.757
h = 1
(b)
Local flow path along D∞ in 3 × 3 window
Cell center
Crossing point along D∞ direction to 3 × 3 window boundary
12 Elevation
4.177 Estimated flow distance from cell center to the target
4.216 Exact flow distance along the slope line
1.054 Local flow path length
1.483 Estimated flow distance from the crossing point to the target
Estimated flow distance inconsistent with the exact value

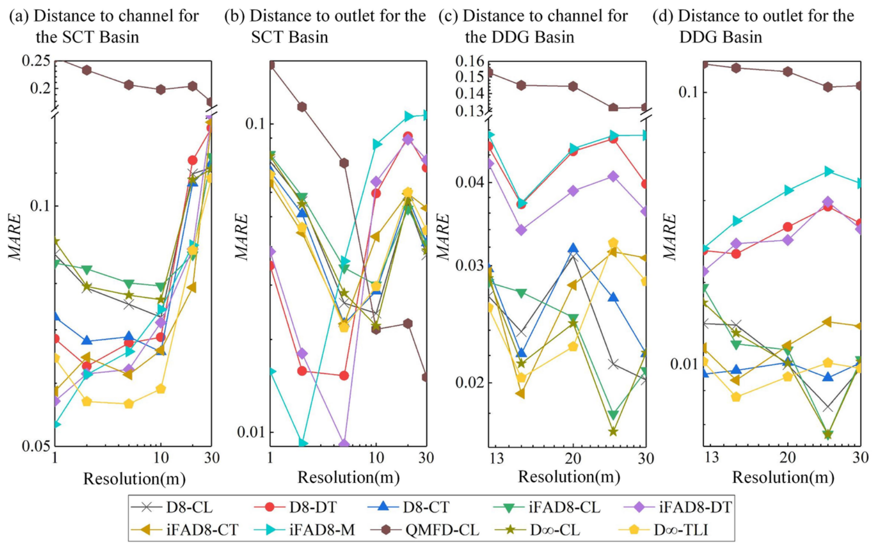

(a)
P8
P1
P2
Vc
Vd
P0
P7
P3
P6
P5
P4
(b)
FD2 = (L4FDvc + L3FDvd)/(L3 + L4)
FDvc (Vc)
FDvd (Vd)
L3
T2
L4
L1
L5
T3
FD1 = (L2FDvc + L1FD0)/(L1 + L2)
T1
FD3 = (L6FDvd + L5FD0)/(L5 + L6)
L2
L6
FD0 (P0)
(c) FD0 > FDvd > FDvc or FD0 < FDvd < FDvc
(d) FD0 > FDvc > FDvd or FD0 < FDvc < FDvd
(e) FDvc > FD0 > FDvd or FDvc < FD0 < FDvd
(f) FD0 > FDvd = FDvc or FD0 < FDvd = FDvc
(g) FD0 = FDvc > FDvd or FD0 = FDvc < FDvd
(h) FDvc > FD0 = FDvd or FDvc < FD0 = FDvd

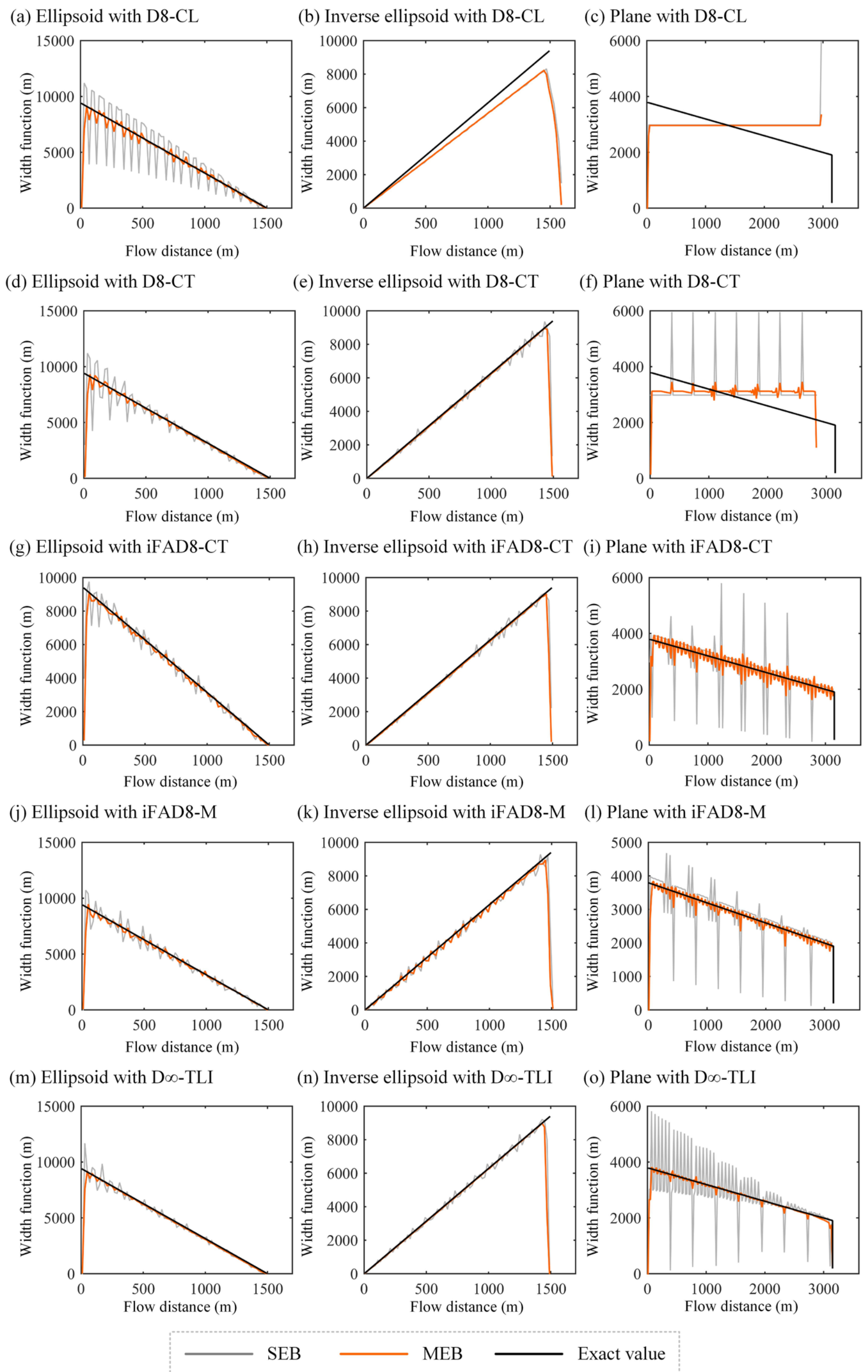

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
106°28'0"W
106°26'0"W
35°49'0"N
35°47'0"N
91°8'0"E
91°12'0"E
29°48'0"N
29°44'0"N
Selected Point
Main Channel
Branch
0 0.5 1 2 Kilometers
0 0.75 1.5 3 Kilometers
Google Earth
200 m
400 m
N

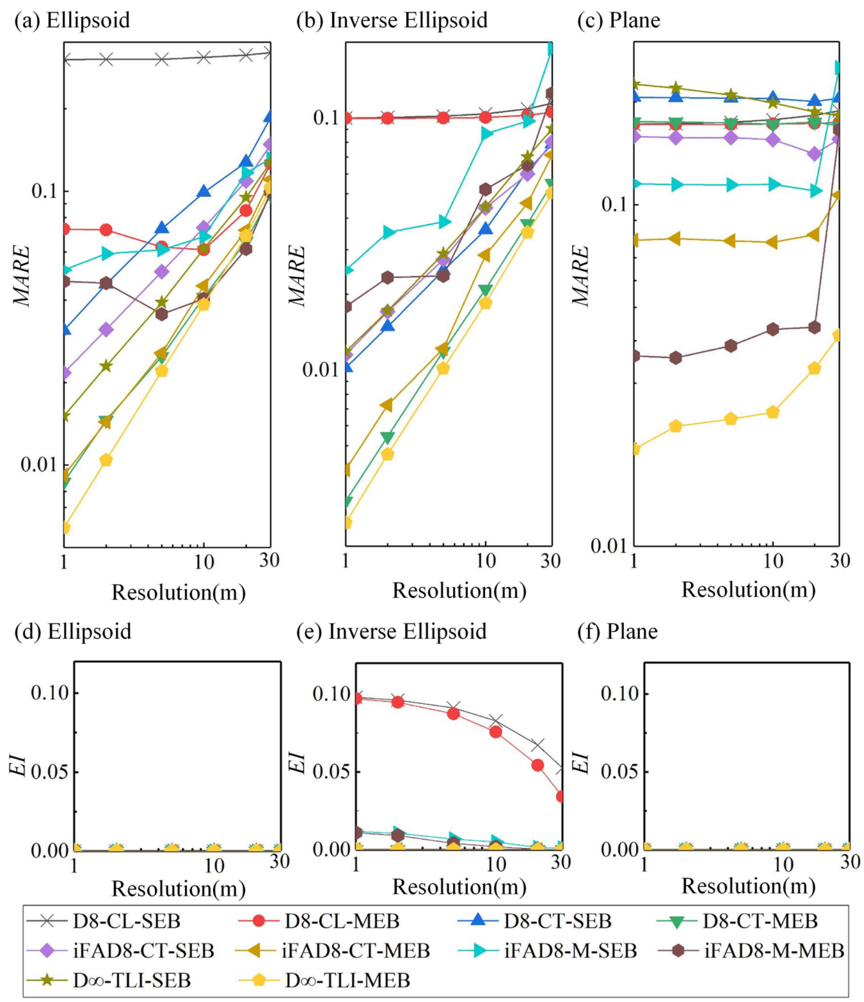

(a)
Partial Ellipsoid
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(b)
Inverse Ellipsoid
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(c)
Plane
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(d)
Saddle
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI


(a) Partial Ellipsoid
(b) Inverse Ellipsoid
(c) Plane
(d) Saddle
MARE
Resolution(m)
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI

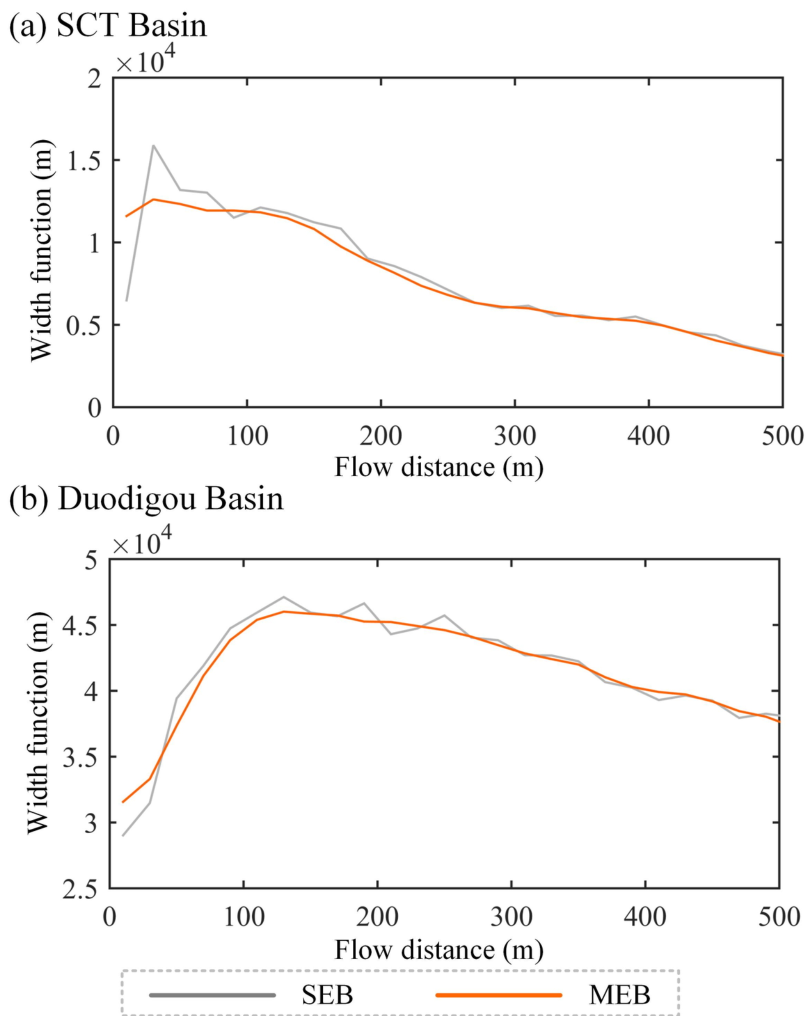

(a) SCT Basin
c
(b) Duodigou Basin
d
> 3000
2500
2000
1500
1000
500
0
(c)
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(d)
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI


(a) Distance to channel for the SCT Basin
(b) Distance to outlet for the SCT Basin
(c) Distance to channel for the DDG Basin
(d) Distance to outlet for the DDG Basin
MARE
Resolution(m)
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI


(a) Ellipsoid with D8-CL
(b) Inverse ellipsoid with D8-CL
(c) Plane with D8-CL
(d) Ellipsoid with D8-CT
(e) Inverse ellipsoid with D8-CT
(f) Plane with D8-CT
(g) Ellipsoid with iFAD8-CT
(h) Inverse ellipsoid with iFAD8-CT
(i) Plane with iFAD8-CT
(j) Ellipsoid with iFAD8-M
(k) Inverse ellipsoid with iFAD8-M
(l) Plane with iFAD8-M
(m) Ellipsoid with D∞-TLI
(n) Inverse ellipsoid with D∞-TLI
(o) Plane with D∞-TLI
Width function (m)
Flow distance (m)
SEB
MEB
Exact value



(a) Ellipsoid
(b) Inverse Ellipsoid
(c) Plane
MARE
0.1
0.01
1
10
30
Resolution(m)
(d) Ellipsoid
(e) Inverse Ellipsoid
(f) Plane
EI
0.10
0.05
0.00
D8-CL-SEB
D8-CL-MEB
D8-CT-SEB
D8-CT-MEB
iFAD8-CT-SEB
iFAD8-CT-MEB
iFAD8-M-SEB
iFAD8-M-MEB
D∞-TLI-SEB
D∞-TLI-MEB



(a) SCT Basin
×10^4
Width function (m)
Flow distance (m)
0
0.5
1
1.5
2
0
100
200
300
400
500
(b) Duodigou Basin
×10^4
Width function (m)
Flow distance (m)
2.5
3
3.5
4
4.5
5
0
100
200
300
400
500
SEB
MEB

Multiple Equidistant Belt Technique for Width Function Estimation through A Two-Segmented-Distance Strategy

Pengfei Wu[1,2], Jintao Liu[1,2*], Meiyan Feng[1,2], Hu Liu[3]

---

[1]State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210024, China

[2]College of Hydrology and Water Resources, Hohai University, Nanjing 210024, China

[3]Linze Inland River Basin Research Station, Chinese Ecosystem Research Network, Lanzhou 730000, China

---

* Corresponding author: jtliu@hhu.edu.cn (J. Liu).

Key Points

- Triangulation with linear interpolation (TLI) is adopted to approximate the flow distance values for a cell region except for its center.
- Two segmented-distances inside and outside a local 3 × 3 window are severally modeled by D∞ and TLI.
- Estimation of width function can be essentially improved with possible multiple equidistant belts technique over a cell.

