

以 DEM 提取流域水系河源的最小误差分析

陈冬平, 陈莹, 陈兴伟*

(福建师范大学 地理科学学院, 福州 350007)

摘要: 目前, 以水文模型提取流域水系已成为水文科学研究中的热点问题。但用 DEM 提取流域水系的方法不能解决流域水系唯一性的问题。故本文在没有流域纸质水系图或遥感数据的前提下, 提出以 DEM 提取流域水系的河源最小误差法, 即通过建立实际河源与提取水系河源两点之间的距离误差与网格数大小的相互关系, 以两点之间误差最小为原则解决流域水系提取唯一性问题。以晋江流域为例, 30m 分辨率的数字高程模型 (DEM) 为基础数据, 采用河源最小误差法提取了晋江流域水系。结果表明, 最小河流长度为 42m 时, 实际河源与提取水系河源的距离误差达到最小, 相应的晋江流域的水系分维值为 1.389, 表明该法是提取流域水系较合理可行的算法。

关键词: 流域水系; DEM; 河源最小误差法; 网格数; 晋江流域

DOI: 10.3724/SP.J.1047.2011.00240

1 引言

以 DEM 提取流域水系是近年来水文科学研究中的热点^[1]。传统流域水系的提取, 通常是基于纸质的地形图或水系图通过数字化方式提取。何钢^[2]等基于 1:450 万中华人民共和国地形图和 1:1 200 万“陆地卫星影像中国地学分析图集”, 采用 MapInfo 提取水系, 王大鹏^[3]等以 1:5 万的河流数据提取太湖流域的河网水系。这种方法提取的流域河网水系的精度主要受限于已有的地形图或水系图。

随着数字技术的发展, 数字高程模型 (DEM) 展现其在水文学研究中的独特优势。Hutchinson^[4]首次提出利用河道矢量数据来丰富 DEM 提取的河网水系信息。Saunders^[5]等通过人工输入的河道数字地图来修正提取的流域水系。Turcotte^[6]等通过数字河流与湖泊网络 DRLN (digital river and lake network) 方法提高湖泊与平原的处理能力。Xu Xinliang^[7]等用 DEM 自动提取流域水系的方法。在国内, 以 DEM 提取河流水系主要有两种方法, 一是参照纸质水系图或河道矢量图层, 将 DEM 提取的流域水系与河道数字化图层相叠加提取流域水

系, 如王加虎^[8]等引入矢量河网和栅格河网提取海河支流——滦河流域的水系。许悍卫^[9]等以遥感信息提取秦淮河流域的河网水系, 郑子彦^[10]等附加主要的河道数字化图层提取临沂市的河网水系。该法提取的精度也取决于纸质水系图或河道矢量图层的精度; 二是在没有流域水系图层的条件下, 孔凡哲^[11]等利用集水面积阈值与河网密度或河源密度的相互关系提取沿渡河流域水系, 熊立华^[12]等通过临界点支撑面积与河网平均坡度和河网总长度的相互关系提取加拿大 Paddle 流域的河网水系, 林峰^[13]等和王林^[14]等提出根据集水面积阈值与分维值的相互关系, 分别确定九龙江流域水系和晋江流域水系。然而, 文献[11-14]都是通过“拐点”假定来提取流域水系, “拐点”假定存在无标度区间的选择, 区间不同, “拐点”不同, 因此, 并没有解决流域水系唯一性的问题。自然界一条河流的水系应当是唯一的, 不论基于何种资料, 通过何种方法提取的水系最为接近于真实情况才能算合理。而对提取水系的合理性, 在前人的研究中并没有提出一个确切的概念, 也没有形成统一的认识。

鉴于上述 3 种提取水系方法存在的不足, 本文提出以 DEM 提取的水系河源与实际河源距离误差

收稿日期: 2010-06-02; 修回日期: 2010-07-03.

