

王萍, 闫春遐, 徐少朋, 等. 2015. 暖脊特征点提取及暖脊线自动绘制方法研究. 气象, 41(6): 771-777.

暖脊特征点提取及暖脊线自动绘制方法研究^{*1}

王 萍¹ 闫春遐¹ 徐少朋¹ 卢焕珍²

1 天津大学电气与自动化工程学院, 天津 300072

2 天津市气象台, 天津 300072

提 要: 低空暖脊的出现意味着一种天气不稳定因素的形成, 因此, 基于温度场的暖脊线的正确识别对于天气预报具有重要意义。本文在分析山脊线算法用于提取暖脊线时暴露出的问题的基础上, 提出一种纠偏算法, 配合使用为反映暖脊线气象特点所构建的规则, 提取到干扰较少的暖脊特征点; 根据最小二乘法点拟合线的思想, 提出一种光滑中轴算法, 得到质量较高的暖脊线。实验表明, 本文算法绘制的暖脊线与气象专家手工绘制的暖脊线吻合较高, 击中率高达 97%, 未发现误识。

关键词: 暖脊, 断面积值纠偏法, 平滑中轴, 特征线识别

中图分类号: P49; P409

文献标志码: A

doi: 10.7519/j.issn.1000-0526.2015.06.012

Study on Methods of Warm Ridge Feature Extraction and Warm Ridge Line Automatic Drawing

WANG Ping¹ YAN Chunxia¹ XU Shaopeng¹ LU Huanzhen²

1 School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072

2 Tianjin Meteorological Observatory, Tianjin 300072

Abstract: The appearance of warm ridge in low attitude implies that an unstable weather factor is forming. Therefore, recognizing the warm ridge correctly based on temperature field data is important for weather forecasting. In view of the exposed problems when extracting the warm ridge lines using the ridge line algorithm, this paper proposes a section extreme value correction (SEVC) algorithm, by which candidate warm ridge feature points are extracted, and establishes rules based on weather characteristics to exclude non warm ridge feature points. According to the idea of least squares method used in fitting curve, this paper puts forward a kind of smooth medial axis fitting (SMAF) algorithm to get higher quality of warm ridge lines, which coincide with that weather forecasters manually draw. A large number of experiments show that the recognition rate is as high as 97% and mistaken recognition rate is zero.

Key words: warm ridge, section extreme value correction algorithm, smooth medial axis fitting algorithm, feature line recognition

引 言

气象上, 通常需要借助温度场来分析对流层低层暖空气和中高层冷空气及其之间的配比, 需要在

对流层低层分析暖脊, 在中层分析冷槽(张杰, 2006; 朱乾根, 2007)。暖脊是指从高温区中延伸出来的较狭长区域, 通常用暖脊线来表征, 即从暖中心出发, 沿系列等温度线曲率极大值点勾勒而成。在我国地域, 暖脊线通常呈南到北、西到东或西南到东北的走

* 公益性行业(气象)科研专项(GYHY201406004)资助

2014 年 8 月 14 日收稿; 2015 年 1 月 15 日收修定稿

第一作者: 王萍, 主要从事图像识别、模式识别应用研究. Email: wangps@tju.edu.cn

向(张小玲等,2010)。在我国的气象数据资料分析处理中,暖脊线的识别和绘制是通过人工或借助计算机软件(如我国的MICAPS气象信息综合分析处理系统;李月安等,2010)提供的辅助工具人工完成的。其绘制质量难免受到人为因素(如经验不足)的影响,本文试图通过构建合理的算法,实现暖脊点的自动提取及暖脊线的自动绘制,并为基于天气图的自动分析、挖掘出更多的客观天气规律奠定基础。

在地质领域,从数字化地形资料中自动提取山脊线和山谷线的研究相对成熟,常见的有适于数字化等高线资料的二维算法和适于高程网格数据的三维算法,分别称为等高线曲率极值法和地形断面极值法(黄培之等,2005),两者均是从数据中提取脊线或谷线上的候选点,再根据一定的规则对这些候选点进行分析和归类。如果将对流层等压面上的温度值等同于地形高度,暖脊线与山脊线便具有了类似的几何性质。