Abstract

The arbitrary adoption of cell center to represent the whole cell is a compromise to the grid structure of the digital elevation models (DEMs), which greatly limits the accuracy of estimating flow distance and width functions. This study uses the triangulation with linear interpolation (TLI) method to approximate the missing flow distance values within a cell except for the cell center. A new flow distance algorithm (D∞-TLI) is proposed to improve the flow distance estimation by using a two-segment-distance strategy. The first segment distance from a cell center to a crossing point at the local 3 × 3 window boundary is modeled by the D∞ method. The second segment distance souring from the crossing point is estimated by the TLI using the flow distance values assigned for the two closest downstream cell centers, while these values have been assigned by iterating from lowest to highest cells. Then, using the continuous flow distance field approximated over a cell region, this cell can be divided into multiple equidistant belts (MEB) to estimate the width function. Four numerical terrains and two real-world terrains are used for assessments. The results demonstrate that D∞-TLI outperforms nine existing flow distance algorithms over any numerical terrains, and it is overall optimal for real-world terrains. Meanwhile, MEB extracts the width function which is less affected by unreasonable artificial fluctuation than the previous method. Hence, MEB combined with D∞-TLI can obtain a high-accuracy estimation of hydro-geomorphological attributes that may be conducive to the application of hydrologic modeling.

Keywords

Flow distance; Width function; Two-segmented-distance strategy; Triangulation with linear interpolation; Multiple equidistant belt technique

1. Introduction

As an important feature of overland flow, flow distance is essential for hydrological, geomorphological, and ecological research, such as runoff or flood analysis (Bogaart & Troch, 2006; Di Lazzaro et al., 2016; Liu et al., 2012; McGuire et al., 2005; Muzik, 1996; Rinaldo et al., 1991; Xu et al., 2018), soil erosion or thickness simulation (Dong et al., 2022; Hickey, 2000; Tesfa et al., 2009), and water quality modeling (Fan et al., 2015). These researches rely on flow distance estimations with different scales, including distance to channel or outlet (Bogaart & Troch, 2006; Van Nieuwenhuizen, 2021), river length (Fan et al., 2015) and uphill slope line length (Dong et al., 2022; Tesfa et al., 2009). Meanwhile, as a form of flow distance distribution, width function of a hillslope or catchment is always used as a hydrologic response function in hydrologic modeling (Bogaart & Troch, 2006; Gupta et al., 1986; Hazenberg et al., 2015; Lapides et al., 2022; Liu et al., 2016; Moussa, 2008; Noël et al., 2014; Ranjram & Craig, 2021; Rigon et al., 2016; Troch et al., 2002, 2003). Thus, there has been renewed interest in algorithms to accurately estimate flow distance as well as width function.

As the discretized representation of terrains by the grid digital elevation models (DEMs), flow distance is always estimated by cumulating the length along the predicted DEM-based flow path (Mayorga et al., 2005). Flow path, according to whether a cell is allowed to drain to more than one downslope cell, can be estimated by two types of flow direction algorithms, i.e., the single flow direction (SFD) algorithms and the multiple flow direction (MFD) algorithms (Wilson et al., 2007).

The SFD path uses a zigzag line whose flow distance can be measured explicitly. However, the MFD path is a dispersive network whose flow distance must be computed implicitly as the weighted average length of all the lines based on the flow proportion distributed from the beginning cell along every line (Bogaart & Troch, 2006).

Once the spatial distribution of flow distance is computed, the width function can always be estimated implicitly (e.g., Liu et al., 2012; Sahoo & Sahoo, 2019b). In this implicit method, the width function is defined as an area distribution function (Moussa, 2008; Veneziano et al., 2000) or probability density function (Bogaart & Troch, 2006; Liu et al., 2012; Sahoo & Sahoo, 2019a, 2019b) of the equidistant belts. Compared with the explicit method that directly considers contour lengths as width functions (Fan & Bras, 1998), the implicit method is more suitable for applications in the real-world hillslopes or catchments (Sahoo & Sahoo, 2019b). Moreover, it can provide width functions with more details when compared with the strategies which obtain monotonic width functions by simplifying the terrains into regular shapes (e.g., Noël et al., 2014).

However, the discretized grid structure of DEM has naturally limited the accuracy of the predicted flow path, it thus further constrains the precision of the flow distance estimation as well as the width function estimation (Liu et al., 2012; Paik, 2008; Paz et al., 2008; Wu et al., 2020). For instance, most flow direction algorithms were designed to fix the flow path out from a local cell center down to one or more other downstream cell centers (e.g., O'Callaghan & Mark, 1984; Quinn et al., 1991;

Orlandini et al., 2003; Shin & Paik, 2017; Tarboton, 1997; Paik, 2008; Wu et al., 2020), but the true path may miss these centers (Paik, 2008). So, the flow distance along a predicted flow path to a given target may be inconsistent with the true flow distance (Paz et al., 2008; Liu et al., 2012). A solution to this problem is to employ some highly accurate flow direction algorithms to track the gravity-driven flow path which is essentially not forced to pass the downstream cell centers (e.g., Zhou et al., 2011). However, this solution does not apply to large scales because the related flow direction algorithms require vast computing time as well as storage space (Zhou et al., 2011).

In order to ensure acceptable computational efficiency, the flow distance assignment for each cell center is better to only search in a local window rather than the whole flow path. Bogaart and Troch (2006) proposed such a two-segmented-distance framework which only models the sub-distance from the cell center to a crossing point at a 3×3 window boundary and adds it to the sub-distance sourcing from the crossing point to a given target (e.g., channel) for the final flow distance value. But this method only adopted center-to-center flow path previously because the true crossing points inconsistent with cell centers may be not assigned with flow distance values, although some strategies can restore the true sub-distance in the 3×3 window accurately (e.g., Butt & Maragos, 1998; De Smith, 2004; Liu et al., 2012; Paz et al., 2008). Hence, the lack of flow distance values at a cell region except for its center seems to limit the improvements in flow distance estimation.

Meanwhile, the missing flow distance values prevent it from dividing a cell covering multiple equidistant belts into correct numbers of belts. So the conventional method adds the whole cell into merely one equidistant belt based on the flow distance value assigned for the cell center. Then imprecise equidistant belt area function will lead to width function with unreasonable artificial fluctuation because some equidistant belts may capture too many regions from adjacent belts as shown in the results of some existing studies (e.g., Moussa, 2008; Liu et al., 2012; Sahoo & Sahoo, 2019b; Veneziano et al., 2000).

In fact, following common geographical studies, the missing flow distance values at a cell region except for its center can be approximated using the values assigned to cell centers with an interpolation method (e.g., Lee & Schachter, 1980; Polidori & Chorowicz, 1993; Schwendel et al., 2012; Yilmaz, 2007). However, there is no existing study trying to adopt the interpolation method to improve the accuracy of flow distance estimation as well as width function estimation. Hence, this study attempts to incorporate an interpolation method, i.e., the triangulation with linear interpolation (TLI) method (Sloan, 1987; Yilmaz, 2007), into the flow distance algorithm and the width function algorithm for better estimation accuracy. Here TLI is selected firstly because of its high precision (Schwendel et al., 2012). In addition, TLI can provide straight isolines in a cell region, dividing the cell into regular equidistant belts whose area can be measured explicitly for the width function. Finally, both the numerical and the real-world terrains with different resolutions are adopted for the comparison between the proposed and the existing algorithms.

2. Methodology and Experiments

2.1. Flow Distance Estimation Combining D∞ And TLI

The new flow distance algorithm requires a DEM with flats and depressions removed. Each cell center at the channel network or any other given target such as outlet is assigned with a flow distance value equal to zero. Then all the cells without flow distance values are sorted into a queue in an ascending elevation order, and the flow distances from cell centers in the queue are calculated with elevation from low to high. This framework inherited from the method by Bogaart and Troch (2006) can guarantee that flow distances from all the neighboring downstream cell centers surrounding the current cell to channel are known. The cells with the same elevation can be processed in an arbitrary order because the flow will not be drained to a neighboring cell with the same elevation when the flats are removed.

Figure 1a shows a sketch to calculate the flow distance from a cell center (P_0) to channel (or other target). Firstly, a 3 × 3 window is built using the current cell and its eight neighbors. The D∞ method proposed by Tarboton (1997) is adopted for the local flow path in the window following some existing literature (Liu et al., 2012; Orlandini et al., 2003; Shin & Paik, 2017). Then the crossing point (i.e., R) along the local flow path to the window boundary can be identified. Based on the D∞ theory, the closest cardinal and diagonal neighboring cell centers (i.e., P_1 and P_2 in Figure 1a) to R are always lower than P_0, so they have been assigned with flow distance values according to the framework iterating from low to high.

Although TLI cannot provide the whole flow distance distribution in the window before the assignment of flow distance for P_0, it can assign the flow distance value for any point on the line P_1P_2 with the existing flow distance values assigned for these two cell centers. The flow distance (FD_R) from R to channel is computed by TLI as following:

$$FD_R = \frac{L_d FD_c + L_c FD_d}{L_c + L_d} \quad (1)$$

where FD_c and FD_d denote the flow distance from the nearest cardinal cell center (i.e., P_1) and the nearest diagonal cell center (i.e., P_2) to channel, respectively. L_c and L_d denote the length from R to the nearest cardinal cell center (i.e., P_1) and the nearest diagonal cell center (i.e., P_2), respectively. L_c and L_d can be calculated as following:

$$L_c = h\tan\alpha \quad (2)$$

$$L_d = h(1 - \tan\alpha) \quad (3)$$

where h denotes the resolution of DEM, and α denotes the angle between the D∞ direction and the closest cardinal direction. There is a special case when the D∞ direction points to a neighboring cell center. In this case, the other nearest cell center may be higher than P_0, and here its flow distance value is unknown. It makes no difference because FD_R is equal to the flow distance value assigned for the cell center pointed by the D∞ direction according to Equation 1 in this case. Then the local flow path length (L_0) by D∞ is calculated as:

$$L_0 = \frac{h}{\cos\alpha} \quad (4)$$

Finally, the flow distance (FD_0) from P_0 to channel is defined as:

$$FD_0 = FD_R + L_0 \quad (5)$$

This new algorithm combines D∞ and TLI, so it is referred as D∞-TLI. A DEM of a plane with a gradient ratio of 3:1 is provided as an example in Figure 1b to show the capacity of D∞-TLI to restore the flow distance to the DEM side. The estimated local flow path length, as well as both the estimated flow distances from the crossing point and the cell center to the DEM side, is labelled for each cell in Figure 1b. Moreover, an enlarged 3 × 3 window in Figure 1c is used to show the detailed assignment of flow distance for a cell center following Equation 1-5.

According to the estimated flow distance distribution (black values) in Figure 1b, the difference between the estimated flow distance and the exact slope line length is little for most cell centers. Here the exact slope line originating from a cell center is consistent with the gravity-driven flow path (Maxwell, 1870; Orlandini et al., 2014). This example only shows the potential of D∞-TLI to provide the flow distance distribution for the plane. However, whether more bias may appear when D∞-TLI is applied to other terrains (e.g., divergent, or convergent terrains) still needs further verifications. Hence, multiple terrains (see in section 2.3.1) with different complexities are adopted to assess D∞-TLI in our experiments.

2.2. Width Function Calculation with TLI

The width function coincides with the area distribution function of equidistant belt (Moussa, 2008), so the proposed algorithm extracts the equidistant belt firstly. In this step, cells in the specific hillslope or catchment are processed one by one. For each cell, the proposed algorithm firstly obtains the equidistant lines whose flow distances to channel are multiple to the belt interval, then each area between two equidistant lines is added to the corresponding equidistant belt. Here the belt interval is always set to be the DEM resolution (e.g., Liu et al., 2012; Moussa, 2008; Sahoo & Sahoo, 2019b).

To implement above scheme, according to some existing flow direction algorithms (e.g., Pilesjö & Hasan, 2014; Tarboton, 1997), a given cell is firstly divided into eight triangular facets as marked in blue in Figure 2a. Then every facet is processed independently. The grey facet in Figure 2a is taken for illustration. There are two vertexes (i.e., V_c and V_d) except the current cell center (i.e., P_0) within this facet. V_c is the midpoint of the given cell center (P_0) and its cardinal neighbor center (P_1 for the selected facet), and V_d is the midpoint of the given cell center (P_0) and its diagonal neighbor center (P_2 for the selected facet).

Equidistant lines (dotted lines in Figure 2c-2h) should be obtained firstly. However, if any one of the neighboring cells (i.e., P_1 or P_2) belong to another hillslope or catchment, this facet will not be further divided to avoid the effects of unreasonable critical lines. Then this facet is added to the belt covering the flow distance from the cell center (P_0). This step ensures that the integration of the width function is equal to the area of the extracted hillslope or catchment.

When all the three vertexes belong to the same hillslope or catchment, the continuous flow distance field for the facet is calculated using the TLI method. Here the flow distance from P_0, P_1, P_2, P_3 are expressed as FD_0, FD_1, FD_2, FD_3, and the mean of these four flow distance values is expressed as *MD*. According to the TLI, the flow distance from V_c to channel (i.e., FD_{vc}) can be calculated following:

$$FD_{vc} = (FD_0 + FD_1)/2 \tag{6}$$

According to Zhou et al. (2011) method, the flow distance from V_d to channel (i.e., FD_{vd}) is calculated as below:

$$FD_{vd} = \begin{cases} (FD_1 + FD_3)/2, \ if \ |MD - (FD_1 + FD_3)/2| \le |MD - (FD_0 + FD_2)/2| \\ (FD_0 + FD_2)/2, \ if \ |MD - (FD_1 + FD_3)/2| > |MD - (FD_0 + FD_2)/2| \end{cases} \quad (7)$$

Then the equations for the flow distance from any point at the facet sides to channel are shown in Figure 2b. These equations can also help to obtain the point position at a side with a given flow distance. Thus, when the given flow distance ranges between the minimum and the maximum flow distances from the three vertexes, only two points owning the given distance can be found at the three sides, and the equidistant line can be approximated by a straight line linking them (Figure 3c-3h). When the flow distances from the three vertexes to channel are different from each other, one point must be located on the side whose two vertexes own the minimum and the maximum flow distances respectively, and the second point can be located on one of the other two sides (Figure 3c-3e). Otherwise, the equidistant line should be parallel to the side linking two vertexes whose flow distances are equal (Figure 3f-3h). Thereupon, the facet can be divided into multiple equidistant belts between the equidistant lines.

After all the cells in the hillslope or catchment are processed, the probability density ($p(x)$) and the area distribution of equidistant belt are generated. Here the width of every equidistant belt is the ratio of the equidistant belt area ($Sp(x)$) to the belt interval (L_u), i.e., $Sp(x)/L_u$, and is defined as the width at the middle flow distance of the belt, where S is the total area of the hillslope or catchment. Hence, the applications of area distribution and probability density function to obtain the width function are consistent. This proposed division method for width function is referred as the multiple equidistant belt (MEB) method, which is different from the conventional method adds a cell into single equidistant belt (SEB).

2.3. Experiment Materials and Assessment Criteria

2.3.1. Numerical and Real-world Terrains

Four numerical terrains and two real-world terrains are adopted for algorithm assessments (Figure 3). The numerical terrains contain an ellipsoid, an inverse ellipsoid, a plane, and a saddle (Figure 3a-3d). These terrains represent the divergent, the convergent, the plain and the complex terrains, respectively. The formulas proposed by Li et al. (2021) are used to build these terrains with six resolutions (1 m, 2 m, 5 m, 10 m, 20 m, 30 m). As mentioned in Section 2.1, the exact flow distance is equal to the slope line length for the numerical terrains. This length can be calculated by integration using the slope line formulas introduced by Li et al. (2021), and then the exact equidistant belt area can be divided for the exact width function. When assessing the flow distance algorithms, only the partial ellipsoid with a square boundary in Figure 3i is used because there is a systematic error between the circle boundary of the complete ellipsoid and the valid DEM cells (Figure S1). But the assessments of the width function use the complete ellipsoid because it is hard to obtain the exact equidistant line or belt with integration when only the partial terrain is used. The saddle is not employed for the width function extraction due to the lack of the exact equidistant belt area.