基金项目: 国家自然科学基金面上项目 (50979015)。

作者简介: 陈冬平 (1986-), 男, 江西宜春人, 硕士。主要研究方向为水资源与水环境。E-mail: chendongping092@126.com

* 通讯作者: 陈兴伟, 男, 福建省福鼎人, 教授, 博导。研究方向为流域水沙污染物过程及其数值模拟。

E-mail: cxwchen215@163.com

最小的方法来确定水系。以晋江流域为例,探讨采用河源最小误差法确定流域河网水系,并在此基础上,利用网格法计算晋江流域水系分维值并分析比较,初步验证河源最小误差法的合理性。

2 河源最小误差法提取晋江流域水系

晋江是福建省中南部河流,位于闽江与九龙江之间,是福建省的第三大河流,因纪念晋朝永嘉南渡故名晋江。晋江河源位于泉州市安溪县境内桃舟乡达新村梯仔岭东南坡^[15],坐标 $117^{\circ}41'42''\text{E}$, $25^{\circ}21'30''\text{N}$,其上游有东溪、西溪两支流,西溪为正源。两溪于南安的双溪口汇合,合流为一,即古称“南安江”。流经永春,安溪,南安,鲤城,丰泽,晋江等市县,于石狮注入台湾海峡。晋江流域所处的地理范围在 $117^{\circ}44'\sim 118^{\circ}47'\text{E}$, $24^{\circ}31'\sim 25^{\circ}32'\text{N}$ 之间(如图 1 所示),河流全长 302km,总面积 $5\,269\text{km}^2$ 。

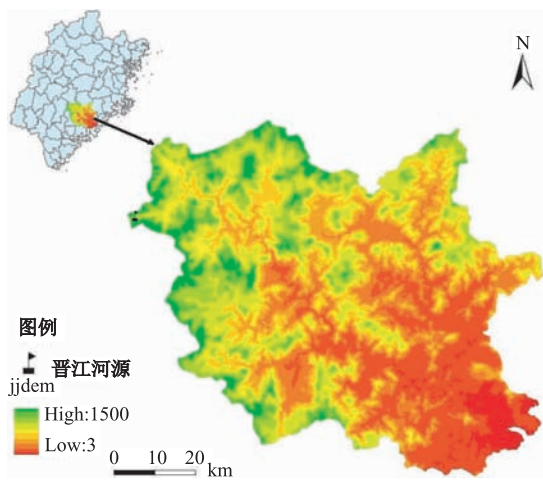


图 1 晋江流域位置及高程分布

Fig. 1 The location and elevation distribution of Jinjiang basin

2.1 河源最小误差法的分析

河源的确定对于河流长度的计算、比降、流量、集水面积的计算确定等都有很大影响,因此,一个精准的河源确定分析对于提取流域水系至关重要。河流河源的确定一般根据“河源唯远”原则^[16],文献^[15]正是根据此原则来确定晋江流域的河源。

基于 DEM 提取流域河网水系,每个网格数或集水面积阈值下可提取相应的水系,且每个水系都有其对应的水系端点或称河源;阈值不同,提取的水系及其河源的位置也不相同。因此提取的河源

与实际河源两点间就存在位置(距离)误差。理论上讲,精确的水系提取,将使两者重合,误差为零。但在实际工作中,两者很难重合,总是存在一定的误差。因此,本文提出以 DEM 提取流域水系的河源最小误差法,即选用实际河源与提取水系河源两点之间的距离误差作为流域水系提取的标准,以两点之间误差最小为原则来解决流域水系唯一性的问题,从而确定流域河网水系。该方法以距离最小的唯一性解决河流水系的唯一性,理论上是可行的。随着网格数的变化,提取水系河源与真实河源的距离误差不断的变化,且存在一定的规律,此规律可通过数学函数来表达,而求取函数的导数可以获得函数的最小值,由此可见在方法上也具有可行性。

2.2 DEM 数据处理与拟合

采用 30m 分辨率的 DEM,对高程范围 3~1 500m 的晋江流域,利用 ArcGIS 9.2 的 Hydro 模块,通过以下几个步骤提取流域水系:(1)DEM 填洼,通过 fill 命令完成,由此可以得到一个与原 DEM 相比“无洼地”新的 DEM;(2)流向的确定,采用传统 D8 方法通过 Flow Direction 命令完成,形成一张新的栅格模型;(3)汇流累积量分析,采用 Flow Accumulation 命令完成,生成汇流栅格;(4)水系的生成,在生成的汇流栅格基础上,设定不同阈值的集水区面积,获取不同密度的河网水系,此过程可由 Stream Link 命令完成。