1 基于断面极值法的暖脊特征点分析

断面极值法从求取数据格点的梯度入手,令梯度方向为过数据点的垂直断面的法线方向,那么,等值线最大曲率点就是此断面上格点数据的局部极值点。

设曲面函数为

$$Z = f(x, y) \quad (1)$$

如果我们关注函数值的变化最大方向,可求梯度:

$$\nabla = \mathbf{n}(f_x, f_y) \quad (2)$$

式中, $f_x = \partial f(x, y) / \partial x$, $f_y = \partial f(x, y) / \partial y$ 。此梯度矢量的方向即为函数值的变化最大方向:

$$\theta_0 = \arctan(f_y / f_x) \quad (3)$$

对于离散的格点数据,可用差分代替微分。设格点 (i, j) 上的数据为 $Z_{i,j}$, 其八邻域的格点坐标及其取值如图1所示,式(3)中的 f_y 和 f_x 的三阶反距离平方权差分计算方法(陈楠等,2007)如下:

$$\begin{cases} f_y = \frac{(Z_7 - Z_1) + 2(Z_8 - Z_2) + (Z_9 - Z_3)}{8\Delta_y} \\ f_x = \frac{(Z_1 - Z_3) + 2(Z_4 - Z_6) + (Z_7 - Z_9)}{8\Delta_x} \end{cases} \quad (4)$$

将 θ_0 的计算范围 $(-90^\circ, 90^\circ)$ 映射到 $(0^\circ, 360^\circ)$, 并将其离散成8个方向,八邻域近似断面如表1所示。

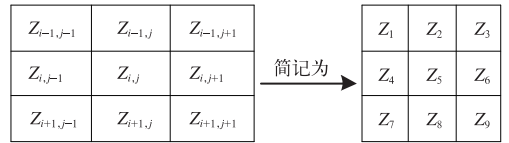


图1 格点 (i, j) 的八邻域坐标及取值

Fig. 1 Eight neighborhood coordinate and value of lattice (i, j)

表1 角度的离散化

Table 1 Discrete angles

θ'_0 的取值范围	离散化 α	断面穿过的三点
$[337.5^\circ, 360^\circ] [0^\circ, 22.5^\circ]$	0°	$(i-1, j) (i, j) (i+1, j)$
$[22.5^\circ, 67.5^\circ]$	45°	$(i-1, j-1) (i, j) (i+1, j+1)$
$[67.5^\circ, 112.5^\circ]$	90°	$(i, j-1) (i, j) (i, j+1)$
$[112.5^\circ, 157.5^\circ]$	135°	$(i+1, j-1) (i, j) (i-1, j+1)$
$[157.5^\circ, 202.5^\circ]$	180°	$(i-1, j) (i, j) (i+1, j)$
$[202.5^\circ, 247.5^\circ]$	225°	$(i-1, j-1) (i, j) (i+1, j+1)$
$[247.5^\circ, 292.5^\circ]$	270°	$(i, j-1) (i, j) (i, j+1)$
$[292.5^\circ, 337.5^\circ]$	315°	$(i+1, j-1) (i, j) (i-1, j+1)$

图2a是根据MICAPS提供的850 hPa下的温度格点数据绘制的等温线和利用断面极值法得到的暖脊候选特征点标识,根据我国温度场中暖脊的走向,只需关注 $0^\circ \sim 90^\circ$ 的特征点, \circ 表示 0° 特征点, \bullet 表示 45° 特征点, \blacktriangle 表示 90° 特征点。图2b展现了天津市气象台预报员卢焕珍在同一张等温线图上手绘制的暖脊线。通过对比,不难发现:

(1) 在真实存在的四条暖脊线附近有一些暖脊候选特征点被提取出来,特别称这些点为有效点。

(2) 有较多的候选特征点不能形成对暖脊线的支撑,称这些点为无效点。就图2a而言,无效点占全部候选特征点的比例高达52.3%(46/88)。

将图2a中的有效点和无效点进行对比发现(参见点A和点B),有效点附近等温线两侧下滑迅速,

无效点附近等温线下降平缓且等温线密集。在此,特别称前者为陡峰点,后者为坡峰点。

(3) 在有效点中,有些点的方向与暖脊线走向不一致(差 45°),如图 2a 中的点 C,对其做局部水平剖线,将剖线上的各点温度值如图 3,可以发现,这类点两侧温度值变化一缓一急,称这类点为坡-陡峰点。面对坡-陡峰点,断面极值法给出的特征点对暖脊而言出现了两点偏差,即方向偏差和位置偏差。

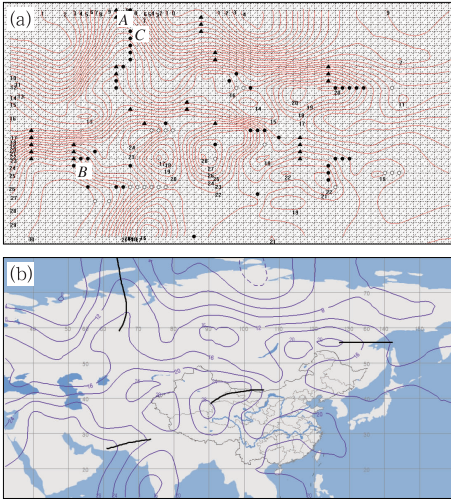


图 2 特征点和实际暖脊线

(a) 基于断面极值法的暖脊候选特征点, (b) 实际暖脊线

Fig. 2 Feature points and actual warm ridge lines

(a) candidate feature points of warm ridge lines using section extreme value method,