It is difficult to obtain the exact flow distances from most positions over a real-world terrain to channel or outlet. However, flow distance from a point in a channel or gully can be measured along the overland flow trajectory. So, some channels (or gullies) in two real-world terrains, including a sub-basin of the Spruce Canyon (Figure 3e) and the Duodigou Basin (Figure 3f), are mapped using the images from the Google Earth for the assessments (e.g., Figure 3g and 3h).

The Spruce Canyon is in New Mexico, USA, while the selected tributary sub-basin (referred as SCT Basin) owns a drainage area of 7.0 km^2 with the elevation ranging from 2292 m to 3027 m. This basin has a relatively low mean slope (17.2°), while the valley bottoms and the channels are narrower than 100 m and 3 m, respectively. The downstream channels of thirty points are mapped. To avoid repetitive computation bias, only two longest channels are selected as the main channels, while other twenty-eight channels end up at the main channel and are regarded as branches. The lengths of two main channels are 3475 m and 3496 m, while the lengths of the branches range from 75 m to 687 m. Bare earth DEM data with 1 m resolution is provided by the Jemez River Basin Snow-off LiDAR Survey, and is resampled to five coarser resolutions (2 m, 5 m, 10 m, 20 m, 30 m) consistent with the selected resolutions of the numerical terrains.

The Duodigou Basin (referred as DDG Basin) is in the Tibetan Plateau, China (Fei et al., 2022). It is a steep alpine terrain covering 56.6 km^2 with the elevation ranging from 3719 m to 5425 m and the mean slope equal to 28.0°. The valley bottom is narrow (< 20 m) at upstream and wide (> 900 m) at downstream. The downstream channels of ten points are mapped. The longest channel (11386 m) is selected as the main channel. The lengths of the branches range from 936 m to 5979 m. These channels are narrower than 10 m. The 12.5 m-resolution DEM of this basin is obtained from the Advanced Land Observing Satellite (ALOS). Limited by the coarse initial resolution, the DEM is resampled to four resolutions (15 m, 20 m, 25 m, 30 m) with smaller intervals than those of other terrains to show the influence of the resolution.

2.3.2. Algorithm Assessments

Ten flow distance algorithms including D∞-TLI and nine other algorithms are adopted to assess the estimated flow distance distribution over the numerical and the real-world terrains, and their information are listed in Table 1. Here the traditional cumulative length (CL) method, the distance transform (DT) method by Paz et al. (2008), and the cosine transform (CT) method by Liu et al. (2012) are combined with two SFD algorithms, i.e., the classical D8 algorithm (O'Callaghan & Mark, 1984) and a highly accurate algorithm named iFAD8 (Wu et al., 2020). The merging (M) method by Dong et al. (2022) is also employed and combined with iFAD8. The cumulative length methods based on D∞ and QMFD proposed by Bogaart and Troch (2006) are also adopted. Hence, ten algorithms for comparison are D8-CL, D8-DT, D8-CT, iFAD8-CL, iFAD8-DT, iFAD8-CT, iFAD8-M, D∞-CL, QMFD-CL and D∞-TLI. The flow direction algorithms (D8, iFAD8, D∞ and QMFD) are selected due to their applicability for flow distance measurements. Although some other flow direction algorithms are shown to be more effective in other applications (e.g., Pilesjö & Hasan, 2014; Wu et al., 2022), they are not suitable to this study because they provide flow path out from a non-point source.

The mean absolute relative error (*MARE*) is used to assess the deviations between the estimated and the exact flow distances, which is defined as follows:

$$RE_i = \frac{PV_i - EV_i}{EV_i} \tag{8}$$

$$MARE = \frac{1}{n}\sum_{i=1}^{n}|RE_i| \tag{9}$$

where PV_i and EV_i are the estimated and the exact values of the *i*th cell, respectively. RE_i denotes the relative error of the *i*th cell, and n denotes the number of cells considered for the assessment.

When width function is estimated for a terrain, the exact widths of some estimated equidistant belts may be zero due to the possible overestimation of the flow distance. This phenomenon can limit the direct application of *MARE* to assess the deviations between the estimated and the exact width functions. Hence, two valid assessment criteria are adopted here. Firstly, a part of the equidistant belts with flow distances not exceeding the maximum exact flow distance are selected to calculate the *MARE* following Equation 8 and 9. Here PV_i and EV_i in Equation 8 are the *i*th estimated and the *i*th exact widths, respectively. Then the exceeding index (*EI*) is adopted to represent the ratio of the widths whose estimated flow distances are longer than the exact maximum flow distance.

$$EI = \frac{\sum_{j=1}^{m} W_j}{T} \tag{10}$$

where m denotes the number of the equidistant belts exceeding the maximum exact flow distance, and W_j denotes the *j*th exceeding equidistant belt width. T denotes the total width of all exact equidistant belts. It is obvious that an accurate width function should possess both low *MARE* and *EI* values.

3. Results

3.1. Assessments of the Flow Distance Algorithms

3.1.1. Performances over the Numerical Terrains

To show the difference between the exact and the estimated flow distance distribution clearly, the partial enlarged details over the 20 m-resolution terrains are shown in Figure 4. Meanwhile, the flow distance distributions over the whole terrains are provided in Figure S2. According to Figure 4b and 4d, D∞-TLI can reproduce the exact flow distance correctly for the inverse ellipsoid and the saddle. The isolines of flow distance by D∞-TLI are smooth and parallel to the exact isolines over the partial ellipsoid, while the deviation of D∞-TLI is as low as iFAD8-CT or iFAD8-M (Figure 4a). However, D∞-TLI underestimates the flow distance where the exact isolines facing two directions intersect over the plane (see in Figure 4c). For other algorithms, the results by the D8-based algorithms (including D8-CL, D8-DT and D8-CT) are unreasonable. D∞-CL and QMFD-CL overestimate flow distance everywhere. Although iFAD8-CL and iFAD8-DT lead to large deviations over all the numerical terrains, the other two iFAD8-based algorithms (iFAD8-CT and iFAD8-M) can reproduce the exact flow distance distribution more reasonably than other algorithms except D∞-TLI. But the accuracy of iFAD8-CT or iFAD8-M is unsteady as shown by the undulant isolines, and it is obviously lower than the accuracy of D∞-TLI over the inverse ellipsoid and the saddle. Hence, D∞-TLI is shown to be the best choice to reproduce flow distance distributions over the numerical terrains based on the visual assessments, which is also proven by the results over 5 m-resolution terrains (Figure S3).

According to the quantitative assessments (Figure 5), lower MARE appears when a finer resolution is used for most cases. D∞-TLI is shown to outperform other algorithms because it obtains the least *MARE*s with all resolutions over the inverse ellipsoid, the plane or the saddle, while only iFAD8-CT has a similar great performance with D∞-TLI over the partial ellipsoid. Overall, the average *MARE* of D∞-TLI is only 2.31 % over four numerical terrains. iFAD8-CT and iFAD8-M are the two algorithms only second to D∞-TLI, and have obvious improvements to iFAD8-CL and iFAD8-DT while iFAD8-DT outperforms iFAD8-CL. D8-DT and D8-CT outperform D8-CL over the partial ellipsoid and the inverse ellipsoid, but underperform D8-CL over the saddle. The D8-based algorithms have similar performances over the plane. Consistent with the results in Figure 4, D∞-CL and QMFD-CL obtain great errors over all the terrains, and D∞-CL seems to be more accurate than QMFD-CL.

3.1.2. Real-world Applications

The distribution of flow distance to the mapped channels is calculated in both the SCT Basin and the DDG Basin, and the visual results by different algorithms are shown in Figure 6. Here some short channels in the SCT Basin are ignored and only four channels are adopted. Enlarged details in Figure 6c and 6d show that the MFD-based algorithms including D∞-TLI, D∞-CL and QMFD-CL provide smoother isolines of flow distance than the selected seven SFD-based algorithms. Compared with other algorithms, D∞-CL and QMFD-CL always overestimate the flow distance. The strategy DT, CT and M can shorten the results of CL no matter which SFD algorithm (D8 or iFAD8) is selected, which is shown clearly by the isoline of 100 m over the SCT Basin (Figure 6c) and the isoline of 600 m over the DDG Basin (Figure 6d).

Two cases are considered for the quantitative assessments, i.e., the flow distances from the selected points to the main channels shown in Figure 3e and 3f, as well as the flow distances from all the selected points to the basin outlet. The wavy *MARE*s show that the ability of every selected algorithm is unsteady over the real-world terrains with different resolutions (Figure 7). However, compared with other algorithms, D∞-TLI can always obtain acceptable *MARE*s. For any case in Figure 7, the average *MARE* of all the resolutions is listed in Table 2. D∞-TLI is always one of the three best choices for any case. In addition, the average *MARE* of all the cases and resolutions are calculated, and the value of D∞-TLI (4.01 %) is the lowest, while the values of D8-CT (4.30 %) and iFAD8-CT (4.29 %) are lower than other algorithms except D∞-TLI. The average *MARE*s of QMFD-CL are too high over the real-world terrains. The performances of D8-CL, D8-DT, iFAD8-CL, iFAD8-M and D∞-CL are similar according to the results in Table 2.

3.2. Assessments of the Width Function

According to the results in Section 3.1, only five flow distance algorithms are selected to provide the flow distance distribution for the width function algorithm assessments, including D∞-TLI, classical D8-CL, and three algorithms with acceptable performances over the numerical or the real-world terrains (i.e., D8-CT, iFAD8-CT and iFAD8-M). Figure 8 shows the width functions estimated by two width function algorithms (i.e., the conventional SEB and the proposed MEB) with different estimated flow distance distributions over three 20 m-resolution numerical terrains. Here the flow distance interval of the equidistant belt for the width function is equal to the DEM resolution (i.e., 20 m) following some existing studies (e.g., Liu et al., 2012; Moussa, 2008; Sahoo & Sahoo, 2019b).

As shown in Figure 8, while the exact width functions for the numerical terrains are smooth, SEB causes artificial fluctuations for the estimated width functions in most cases. The MEB algorithm can decrease these unreasonable artificial fluctuations successfully. The quantitative assessment results in Figure 9 also show that MEB improves the accuracy of the estimated width function with the lower *MARE* than SEB for any selected flow distance distribution. Meanwhile, *EI* is always equal to zero over the ellipsoid or the plane, and slightly larger than zero over the inverse ellipsoid when the flow distance distribution by D8-CL or iFAD8-M is adopted. The estimated width function combining MEB and D∞-TLI is highly consistent with the exact width function (Figure 8m-8o), and generally obtains the lowest *MARE* over all the numerical terrains (Figure 9). The average *MARE* is 2.97 % for this combination but higher than 5% for other combinations. Other estimated flow distance distributions except the distribution by D∞-TLI can restore the trend of the exact width function over the ellipsoid with SEB or MEB (Figure 8a, 8d, 8g and 8j). However, the deviation of D8-CL is great over the inverse ellipsoid (Figure 8b) and the plane (Figure 8c), while the deviation of D8-CT is great over the plane (Figure 8f). These unreasonable deviations can also be identified from Figure 9.

The application over the real-world terrain also shows that MEB can overcome the artificial fluctuations (Figure 10). To obtain the exact width function over a real-world terrain is a great challenge. However, if the flow distance to channel is close to zero, the estimated width can be assumed to be double the channel length because the equidistant belt has a small interval and is close to both the channel banks. This is not a very disciplined assessment method, but can provide a reference for the application over the real-world terrain. Here the SCT Basin with 1 m-resolution DEM is selected for real-world applications and the width function to channel (including four channels in Figure 6a) with an equidistant belt interval of 1 m is calculated. No matter which algorithm is adopted to determine the flow distance distribution, the area of the first equidistant belt with flow distance ranging from 0 m to 1 m should be much smaller than the exact area as shown in Figure 8. This is because the exact first equidistant belt is covered by both the hillslope and the channel cells while the proposed algorithm only estimates width function using the hillslope cells. Hence, the next belt ranging from 1 m to 2 m (i.e., the width at 1.5 m flow distance) is used to predict the total channel length (7546 m), and the results are shown in Table 3.

The relative errors of all the predicted lengths by SEB exceed 10 %, while the relative errors by MEB are lower than 10 %. Combined with MEB, the relative errors of D∞-TLI and D8-CL are lower than 3 %. Meanwhile, Table 3 shows the widths at three neighboring flow distances (i.e., 1.5 m, 2.5 m, 3.5 m) by different combinations.

Although slight fluctuation in width function is normal, the fluctuation of the selected widths by any SEB-based combination is too strong for such a small flow distance interval. This unreasonable fluctuation is obviously artificial which is caused by SEB.

4. Discussions

4.1. Different Algorithms on Flow Distance Estimation Accuracy

The flow direction algorithm selected seems to be the major influence factor to the accuracy according to the results in Section 3.1. All the D8-based algorithms (i.e., D8-CL, D8-DT and D8-CT) provides abnormal flow distance distributions over the partial ellipsoid, the plane and the saddle, but more effective distributions over the inverse ellipsoid (Figure 4). This is because D8 can provide false flow paths directing the flow to incorrect targets over the divergent and the plain terrains, but has an acceptable performance over the convergent terrain (Wu et al., 2022). Meanwhile, D8-CL and D8-CT can provide reasonable flow distances to outlet for the selected points over two selected real-world terrains (Table 2), because most portion of the flow path from a selected point to the outlet is in the convergent valley where the D8 algorithm works effectively.

iFAD8 can provide reasonable zigzag flow paths out from the cell center (Wu et al., 2020), so the results by any iFAD8-based algorithm (i.e., iFAD8-CL, iFAD8-DT, iFAD8-CT and iFAD8-M) can approximately reflect the features of the exact flow distance distributions. D∞ and QMFD provide dispersive flow paths and may drain a part of the flow into the channel at some unusually distant locations, so D∞-CL and QMFD-CL always overestimate the flow distance (Figure 4 and 6). D∞ is less dispersive than QMFD (Orlandini et al., 2012), so D∞-CL obtains better results than QMFD-CL over all the terrains. Generally, the SFD-based algorithms get better application results than the existing MFD-based algorithms (i.e., D∞-CL and QMFD-CL).

D∞-TLI outperforms any other selected algorithm by estimating the generally most accurate flow distance over both the numerical and the real-world terrains. Although the traditional D∞ method is treated as a MFD method, D∞-TLI neglects the dispersive global flow path of D∞ and only employs D∞ for the local drainage direction. Hence, D∞-TLI does not suffer the serious problem of dispersive flow path by the MFD methods and can provide more accurate flow distance.

There are some strategies to improve the precision of the local drainage direction by D∞ (e.g., Hooshyar et al., 2016; Wu et al., 2020). These new methods may obtain the potential to further improve the accuracy of the estimated flow distance. But when the new infinite direction (ND∞) method proposed by Wu et al. (2020) is adopted to replace the D∞ direction in D∞-TLI, no obvious improvement appears to the accuracy over the real-world terrains (Figure S4). That is because the limited improvement of ND∞ to D∞ can be offset by other errors, such as the errors in TLI or the DEM generation. Hence, the applicability of the improved strategies to D∞ requires more assessments in further studies, and the tradition D∞ direction is recommended in this study due to its simplicity and popularization.