水系提取关键在于网格数的选取,真实河源与提取水系河源的距离误差是随着网格数的变化而变化,通过拟合网格数与距离误差的函数,得到距离误差最小值所对应的网格数,并通过此网格数提取流域水系。在 500~20 000 网格数范围内,选取若干网格数分别生成水系图。在 ArcGIS 9.2 下加载晋江河源数据与提取的水系图,逐一量测提取水系河源与实际河源坐标的距离误差,从而得到距离误差与网格数的相关关系,如图 2 所示。

从图 2 看出,当网格数增加到 6 000 左右时,实际河源与提取水系河源之间的距离误差变化呈由大到小的趋势,而当网格数从 6 000 左右继续增加时,距离误差又呈现由小到大的变化趋势。该结果表明,当网格数位于 6 000 附近,实际河源与提取水系河源之间的距离误差最小。为了进一步确定最小误差及其对应的网格数,通过 Excel 中的函数拟合,

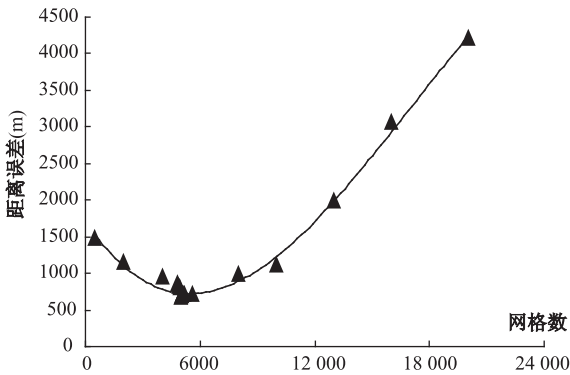


图2 网格数 500 - 20 000 时误差变化曲线图
Fig. 2 The relationship between grid number and distance error

采用曲线拟合法对图 2 所示的数据进行拟合得到:

$$y = (-8) \times 10^{-10} x^3 + 4 \times 10^{-5} x^2 - 0.3847x + 1689$$

(1)

考虑到函数的复杂程度与偏导数的计算,采用多项式拟合,选取函数最高阶为三阶。式中 y 为距离误差, x 为网格数,相关系数 $R=0.995$ 。继而求取式(1)的一阶偏导数,并令其一阶偏导数等于 0,再根据网格数的变化区间,确定其最小值所对应的网格数为 5 814,即网格数为 5 814 时所对应的水系即为晋江流域水系(如图 3)。



图3 网格数为 5814 时晋江流域的水系图
Fig. 3 River networks of Jinjiang basin with grid number of 5814

2.3 分维值计算与分析

本文通过计算水系分维值对提取流域水系的合理性进行了分析。河流水系分维值的计算方法有很多,本文采用网格法或称盒子数法;设用来覆盖测量河流片断的粗视化网格边长为 r ,而其中

包含有河流片断的网格数目为 $N(r)$;随着粗视化的网格边长 $r_1, r_2, r_3 \cdots r_i$ 的变化,覆盖河流片断的网格数目 $N(r_1), N(r_2), N(r_3) \cdots N(r_i)$ 也会随之变化。其中, B 为无量纲量, D_H 代表分维值。

$$N(r) \propto Br^{-D_H}$$

(2)

对公式两边取对数:

$$\ln(N(r)) \propto \ln B - D_H \ln(r)$$

(3)

在坐标图上以 $\ln(r)$ 为横坐标, $\ln(N(r))$ 为纵坐标,可以描绘出两者相互关系的坐标点

通过最小二乘法可以拟合出一条直线:

$$\ln(N(r)) = A - D_H \ln(r)$$

(4)