4.2. Width Function Estimation

It is possible that some cells cover multiple equidistant belts when their cell centers are located on the same equidistant belt. Then SEB adds all these cell areas into one equidistant belt while other neighboring equidistant belts receive no cell area, which draws the artificial fluctuations in Figure 8. The originality of MEB is to attempt to divide a cell into correct equidistant belts, so it is unsurprising to find that MEB leads to slighter artificial fluctuation. However, the accuracy of the estimated width function depends not only on the width function algorithm (i.e., SEB or MEB), but also on the selected flow distance algorithm. With a specific flow distance distribution, MEB can estimated the width function more accurately than SEB (Figure 9). For a specific width function algorithm (SEB or MEB), its accuracy can be greatly improved by using a more accurate flow distance algorithm. However, the combination of SEB with a more accurate flow distance algorithm may outperform the combination of MEB with a flow distance algorithm of which the accuracy is not that high. As an instance, iFAD8-CT causes higher *MARE* than iFAD8-M when providing flow distance distribution for the plane (Figure 5c), but iFAD8-CT-MEB obtains a better width function than iFAD8-M-SEB (Figure 9c). While D∞-TLI is shown to be the best choice for flow distance estimation, the combination of D∞-TLI with MEB is optimal, which is demonstrated by the steady great performance over the numerical and the real-world terrains (Figure 9 and Table 3).

The results in Section 3.2 are based on a traditional precondition that the equidistant belt interval for the width function is equal to the DEM resolution. This is a small interval, so the region not belonging to the correct equidistant belt may occupy a large proportion of the whole cell region. This precondition may increase the artificial fluctuation. If the interval is set to be larger, more area in a cell can belong to the same equidistant belt as the cell center, then the SEB method will suffer a slighter artificial fluctuation. However, MEB can still optimize the accuracy of SEB to a degree in this case as shown in Figure S5.

4.3. Computational Efficiency

All the flow distance algorithms used in this study are implemented following the two-segmented-distance strategy. The runtimes for the plane with different resolutions show that this strategy can guarantee acceptable computational efficiency (Table 4). D∞-TLI can process a DEM with more than 9×10^6 cells in less than 20.0 s. This runtime is similar to D∞-CL, and is longer than the D8-based algorithms as well as QMFD-CL, while is shorter than the iFAD8-based algorithms. This is similar to the difference in efficiency of the selected flow direction algorithms (D8, iFAD8, D∞ and QMFD) as shown by Wu et al. (2022). So, the selected flow direction algorithm seems to be the main factor affecting runtime. The computation efficiency of MEB is also acceptable with less than 14.0 s required to process the same 1 m-resolution DEM using any given flow distance distribution. This runtime is much longer than the SEB algorithm (0.08 s), but is shorter than the runtimes of most flow distance algorithms.

5. Concluding Remarks

A new method to estimate flow distance, as well as width function based on grid DEMs, is proposed in this study. The new flow distance algorithm (D∞-TLI) adopts a two-segmented-distance strategy that divides the flow distance into two segments whose sub-distances are approximated using D∞ and TLI, respectively. Then, the continuous flow distance field is approximated over each cell region, so this cell area can be divided into multiple equidistant belts (MEB) for the width function.

Four numerical terrains and two real-world terrains with multiple resolutions are adopted for assessments. The results show that D∞-TLI generally outperforms nine existing flow distance algorithms and causes low average *MARE*s of 2.31 % and 4.01% for the estimated flow distance distribution over the numerical and the real-world terrains, respectively. Compared with the traditional method, MEB can effectively decrease the artificial fluctuations in the estimated width function. The combination of D∞-TLI with MEB (i.e., D∞-TLI-MEB) outperforms other combinations by providing estimated width functions with an average *MARE* of 2.97 % for the numerical terrains, while it also works well over real-world terrains. Meanwhile, all the strategies used in this study (including the two-segmented-distance strategy, D∞-TLI, and MEB) show acceptable computational efficiency. Therefore, D∞-TLI and MEB have great potential to provide hydro-geomorphological attributes for hydrological models.

In further studies, more experiments can be conducted to show how much our method can improve hydrological modeling, although the numerical accuracy advantage of the flow distance and width function estimated by the newly proposed method has been provided here. D∞ and TLI are selected here due to their great accuracy in common geographical studies, but the performance of other local drainage direction methods and interpolation methods can be tested and evaluated by replacing D∞ or TLI in this method.

Acknowledgements

This work was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP; grant no. 2019QZKK0207-02), the National Natural Science Foundation of China (NSFC; grant no. 92047301), and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (grant no. KYCX22_0636).

Open Research

Data and Code Availability Statement

The DEMs of the numerical terrains are available at: https://doi.org/10.6084/m9.figshare.16909321.v1 (Wu et al., 2022). The DEM containing the SCT Basin distributed by OpenTopography is available at: https://doi.org/10.5069/G9RB72JV. The ALOS DEM containing the DDG Basin is available at: https://search.earthdata.nasa.gov. The real-world basin domains as well as the mapped channels are available at: https://doi.org/10.6084/m9.figshare.22004444.v1. The Java codes of the proposed algorithms are available at: https://doi.org/10.6084/m9.figshare.22010132.v1.

References

Bogaart, B. W., & Troch, P. A. (2006). Curvature distribution within hillslopes and catchments and its effect on the hydrological response. *Hydrology and Earth System Sciences*, *10*, 925-936. https://doi.org/10.5194/hess-10-925-2006

Butt, M. A., & Maragos, P. (1998). Optimum design of chamfer distance transforms. *IEEE Transactions on Image Processing*, *7*(10), 1477-1484.

https://doi.org/10.1109/83.718487

De Smith, M. (2004). Distance transforms as a new tool in spatial analysis, urban planning, and GIS. *Environment and Planning B: Planning and Design*, *31*(1), 85-104. https://doi.org/10.1068/b29123

Di Lazzaro, M., Zarlenga, A., & Volpi, E. (2016). Understanding the relative role of dispersion mechanisms across basin scales. *Advances in Water Resources*, *91*, 23-36. http://doi.org/10.1016/j.advwatres.2016.03.003

Dong, L., Ge, C., Zhang, H., Liu, Z., Yang, Q., Jin, B., Ritsema, C. J., & Geissen, V. (2022). An optimized method for extracting slope length in RUSLE from raster digital elevation. *Catena*, *209*, 105818. https://doi.org/10.1016/j.catena.2021.105818

Fan, Y., & Bras, R. L. (1998). Analytical solutions to hillslope subsurface storm flow and saturation overland flow. *Water Resources Research*, *34*, 921-927. https://doi.org/10.1029/97WR03516

Fan, F. M., Fleischmann, A. S., Collischonn, W., Ames, D. P., & Rigo, D. (2015). Large-scale analytical water quality model coupled with GIS for simulation of point sourced pollutant discharges. *Environmental Modelling & Software*, *64*, 58-71. http://doi.org/10.1016/j.envsoft.2014.11.012

Fei, J., Liu, J., Ke, L., Wang, W., Wu, P., & Zhou, Y. (2022). A deep learning-based method for mapping alpine intermittent rivers and ephemeral streams of the Tibetan Plateau from Sentinel-1 time series and DEMs. *Remote Sensing of Environment*, *282*, 113271. https://doi.org/10.1016/j.rse.2022.113271

Gupta, V., Waymire, E., & Rodíguez-Iturbe, I. (1986). On scales, gravity and network structure in basin runoff. In Gupta, V., Waymire, E., & Rodíguez-Iturbe, I. (Eds), *Scale Problems in Hydrology* (pp. 159-184). D. Reidel Publishing Co. http://doi.org/10.1007/978-94-009-4678-1_8

Hazenberg, P., Fang, Y., Broxton, P., Gochis, D., Niu, G. Y., Pelletier, J. D., Troch, P. A., & Zeng, X. (2015). A hybrid-3D hillslope hydrological model for use in Earth system models. *Water Resources Research*, *51*, 8218-8239. https://doi.org/10.1002/2014WR016842

Hickey, R. (2000). Slope angle and slope length solutions for GIS. *Cartography*, *29*(1), 1-8. https://doi.org/10.1080/00690805.2000.9714334

Hooshyar, M., Wang, D., Kim, S., Medeiroset, S. C., & Hagen, S. C. (2016). Valley and channel networks extraction based on local topographic curvature and k-means clustering of contours. *Water Resources Research*, *52*, 8081-8102. https://doi.org/10.1002/2015WR018479

Lapides, D., Sytsma, A., O'Neil, G., Djokic, D., Nichols, M., & Thompson, S. (2022). Arc Hydro Hillslope and Critical Duration: New tools for hillslope-scale runoff analysis. *Environmental Modelling and Software*, *153*, 105408. https://doi.org/10.1016/j.envsoft.2022.105408

Lee, D. T., & Schachter, B. J. (1980). Two algorithms for constructing a Delaunay triangulation. International *Journal of Computer & Information Sciences*, *9*(3), 219-242. https://doi.org/10.1007/BF00977785

Li, Z., Yang, T., Wang, C., Shi, P., Yong, B., & Song, Y. (2021). Assessing the precision of total contributing area (TCA) estimated by flow direction algorithms based on the analytical solution of theoretical TCA on synthetic surfaces. *Water Resources Research*, *57*, e2020WR028546. https://doi.org/10.1029/2020WR028546

Liu, J., Chen, X., Zhang, X., & Hoagland, K. D. (2012). Grid digital elevation model based algorithms for determination of hillslope width functions through flow distance transforms. *Water Resources Research*, *48*, W04532. http://doi.org/10.1029/2011WR011395

Liu, J., Han, X., Chen, X., Lin, H., & Wang, A. (2016). How well can the subsurface storage-discharge relation be interpreted and predicted using the geometric factors in headwater areas?. *Hydrological Processes*, *30*(25), 4826-4840. https://doi.org/10.1002/hyp.10958

Maxwell, J. C. (1870). On hills and dales. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *40*(269), 421-427.

Mayorga, E., Logsdon, M. G., Ballester, M., & Richey, J. E. (2005). Estimating cell-to-cell land surface drainage paths from digital channel networks, with an application to the amazon basin. *Journal of Hydrology*, 315, 167-182. http://doi.org/10.1016/j.jhydrol.2005.03.023

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., & Seibert, J. (2005). The role of topography on catchment-scale water residence time. *Water Resources Research*, *41*, W05002. http://doi.org/10.1029/2004WR003657

Moussa, R. (2008). What controls the width function shape, and can it be used for channel network comparison and regionalization?. *Water Resources Research*, *44*, W08456. http://doi.org/10.1029/2007WR006118

Muzik, I. (1996). Flood modelling with GIS-derived distributed unit hydrographs. *Hydrological Processes*, *10*, 1401-1409. https://doi.org/10.1002/(SICI)1099-1085(199610)10:10<1401::AID-HYP469>3.0.CO;2-3

Noël, P., Rousseau, A. N., Paniconi, C., & Nadeau, D. F. (2014). Algorithm for delineating and extracting hillslopes and hillslope width functions from gridded elevation data. *Journal of Hydrologic Engineering*, *19*(2), 366-374. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000783

O'Callaghan, J. F., & Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics, and Image Processing*, *28*(3), 323-344. https://doi.org/10.1016/S0734-189X(84)80011-0

Orlandini, S., Moretti, G., Corticelli, M. A., Santangelo, P. E., Capra, A., Rivola, R., & Albertson, J. D. (2012). Evaluation of flow direction methods against field observations of overland flow dispersion. Water Resources Research, 48, W10523. https://doi.org/10.1029/2012WR012067

Orlandini, S., Moretti, G., Franchini, M., Aldighieri, B., & Testa, B. (2003). Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models. *Water Resources Research*, *39*(6), 1144. https://doi.org/10.1029/2002WR001639

Orlandini, S., Moretti, G., & Gavioli, A. (2014). Analytical basis for determining slope lines in grid digital elevation models. *Water Resources Research*, *50*, 526-539. https://doi.org/10.1002/2013WR014606

Paik, K. (2008). Global search algorithm for nondispersive flow path extraction. *Journal of Geophysical Research*, *113*, F04001. https://doi.org/10.1029/2007JF000964

Paz, A. R., Collischonn, W., Risso, A., & Mendes, C. A. B. (2008). Errors in river lengths derived from raster digital elevation models. *Computers & Geosciences*, *34*, 1584-1596. https://doi.org/10.1016/j.cageo.2007.10.009

Pilesjö, P., & Hasan, A. (2014). A triangular form-based multiple flow algorithm to estimate overland flow distribution and accumulation on a digital elevation model. *Transactions in GIS*, *18*(1), 108-124. https://doi.org/10.1111/tgis.12015

Polidori, L., & Chorowicz, J. (1993). Comparison of bilinear and Brownian interpolation for digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, *48*(2), 18-23. https://doi.org/10.1016/0924-2716(93)90036-m

Quinn, P., Beven, K., Chevallier, P., & Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, *5*, 59-79. https://doi.org/10.1002/hyp.3360050106

Ranjram, M., & Craig, J. R. (2021). Use of an efficient proxy solution for the hillslope-storage Boussinesq problem in upscaling of subsurface stormflow. *Water Resources Research*, *57*, e2020WR029105.

https://doi.org/10.1029/2020WR029105

Rigon, R., Bancheri, M., Formetta, G., & de Lavenne, A. (2016). The geomorphological unit hydrograph from a historical-critical perspective. *Earth Surface Processes and Landforms*, *41*, 27-37. https://doi.org/10.1002/esp.3855

Rinaldo, A., Marani, A., & Rigon, R. (1991). Geomorphological dispersion. *Water Resources Research*, *27*(4), 513-525. https://doi.org/10.1029/90WR02501

Sahoo, S., & Sahoo, B. (2019a). A geomorphology-based integrated stream-aquifer interaction model for semi-gauged catchments. *Hydrological Processes*, *33*, 1362-1377. https://doi.org/10.1002/hyp.13406

Sahoo, S., & Sahoo, B. (2019b). Modelling the variability of hillslope drainage using grid-based hillslope width function estimation algorithm. *ISH Journal of Hydraulic Engineering*, *25*(1), 71-78. https://doi.org/10.1080/09715010.2018.1441750

Schwendel, A. C., Fuller, I. C., & Death, R. G. (2012). Assessing DEM interpolation methods for effective representation of upland stream morphology for rapid appraisal of bed stability. *River Research and Applications*, *28*(5), 567-584. https://doi.org/10.1002/rra.1475

Shin, S., & Paik, K. (2017). An improved method for single flow direction calculation in grid digital elevation models. *Hydrological Processes*, *31*(8), 1650-1661. https://doi.org/10.1002/hyp.11135

Sloan, S. W. (1987). A fast algorithm for constructing Delaunay triangulations in the plane. *Advances in Engineering Software*, *9*(1), 34-55.

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, *33*(2), 662-319. https://doi.org/10.1029/96WR03137

Tesfa, T. K., Tarboton, D. G., Chandler, D. G., & McNamara, J. P. (2009), Modeling soil depth from topographic and land cover attributes. *Water Resources Research*, *45*, W10438. https://doi.org/10.1029/2008WR007474

Troch, P. A., Paniconi, C., & van Loon, E. E. (2003). Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. formulation and characteristic response. *Water Resources Research*, *39*(11), 1316. https://doi.org/10.1029/2002WR001728

Troch, P., van Loon, E., & Hilberts, A. (2002). Analytical solutions to a hillslope-storage kinematic wave equation for subsurface flow. *Advances in Water Resources*, *25*, 637-649. https://doi.org/10.1016/S0309-1708(02)00017-9

Van Nieuwenhuizen, N., Lindsay, J. B., & DeVries, B. (2021). Smoothing of digital elevation models and the alteration of overland flow path length distributions. *Hydrological Processes*, *35*(7), e14271. https://doi.org/10.1002/hyp.14271

Veneziano, D., Moglen, G. E., Furcolo, P., & Iacobelli, V. (2000). Stochastic model of the width function. *Water Resources Research*, *36*(7), 1143-1157. https://doi.org/10.1029/2000WR900002

Wilson, J. P., Lam, C. S., & Deng, Y. (2007). Comparison of the performance of flow-routing algorithms used in GIS-based hydrologic analysis. *Hydrological Processes*, *21*, 1026-1044. https://doi.org/10.1002/hyp.6277

Wu, P., Liu, J., Han, X., Feng, M., Fei, J., & Shen, X. (2022). An improved triangular form-based multiple flow direction algorithm for determining the nonuniform flow domain over grid networks. *Water Resources Research*, *58*, e2021WR031706. https://doi.org/10.1029/2021WR031706

Wu, P., Liu, J., Han, X., Liang, Z., Liu, Y., & Fei, J. (2020). Nondispersive drainage direction simulation based on flexible triangular facets. *Water Resources Research*, *56*, e2019WR026507. https://doi.org/10.1029/2019WR026507

Yilmaz, H. M. (2007). The effect of interpolation methods in surface definition: an experimental study. *Earth Surface Processes and Landforms*, *32*, 1346-1361. https://doi.org/10.1002/esp.1473

Xu, H., Ma, C., Lian, J., Xu, K., & Evance, C. (2018). Urban flooding risk assessment based on an integrated k-means cluster algorithm and improved entropy weight method in the region of Haikou, China. *Journal of Hydrology*, 563, 975-986. https://doi.org/10.1016/j.jhydrol.2018.06.060

Zhou, Q., Pilesjö, P., & Chen, Y. (2011). Estimating surface flow paths on a digital elevation model using a triangular facet network. *Water Resources Research*, *47*, W07522. https://doi.org/10.1029/2010WR009961

Tables

Table 1. The flow distance algorithms adopted for the comparison in this study.

Algorithm	Origin of the cumulative distance algorithm	Adopted flow direction algorithm	Origin of the flow direction algorithm
D8-CL	-	D8	O'Callaghan and Mark (1984)
D8-DT	Paz et al. (2008)	D8	O'Callaghan and Mark (1984)
D8-CT	Liu et al. (2012)	D8 / D∞	O'Callaghan and Mark (1984) / Tarboton (1997)
iFAD8-CL	-	iFAD8	Wu et al. (2020)
iFAD8-DT	Paz et al. (2008)	iFAD8	Wu et al. (2020)
iFAD8-CT	Liu et al. (2012)	iFAD8 / D∞	Wu et al. (2020) / Tarboton (1997)
iFAD8-M	Dong et al. (2022)	iFAD8	Wu et al. (2020)
QMFD-CL	Bogaart and Troch (2006)	QMFD	Quinn et al. (1991)
D∞-CL	Bogaart and Troch (2006)	D∞	Tarboton (1997)
D∞-TLI	Current study	D∞	Tarboton (1997)

Table 2. The average *MARE* of the estimated flow distance by different algorithms.

	D8-CL	D8-DT	D8-CT	iFAD8-CL	iFAD8-DT	iFAD8-CT	iFAD8-M	QMFD-CL	D∞-CL	D∞-TLI
Distance to channel for the SCT Basin	8.93	8.44	8.24	8.84	**7.84**	**7.62**	7.92	21.26	9.04	**7.22**
Distance to outlet for the SCT Basin	**4.58**	4.81	**4.52**	4.94	4.93	4.78	5.99	6.71	4.65	**4.52**
Distance to channel for the DDG Basin	2.61	4.32	2.79	**2.48**	3.93	2.94	4.65	14.31	**2.52**	**2.57**
Distance to outlet for the DDG Basin	**1.09**	3.09	**0.96**	1.17	2.99	1.20	4.03	11.59	1.12	**0.93**
All the cases above	4.49	5.29	**4.30**	4.57	5.04	**4.29**	5.72	13.47	4.50	**4.01**

Note. The unit is in 10^{-2}, and the three lowest average *MARE*s of every case are bolded.

Table 3. Relative errors between the exact length and the predicted river lengths using the equidistant belt areas by different flow distance algorithms and width function extraction modes.

	D8-CL-SEB	D8-CL-MEB	D8-CT-SEB	D8-CT-MEB	iFAD8-CT-SEB	iFAD8-CT-MEB	iFAD8-M-SEB	iFAD8-M-MEB	D∞-TLI-SEB	D∞-TLI-MEB
Width at distance of 1.5 m (m)	17683	15425	18514	15779	18343	16129	17633	15682	18222	15471
Width at distance of 2.5 m (m)	19731	16118	18123	16546	17994	16745	21262	18300	17528	16277
Width at distance of 3.5 m (m)	17288	16091	16198	16521	16792	16770	17065	16840	17538	16278
Predicted river length (m)	8842	7712	9257	7889	9172	8064	8817	7841	9111	7736
Relative error (%)	17.17	2.21	22.67	4.55	21.54	6.87	16.84	3.91	20.74	2.51

Table 4. The runtimes of different flow distance algorithms to process the plane with different resolutions.

Resolution (m)	Cell Numbers	D8-CL	D8-DT	D8-CT	iFAD8-CL	iFAD8-DT	iFAD8-CT	iFAD8-M	QMFD-CL	D∞-CL	D∞-TLI
30	1.02×10^{4}	0.003	0.004	0.007	0.012	0.015	0.017	0.01	0.005	0.008	0.007
20	2.28×10^{4}	0.01	0.008	0.017	0.027	0.025	0.037	0.02	0.009	0.022	0.018
10	9.06×10^{4}	0.043	0.042	0.064	0.113	0.108	0.137	0.093	0.046	0.068	0.067
5	3.61×10^{5}	0.198	0.227	0.324	0.513	0.551	0.632	0.459	0.269	0.33	0.318
2	2.25×10^{6}	2.097	2.155	2.877	4.379	4.583	5.181	5.554	2.378	2.775	2.895
1	9.01×10^{6}	11.562	10.306	16.936	21.513	24.577	27.01	59.79	13.971	17.333	18.401

Note. The unit is in second. The evaluation is performed on a computer with an Intel Xeon E5-1620 v4 CPU and 80 GB of memory.

Figures

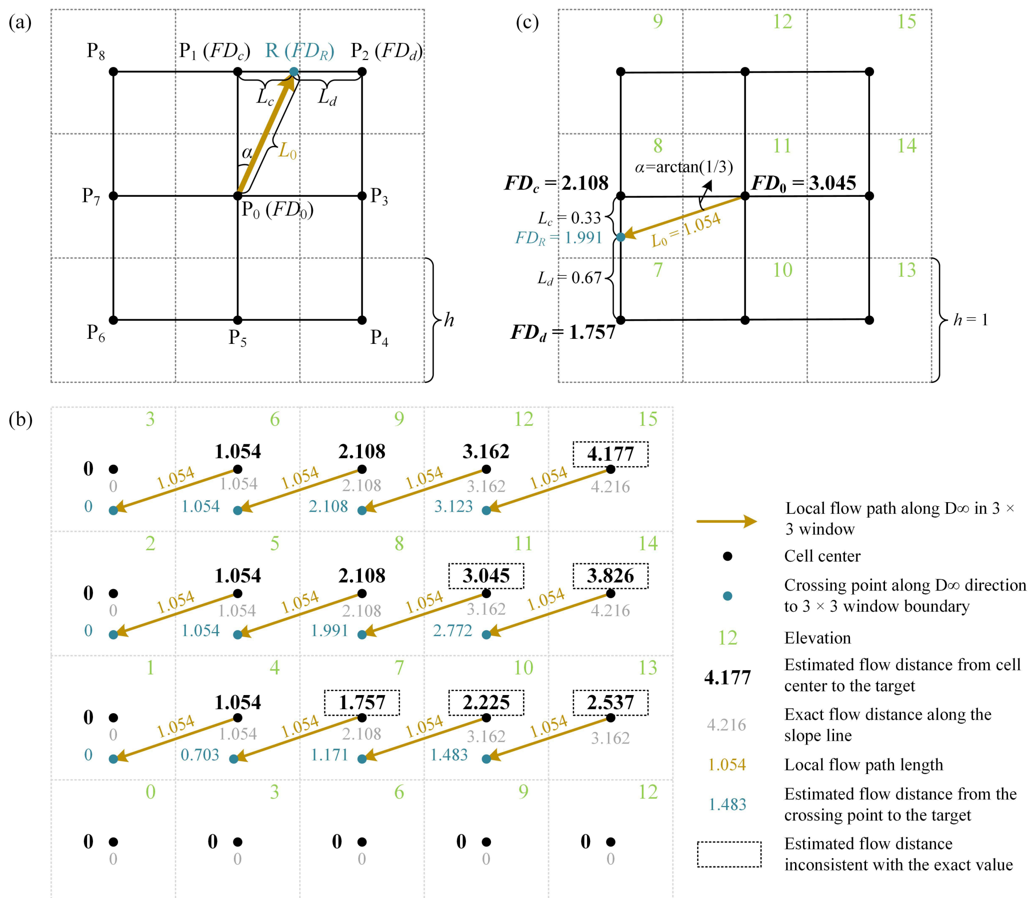


Figure 1. The theory of the new flow distance algorithm. (a) Flow distance (FD_0) from a cell center (P_0) to a downstream target is computed as the sum of the local flow path length (L_0) along D∞ direction in a 3 × 3 window and the estimated flow distance from the crossing point (R) on the window boundary to the same target. (b) A 1 m-resolution DEM of a plane is adopted as an example. Then the estimated distribution of flow distance to DEM side for the cell centers as well as the data generated in the computational process are shown. The exact flow distances which is calculated as the lengths of theoretical slope lines are also shown. Meanwhile, (c) shows the process of the flow distance value assignment for the cell center whose

elevation is equal to 11.

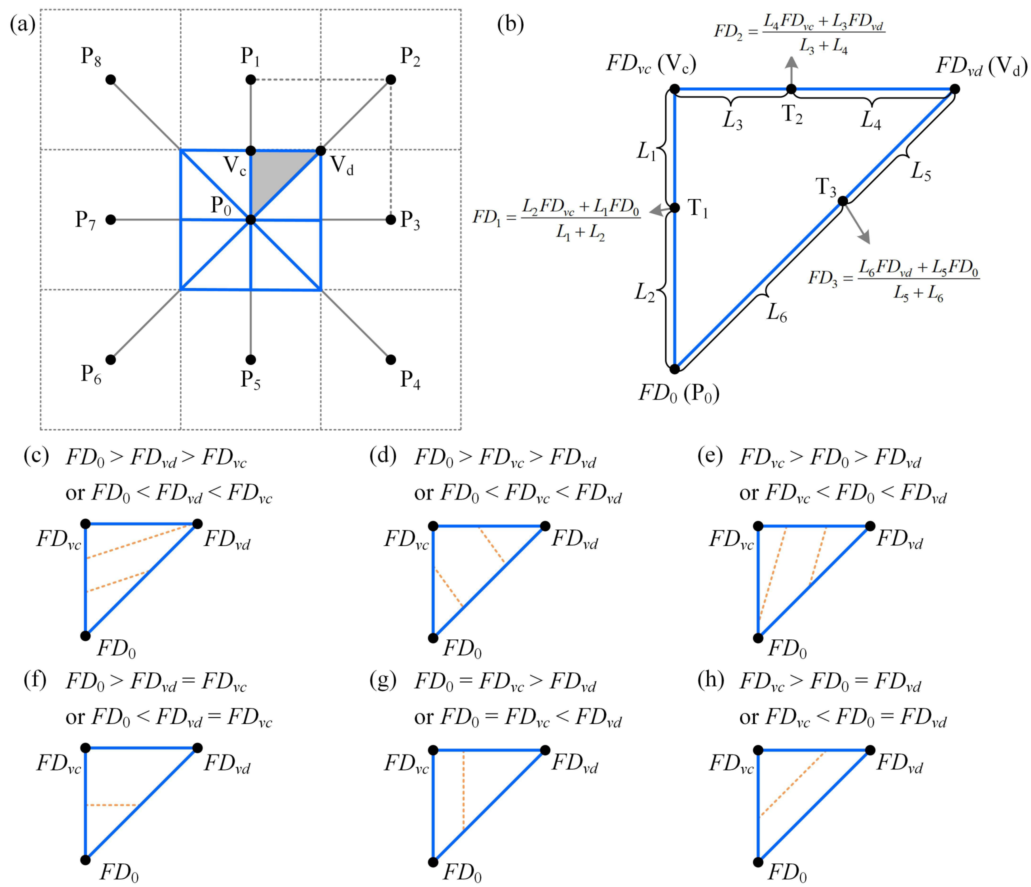


Figure 2. The method to determine the local equidistant belt area which is used to constitute the width function. (a) The cell is divided into eight facets as shown with blue boundary, and the flow distance of any point at the facet boundary can be calculated following (b), where L_1, …, L_6 denote the lengths from the points to the vertexes. (c-h) Each equidistant line is straight and linking two points at the boundary with the same flow distance, and the area between two equidistant lines is added to the corresponding equidistant belt.

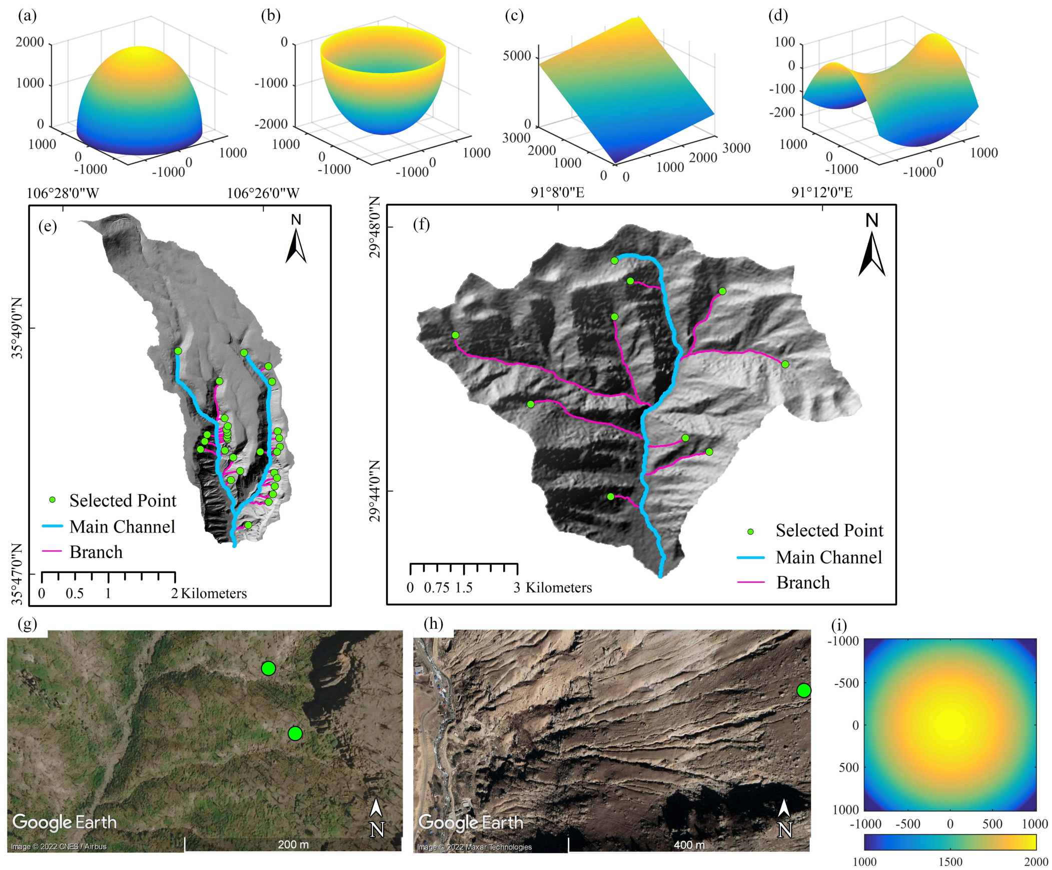


Figure 3. Four numerical and two real-world terrains are used for the assessments, including (a) an ellipsoid, (b) an inverse ellipsoid, (c) a plane, (d) a saddle, (e) the SCT Basin, and (f) the DDG Basin. Local images of (g) the SCT Basin and (h) the Duodigou Basin with several selected source points mapped are used to show the branch channels or gullies. Moreover, the elevation distribution of the square partial region of the ellipsoid (-1020 m < x, y < 1020 m) used for the flow distance assessments is shown in (i).