此直线的斜率的绝对值 D_H 即是水系分维值。
考虑到分形无标度区间的选择与数据分辨率的问题^[17],采用 50~4000m 不等的格网边长对流域水系图(图 3)进行分析,分别得到不同格网边长条件下,流域水系所占的网格数,即 $N(r_i)$ 。继而根据公式(4)得出晋江流域的水系分维值为 1.389,如表 1。

表 1 水系提取结果的分维值

Tab. 1 Fractal dimension of result of river network extraction

流域	网格数	最小河流长度 度(m)	河网密度 (km/km ²)	相关系数	D_H
晋江	5814	42	0.338	0.999	1.389

王林等^[14]基于相同分辨率的晋江流域 DEM 数据,通过最小河源长度与水系分维值相互关系,找出二者变化过程中的“拐点”(图 4),确定晋江流域水系。如图 4 所示,当最小河流长度为 150m 时,二者变化处于“拐点”位置,计算其河网密度为 0.167km/km²,水系分维值为 1.056。然而对于该“拐点”的取值仅凭视觉上的判断,并没有充分的依

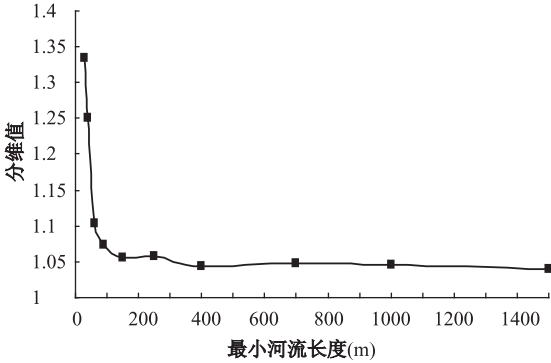


图4 最小河长与河网分维关系图
Fig. 4 Relationship between the minimum river lengths and fractal dimensions

据。该分维值与本研究计算得到的 1.389 存在一定的差距。

从区域水文地质地貌特征来看,晋江流域与闽江流域同属一个水文区划单元,地质地貌特征基本相似,分维值应当比较接近。何隆华^[18]等利用网格法计算闽江流域的水系分维值为 1.67,而本文通过河源最小误差法计算得到的晋江流域水系分维值为 1.389,在数值上与闽江流域的水系分维值更接近。因此,采用河源最小误差法所得的结果是合理客观的。

3 结论与讨论

(1)在没有数字水系与遥感数据资料的条件下,本文提出以 DEM 提取流域水系的河源最小误差法,即以水系河源与实际河源距离误差的最小值来确定河网水系,解决提取流域水系的唯一性问题。

(2)以晋江流域为例,用分辨率 30m 的 DEM 数据,通过河源最小误差法提取了晋江流域水系。结果表明,当网格数为 5 814、最小河流长度为 42m 时,实际河源与提取水系河源的距离误差达到最小。在此基础上,通过计算提取水系的分维值初步说明提取的晋江流域水系是比较合理的。

河源最小误差的缺陷在于更多适应山区性河流的水系提取,而对于平原地区的河网水系提取则需要修正。对于以河源最小误差法提取流域河网水系结果误差的验证,在没有资料的前提下,河源最小误差法的意义在于通过数学函数的最小值能解决流域水系的唯一性。但其在理论与实践,仍有待今后更深层次的研究。

参考文献:

- [1] 芮孝芳. 水文学原理[M]. 北京:中国水利水电出版社, 2004, 18 - 19.
- [2] 何钢,蔡运龙. 不同比例尺下中国水系分维数关系研究[J]. 地理科学, 2006, 26(4): 461 - 465.
- [3] 王大鹏,王周龙,李德一,等. 太湖流域水系分形特征的空间分异研究[J]. 浙江水利科技, 2006, 145, 5 - 7.
- [4] Hutchinson M F. A New Procedure for Gridding Elevation and Stream Line Data with Automatic Removal of