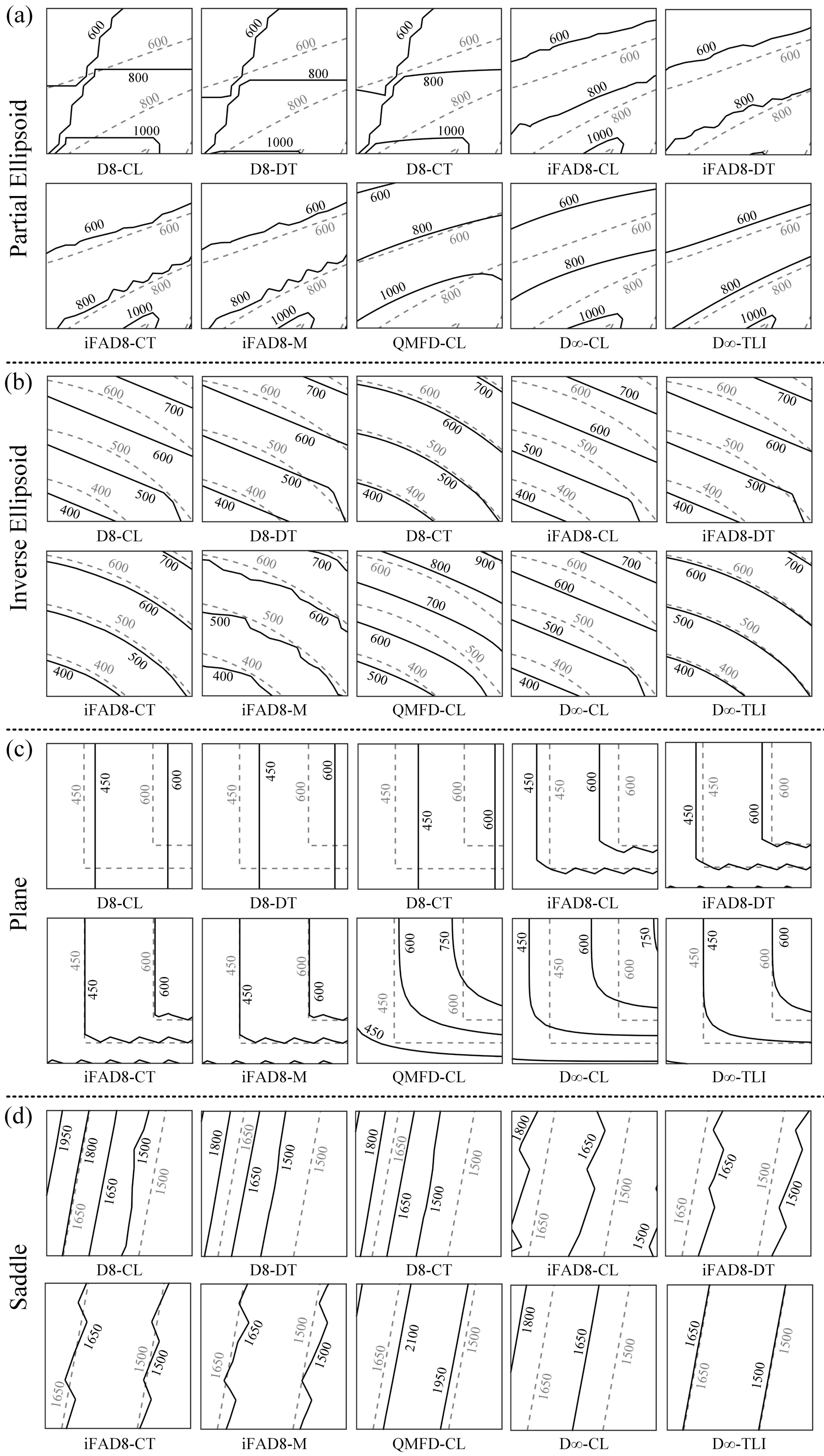


Figure 4. Enlarged windows of the exact flow distance distributions (grey dashed

lines) versus estimated flow distance distributions (black solid lines) to the terrain boundary by different algorithms over four numerical terrains with 20 m resolution. The window ranges of the partial ellipsoid or the inverse ellipsoid are 100 m $< x <$ 400 m and 300 m $< y <$ 600 m. The window ranges of the plane are 350 m $< x <$ 650 m and 100 m $< y <$ 400 m. The window ranges of the partial ellipsoid are 150 m $< x <$ 300 m and 1250 m $< y <$ 1400 m.

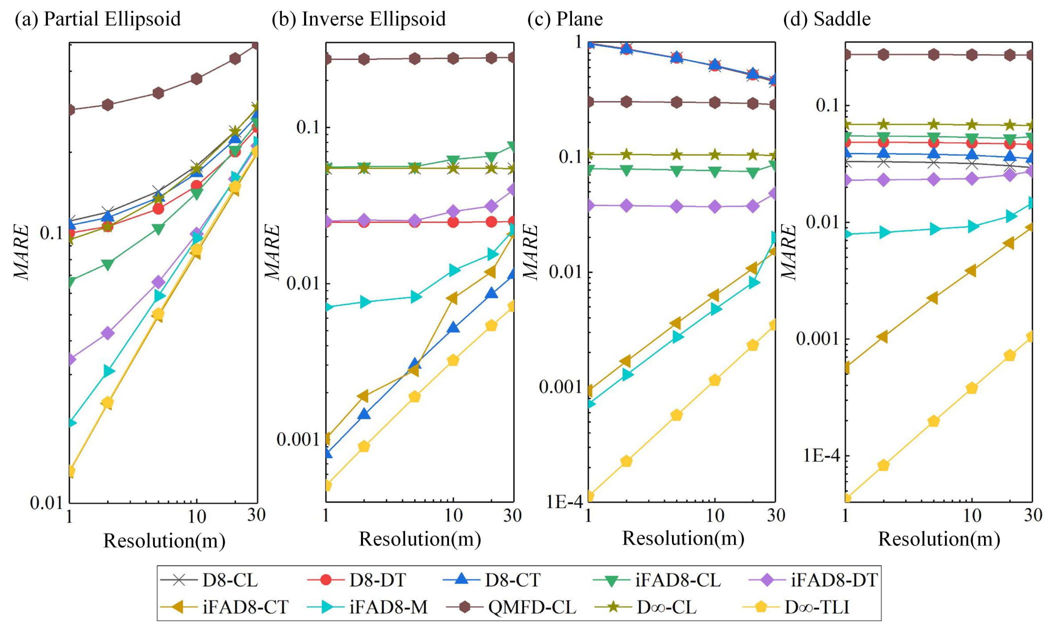


Figure 5. The mean absolute relative error (*MARE*) of the estimated flow distance by different algorithms over (a) the partial ellipsoid, (b) the inverse ellipsoid, (c) the plane, and (d) the saddle with six different resolutions. Here the lines of D∞-TLI and iFAD8-CT are almost coincident for (a) the partial ellipsoid.

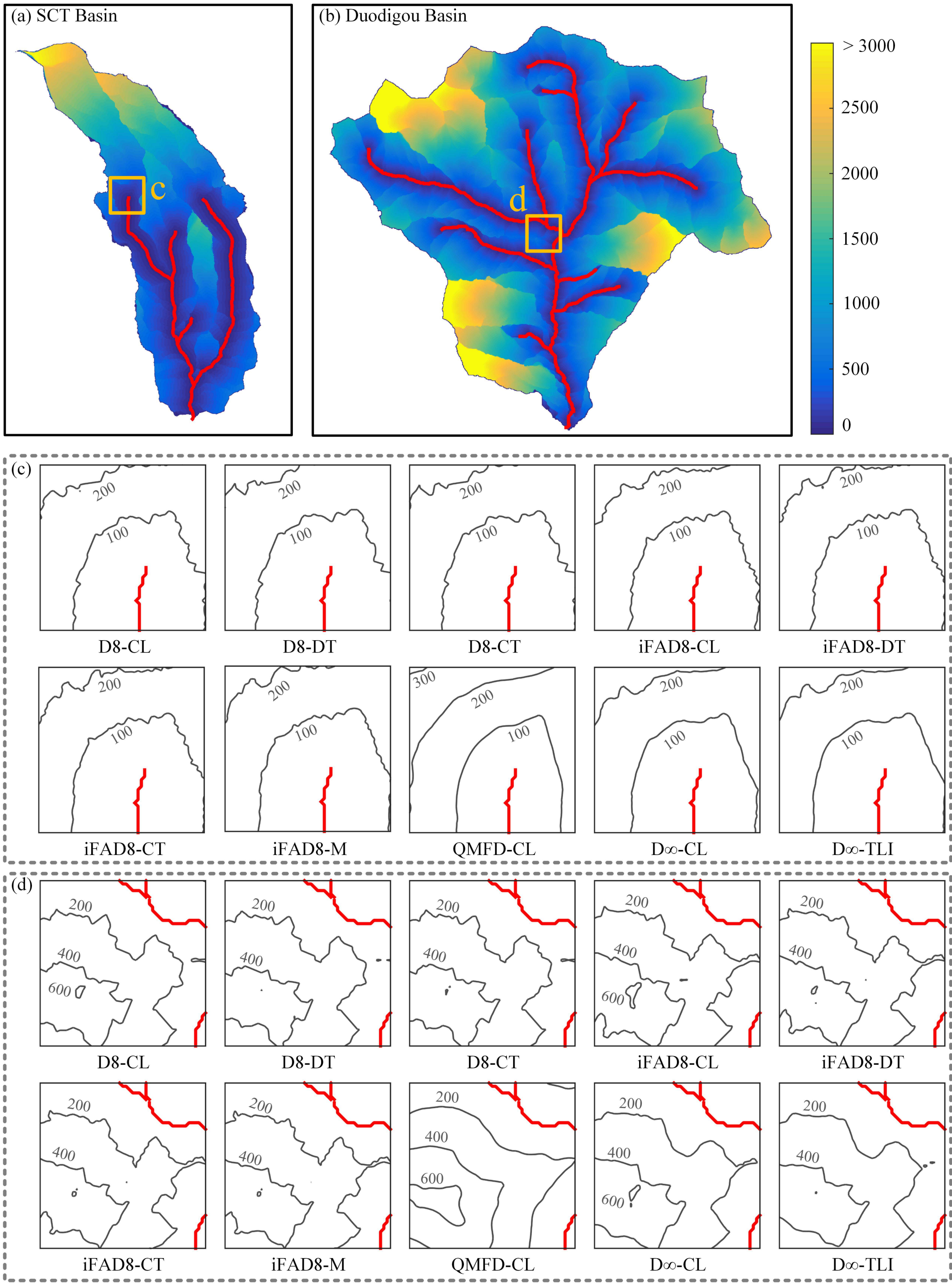


Figure 6. The estimated flow distance distribution (grey lines) to channels (red lines) by D∞-TLI over (a) the SCT Basin and (b) the DDG Basin. For the marked domains in (a) and (b), enlarged windows in (c) and (d) are used to show the difference between the flow distance distributions estimated by ten selected algorithms.

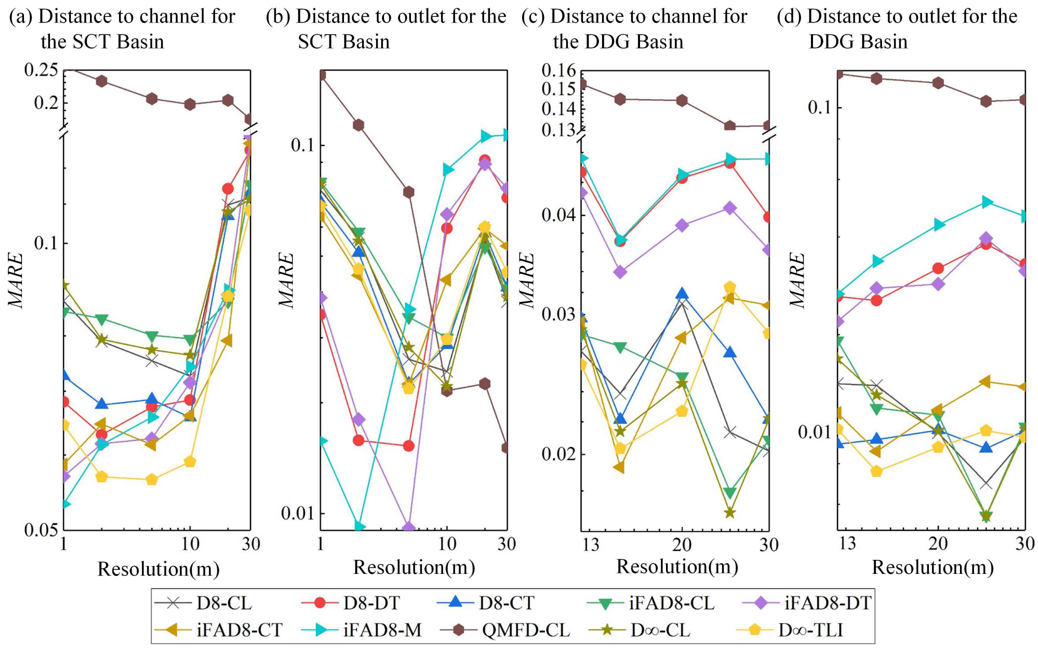


Figure 7. The mean absolute relative error (*MARE*) of the estimated flow distance to channel or outlet for the selected points in the SCT Basin and the Duodigou Basin.

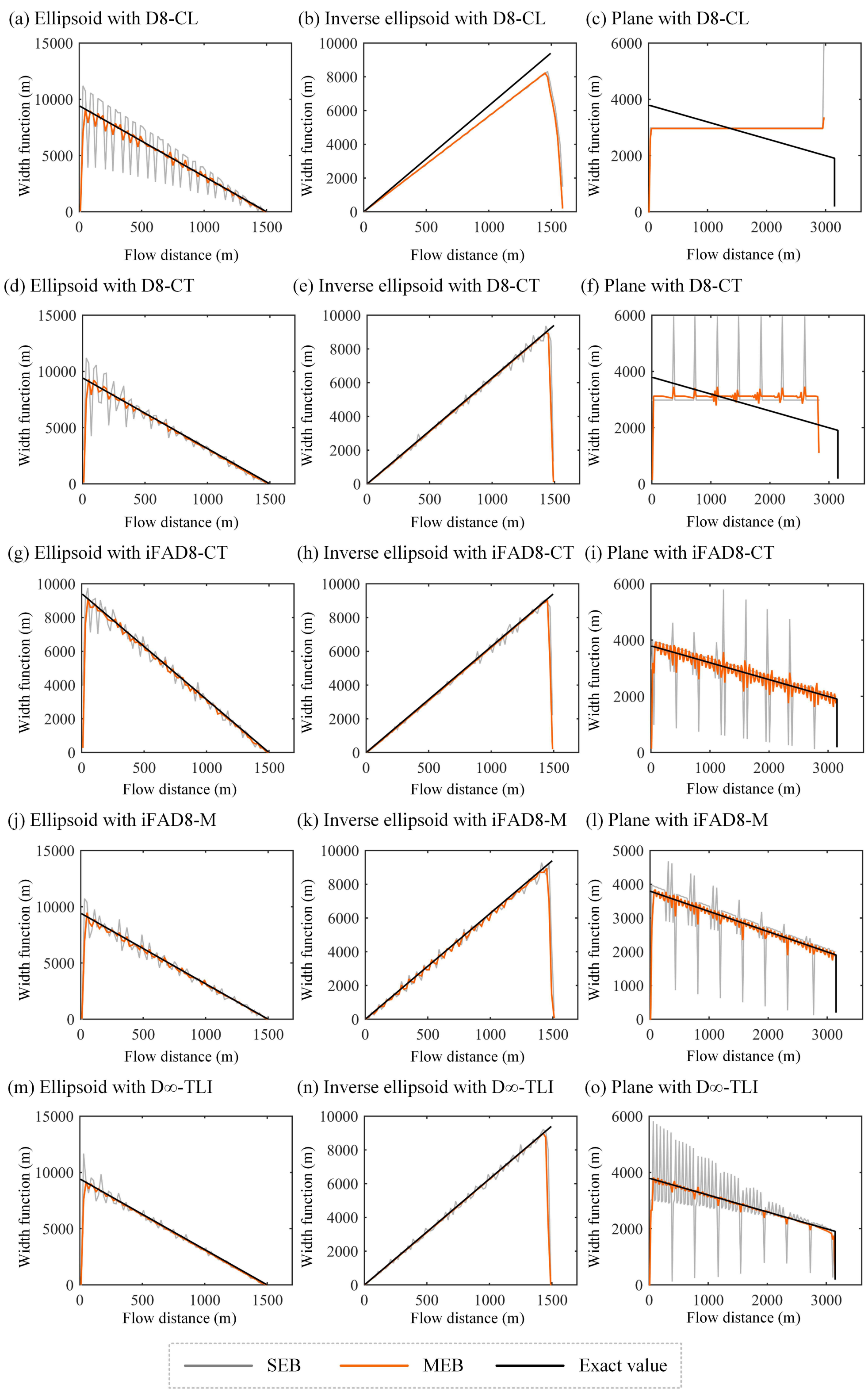


Figure 8. The width functions derived by the SEB and the MEB method with five selected flow distance algorithms and 20 m-resolution numerical terrains. The flow distance interval is equal to the resolution.

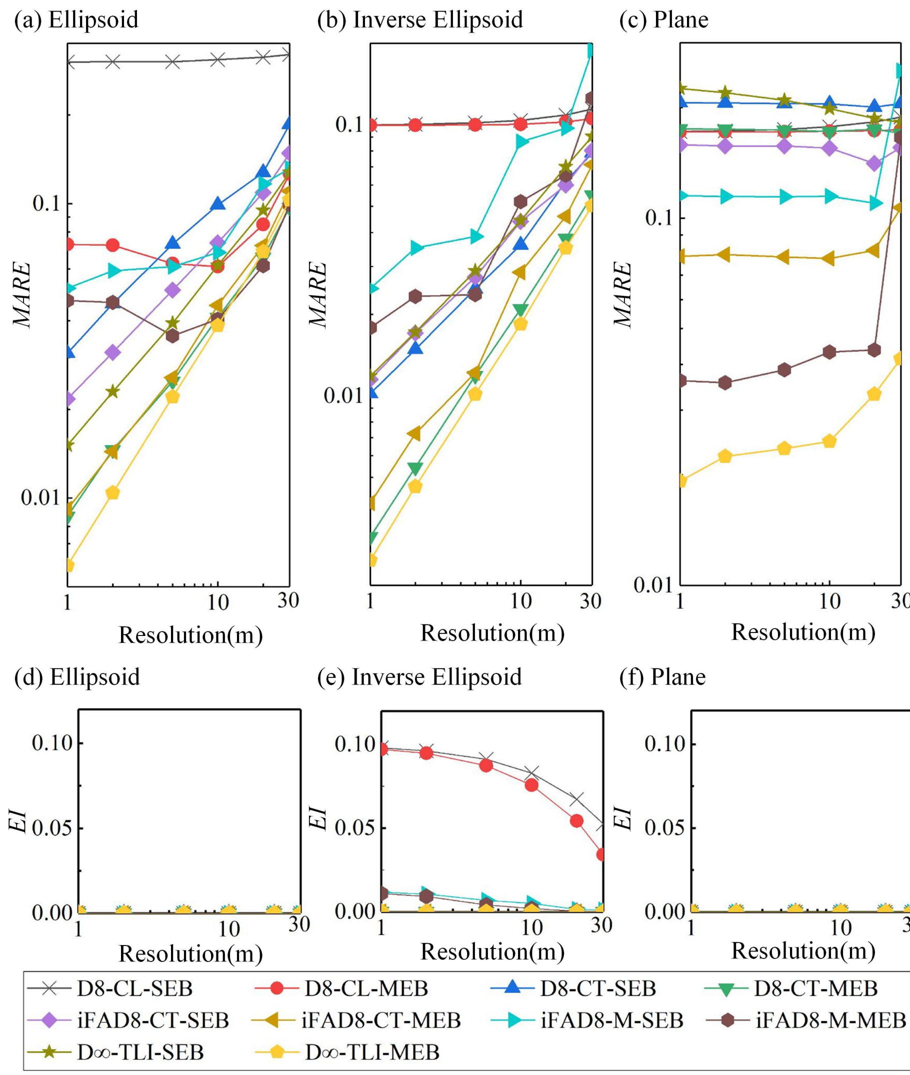


Figure 9. The mean absolute relative error (*MARE*) and the exceeding index (*EI*) of the estimated width function over three numerical terrains with different resolutions. The flow distance interval is equal to the specific resolution.

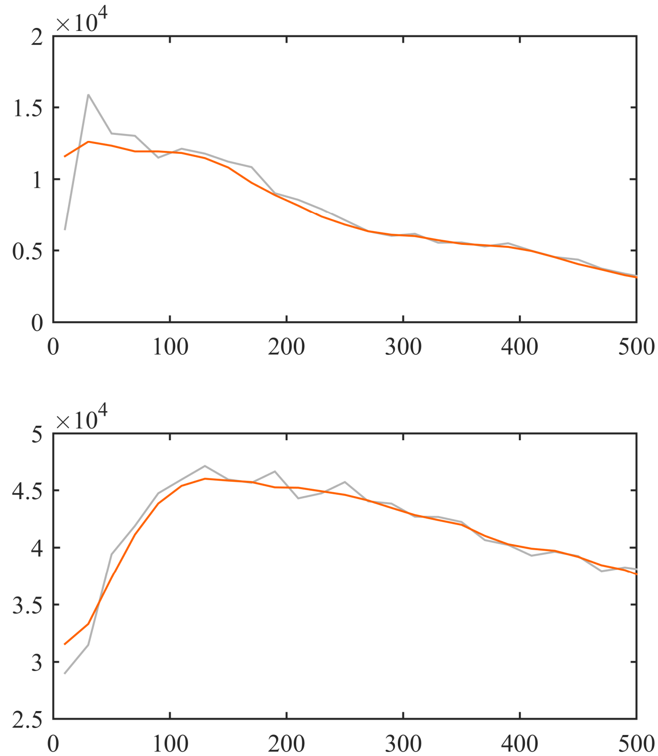


Figure 10. The partial width functions with flow distance shorter than 500 m by the SEB and the MEB method. The distributions of flow distance to channel are estimated by D∞-TLI over (a) the SCT Basin and (b) the DDG Basin with the resolution of 20 m. The flow distance interval of the width function is equal to 20 m.

Figure 1.

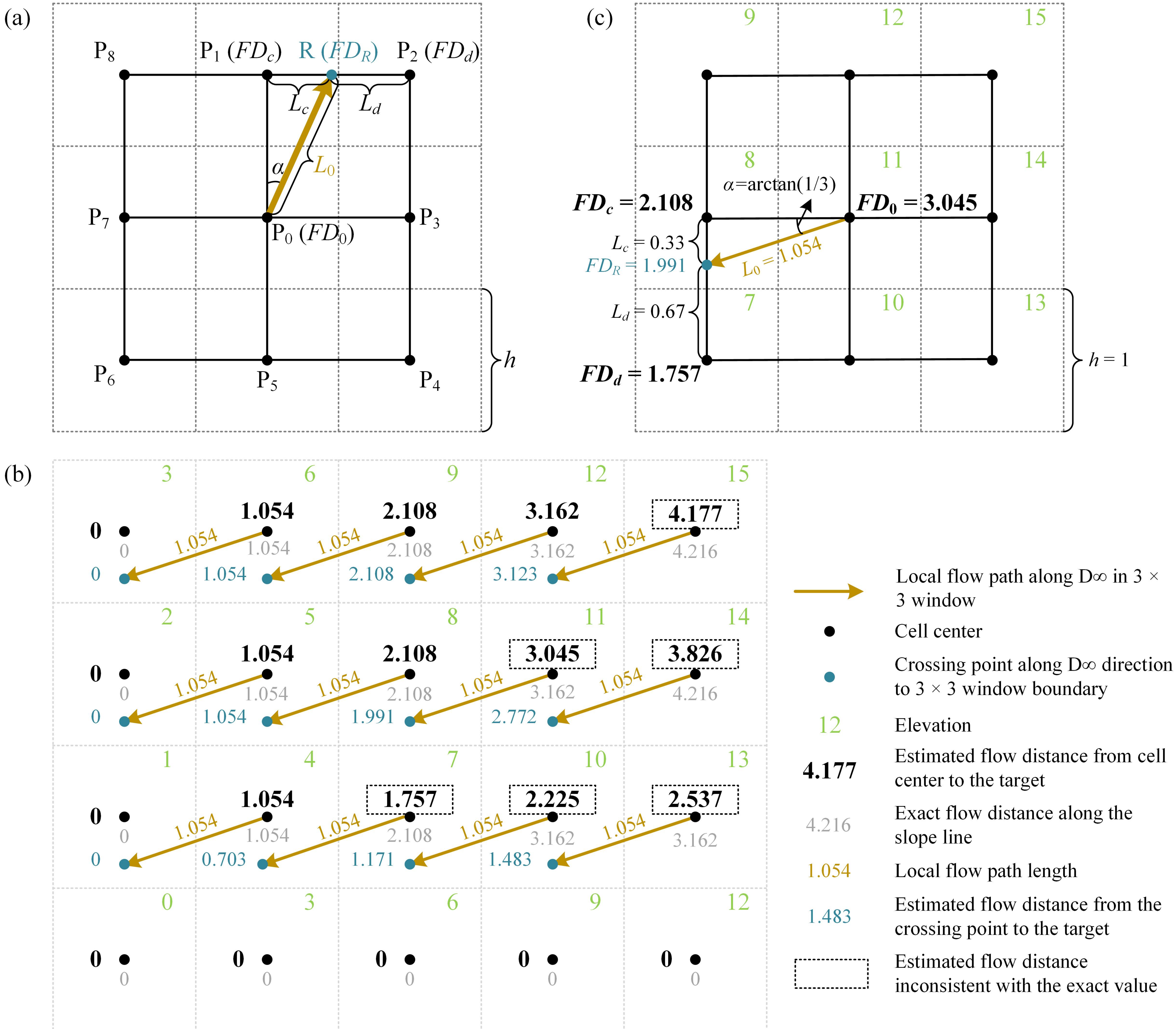

(a)
P8
P1 (FDc)
R (FDR)
P2 (FDd)
Lc
Ld
α
L0
P7
P0 (FD0)
P3
P6
P5
P4
h
(c)
FDc = 2.108
α=arctan(1/3)
FD0 = 3.045
Lc = 0.33
FDR = 1.991
L0 = 1.054
Ld = 0.67
FDd = 1.757
h = 1
(b)
Local flow path along D∞ in 3 × 3 window
Cell center
Crossing point along D∞ direction to 3 × 3 window boundary
Elevation
Estimated flow distance from cell center to the target
Exact flow distance along the slope line
Local flow path length
Estimated flow distance from the crossing point to the target
Estimated flow distance inconsistent with the exact value

Figure 2.

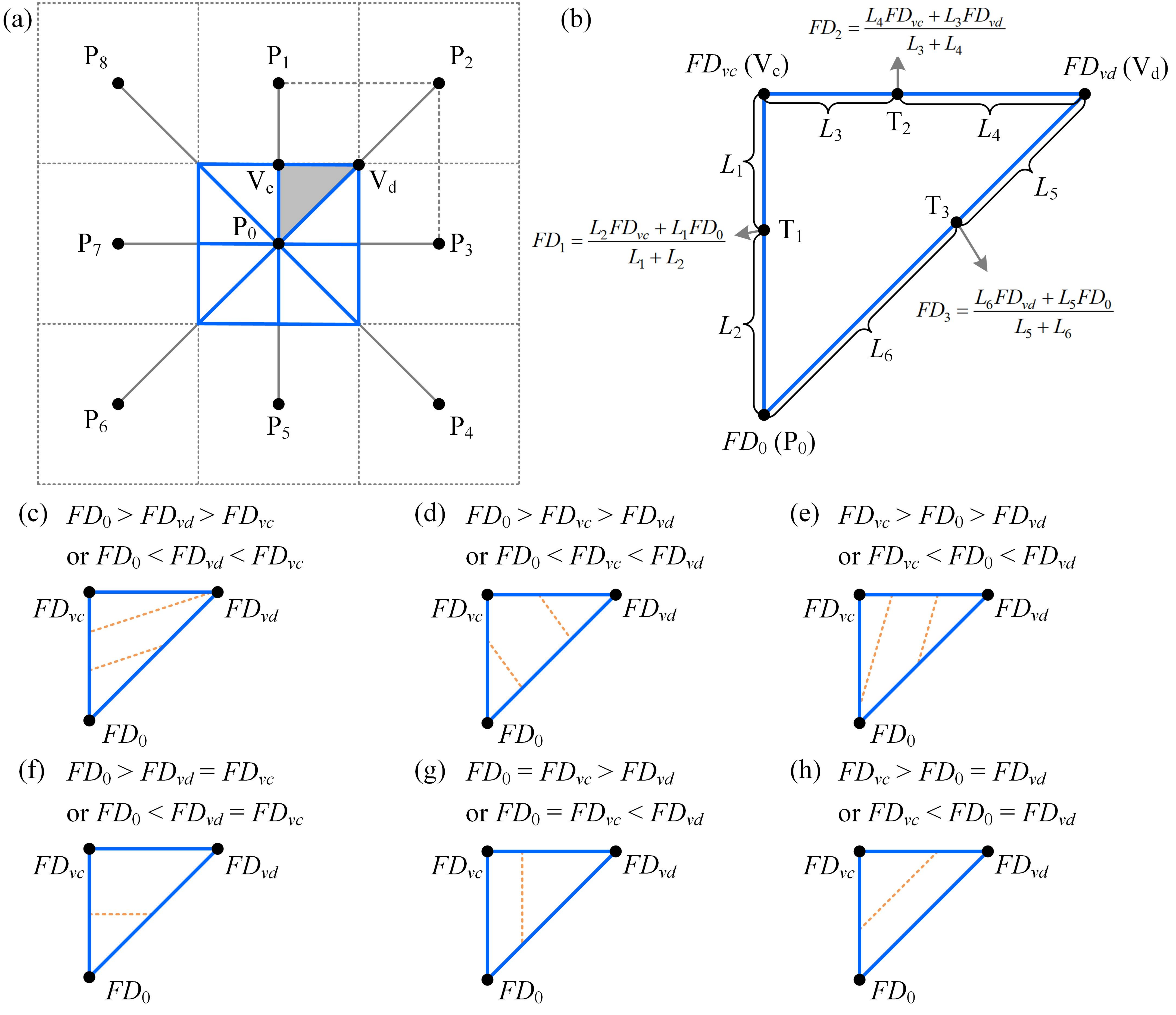

(a)
P8
P1
P2
Vc
Vd
P0
P7
P3
P6
P5
P4
(b)
$FD_2 = \frac{L_4 FD_{vc} + L_3 FD_{vd}}{L_3 + L_4}$
FD_{vc} (Vc)
FD_{vd} (Vd)
L_3
T2
L_4
L_1
L_5
T3
$FD_1 = \frac{L_2 FD_{vc} + L_1 FD_0}{L_1 + L_2}$
T1
$FD_3 = \frac{L_6 FD_{vd} + L_5 FD_0}{L_5 + L_6}$
L_2
L_6
FD_0 (P0)
(c) $FD_0 > FD_{vd} > FD_{vc}$
or $FD_0 < FD_{vd} < FD_{vc}$
(d) $FD_0 > FD_{vc} > FD_{vd}$
or $FD_0 < FD_{vc} < FD_{vd}$
(e) $FD_{vc} > FD_0 > FD_{vd}$
or $FD_{vc} < FD_0 < FD_{vd}$
(f) $FD_0 > FD_{vd} = FD_{vc}$
or $FD_0 < FD_{vd} = FD_{vc}$
(g) $FD_0 = FD_{vc} > FD_{vd}$
or $FD_0 = FD_{vc} < FD_{vd}$
(h) $FD_{vc} > FD_0 = FD_{vd}$
or $FD_{vc} < FD_0 = FD_{vd}$
FD_{vc}
FD_{vd}
FD_0

Figure 3.