- Spurious Pits[J]. Journal of Hydrology, 1989, 106: 21 - 232.
- [5] Saunders W, Maidment D R. Grid-based Watershed and Stream Network Delineation for San Antonio Nueces Coastal Basin[C] // Proceedings of Texas Water 95: A Component Conference of the First International Conference of Water Resources Engineering. San Antonio: ASCE, 1995, 16 - 17.
- [6] Turcotte R, Fortin J P, Rousseau A N. Determination of the Drainage Structure of a Watershed Using a Distal Elevation Model and a Distal River and Lake Network [J]. Journal of Hydrology, 2001, 240: 225 - 242.
- [7] Xu Xinliang, Zhuang Dafang. Automated Extraction of Drainages in China Based on DEM in GIS Environment [J]. Geoscience and Remote Sensing Symposium, 2003, 5: 3344 - 3346.
- [8] 王加虎,郝振纯,李 丽. 基于 DEM 和主干河网信息提取数字水系研究[J]. 河海大学学报(自然科学版), 2005, 33(2): 119 - 122.
- [9] 许悍卫,何 江,余远见. 基于 DEM 与遥感信息的秦淮河流域数字水系提取方法[J]. 河海大学学报(自然科学版), 2008, 36(4): 443 - 447.
- [10] 郑子彦,张万昌,郇庆国,等. 基于 DEM 与数字化河道提取流域河网的不同方案比较研究[J]. 能源科学, 2009, 31(10): 1730 - 1739.
- [11] 孔凡哲,李莉莉. 利用 DEM 提取河网时集水面积阈值的确定[J]. 水电能源科学, 2005, 23(4): 65 - 67.
- [12] 熊立华,郭生练. 基于 DEM 的数字河网生成方法的探讨[J]. 长江科学院院报, 2003, 20(4): 14 - 17.
- [13] 林峰,陈兴伟,王林. 基于 DEM 的九龙江流域水系分维估算[J]. 水资源与水工程学报, 2009, 20(1): 29 - 32.
- [14] 王林,陈兴伟. 基于 DEM 的流域水系分维计算与结果分析[J]. 地球信息科学, 2007, 9(4): 133 - 137.
- [15] 陈斌. 晋江江源考察论证[J]. 人民珠江, 2003(6): 18 - 21.
- [16] 成海宁,刘少创. 关于大河源头确定标准问题的探讨[J]. 科技在线, 2009(6): 24 - 28.
- [17] 巫兆聪. 分形分析中的无标度区确定问题[J]. 测绘学报, 2002, 31(3): 240 - 244.
- [18] 何隆华,赵 宏. 水系的分形维数及意义[J]. 地理科学, 1996, 16(2): 124 - 128.

Analysis of Minimum Error at River Source to Extract River Network Based on DEM

CHEN Dongping, CHEN Ying, CHEN Xingwei

(School of Geographical Science, Fujian Normal University, Fuzhou 350007, China)

Abstract: With the development of hydrological model, extraction of river drainage network has been a hot topic in hydrology research. River drainage network was extracted based on topographic maps or drainage maps by digitization in the early years, but the result was influenced by data source resolution. There are presently two kinds of methods to extract river drainage network based on DEM. One is to overlay the extracted river drainage network based on DEM on the river digitalized maps which came from drainage maps or vector layer of river, to make the extract drainage network more similar to the actual river networks. But the accuracy of river drainage network depends on the resolution of drainage maps or vector layer of river. The other one is based on “inflection point” to extract river drainage network, however, the assumption of “inflection point” exists the problem of choice of scale-free interval. To solve the above problems, the river source minimum error (RSME) method was presented based on DEM in this paper. First, the relationship between the distance error of the actual river source and the extracted river network source and the size of grid was established; second, the minimum distance error was adopted as the principle to solve the problem of the uniqueness in watershed drainage network extraction, and then the river network was determined. Taking Jinjiang River as an example and using DEM with 30m resolution as data source, the RSME method was adopted to extract Jinjiang River drainage network on the platform of ArcGIS9.2. The result showed that the distance error between the river source and the extracted river network source is the smallest one when the grid numbers are up to 5814 and the minimum river length is 42m, the corresponding fractal dimension is 1.389. Moreover, the result indicated that the proposed RSME method is reasonable to extract watershed drainage network.

Key words: watershed drainage network; DEM; river source minimum error method; grid number; Jinjiang River Basin