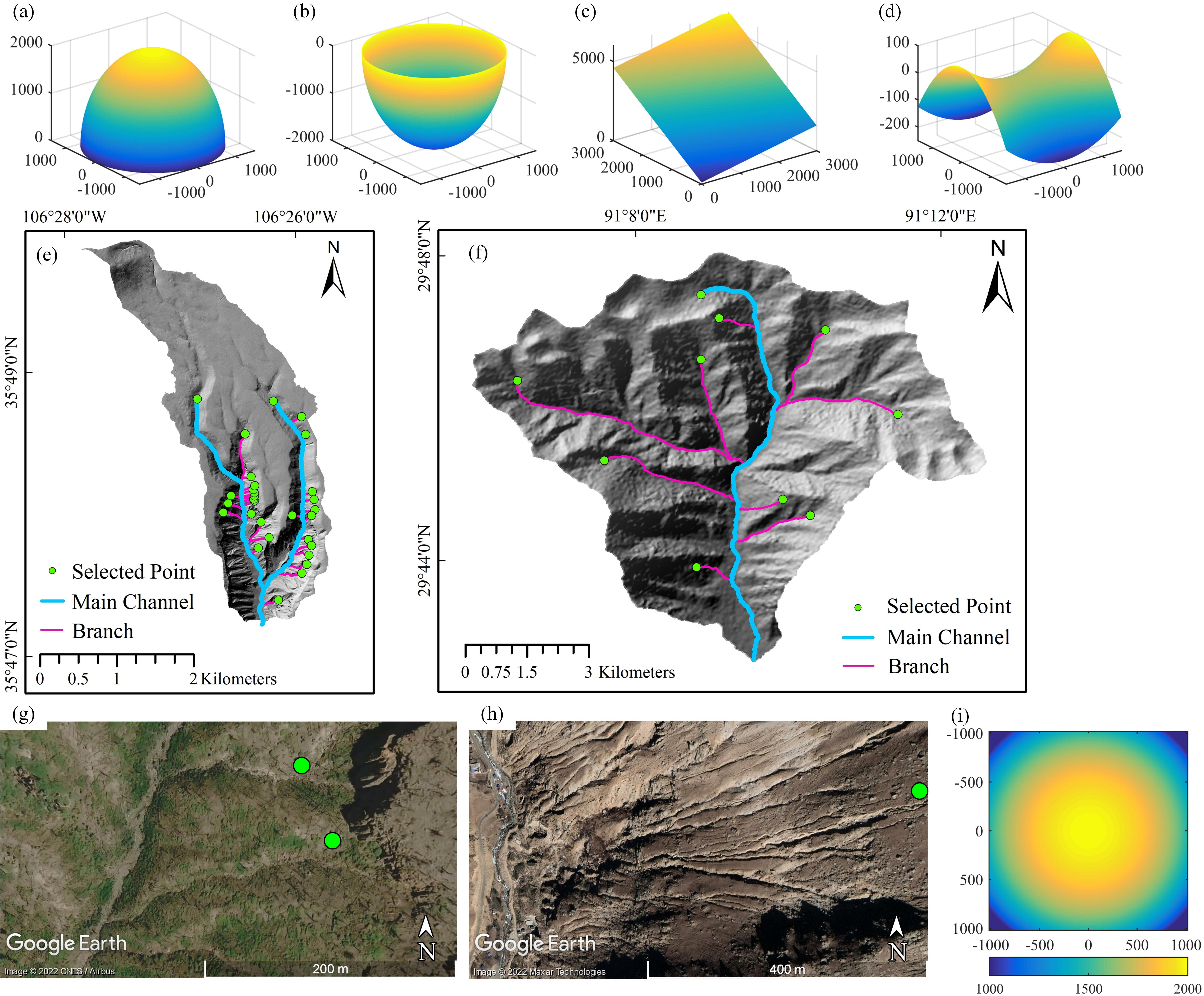

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
106°28'0"W
106°26'0"W
35°49'0"N
35°47'0"N
91°8'0"E
91°12'0"E
29°48'0"N
29°44'0"N
N
Selected Point
Main Channel
Branch
0 0.5 1 2 Kilometers
0 0.75 1.5 3 Kilometers
Google Earth
Image © 2022 CNES / Airbus
200 m
Image © 2022 Maxar Technologies
400 m

Figure 4.


(a)
Partial Ellipsoid
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(b)
Inverse Ellipsoid
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(c)
Plane
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI
(d)
Saddle
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI

Figure 5.

(a) Partial Ellipsoid
(b) Inverse Ellipsoid
(c) Plane
(d) Saddle
MARE
Resolution(m)
0.01
0.1
0.001
1E-4
1
1
10
30
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI

Figure 6.

(a) SCT Basin
c
(b) Duodigou Basin
d
> 3000
2500
2000
1500
1000
500
0
(c)
200
100
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
300
QMFD-CL
D∞-CL
D∞-TLI
(d)
200
400
600
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI

Figure 7.

(a) Distance to channel for the SCT Basin
(b) Distance to outlet for the SCT Basin
(c) Distance to channel for the DDG Basin
(d) Distance to outlet for the DDG Basin
MARE
Resolution(m)
0.25
0.2
0.1
0.05
0.01
0.16
0.15
0.14
0.13
0.04
0.03
0.02
1
10
30
13
20
D8-CL
D8-DT
D8-CT
iFAD8-CL
iFAD8-DT
iFAD8-CT
iFAD8-M
QMFD-CL
D∞-CL
D∞-TLI

Figure 8.

(a) Ellipsoid with D8-CL
(b) Inverse ellipsoid with D8-CL
(c) Plane with D8-CL
(d) Ellipsoid with D8-CT
(e) Inverse ellipsoid with D8-CT
(f) Plane with D8-CT
(g) Ellipsoid with iFAD8-CT
(h) Inverse ellipsoid with iFAD8-CT
(i) Plane with iFAD8-CT
(j) Ellipsoid with iFAD8-M
(k) Inverse ellipsoid with iFAD8-M
(l) Plane with iFAD8-M
(m) Ellipsoid with D∞-TLI
(n) Inverse ellipsoid with D∞-TLI
(o) Plane with D∞-TLI
Width function (m)
Flow distance (m)
SEB
MEB
Exact value

Figure 9.

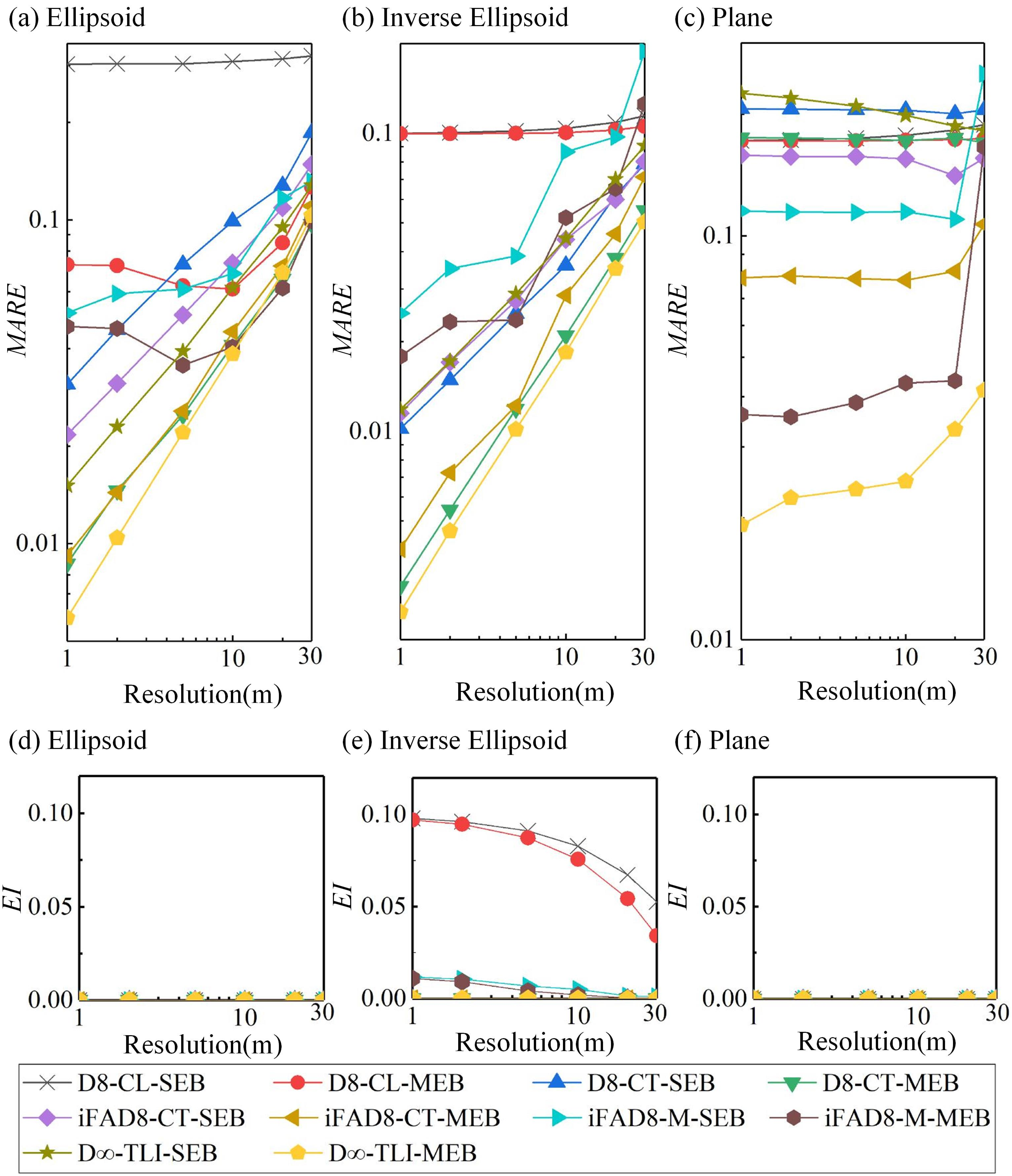
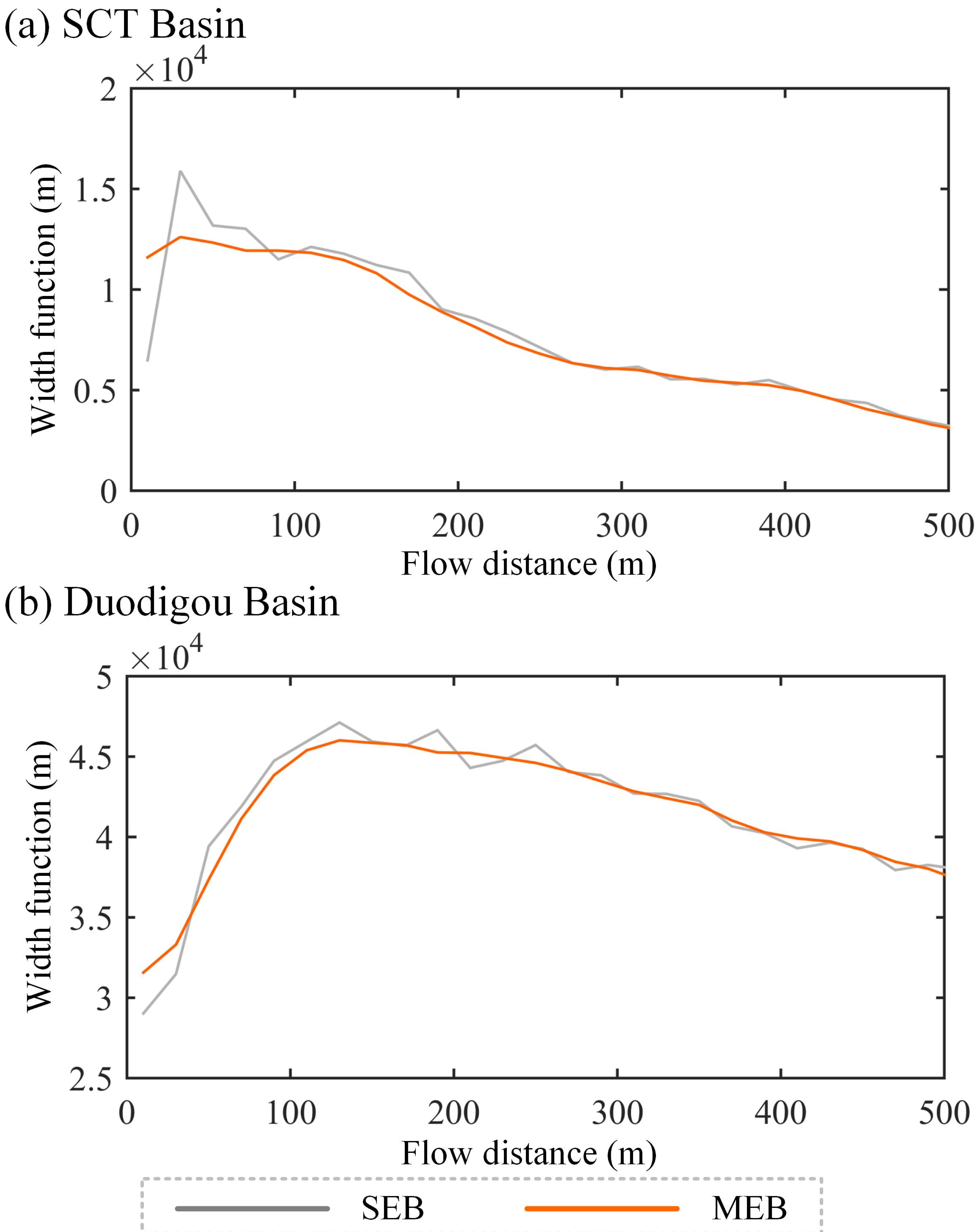

(a) Ellipsoid
(b) Inverse Ellipsoid
(c) Plane
MARE
0.1
0.01
1
10
30
Resolution(m)
(d) Ellipsoid
(e) Inverse Ellipsoid
(f) Plane
EI
0.10
0.05
0.00
D8-CL-SEB
D8-CL-MEB
D8-CT-SEB
D8-CT-MEB
iFAD8-CT-SEB
iFAD8-CT-MEB
iFAD8-M-SEB
iFAD8-M-MEB
D∞-TLI-SEB
D∞-TLI-MEB

Figure 10.



(a) SCT Basin
×10^4
Width function (m)
Flow distance (m)
(b) Duodigou Basin
×10^4
Width function (m)
Flow distance (m)
SEB
MEB