(b) actual warm ridge lines

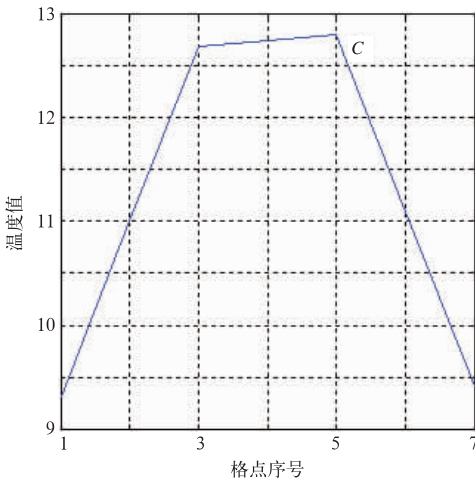


图 3 图 2a 中点 C 附近水平剖线上温度值分布

Fig. 3 Temperature distribution on horizontal section line near point C in Fig. 2a

需要强调的是,上述定义的陡峰点和坡峰点是针对特征点处等值线形态而言的,而坡-陡峰点则指的是特征点处水平或垂直方向温度值沿其两侧下降的特点。

2 暖脊特征点识别

2.1 对坡-陡峰点进行纠偏

针对上述方向偏差和位置偏差,本文将单侧梯度扩展到双侧梯度,同时扩大计算中心极值点两侧梯度的范围。首先需要将传统断面极值法划分的梯度方向范围进行扩充,为叙述方便,首先给出如下定义:

定义 1 脊宽:在断面极值点两侧或一侧连续存在略低于该极值点的系列点形成准极值区。

定义 2 显著性脊:在脊的准极值区两侧温度值迅速下降的脊称为显著性脊。

定义 3 伸展性特征点:在特征点两侧,等值线呈迅速下滑态势的特征点称为伸展性暖脊特征点。显然,前述定义的陡峰点即为伸展性暖脊特征点,由多个伸展性暖脊特征点构成的暖脊线一定是沿距离缓慢下降的,而由多个坡峰点构成的“脊”线则沿距离下降迅速,反映为一组密集的等值线,如图 4 所示。

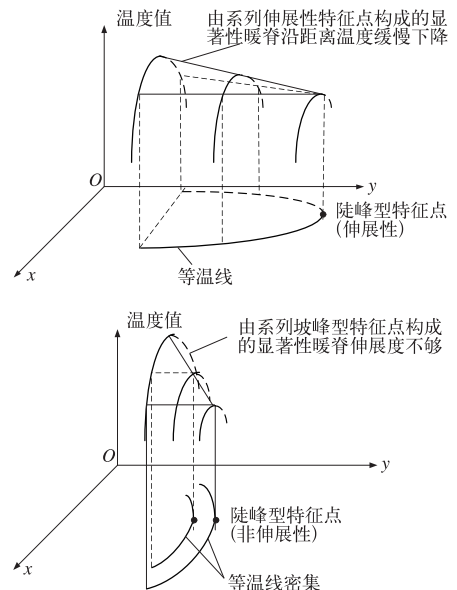


图 4 由伸展/非伸展性特征点构成的局部等温线

Fig. 4 Local contour lines composed of stretch/non stretch feature points

下面以 90° 暖脊线上左坡右陡型坡-陡峰点为例叙述纠偏算法。

(1) 求准极值区确定特征点处的脊宽

① 扫描第 i 行数据,对梯度值落入 $[45^\circ, 135^\circ]$ 范围内的点做过该点的长度为 n 的水平剖面。在剖面范围内找到极大值,将极大值点 (i, j) 标记为水平坐标轴的原点 O ,令原点右侧为正方向;

② 分别将 $(i, j+1)$ 和 $(i, j-1)$ 的温度插值到另一侧的等数值处,获得插值距离 d_R 和 d_L ,令 $r = \text{int}\{\max(d_R, d_L)\}$ 。例如,图 3 中点 C 的水平断面数据如图 5 所示, $(i, j+1)$ 点向左的插值距离 $d_R = 2.947$, $(i, j-1)$ 点向右的插值距离 $d_L = 0.029$,显然, $r=2$;

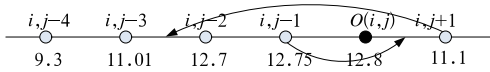


图 5 插值示意图

Fig. 5 Schematic diagram of interpolation

③ 设等温线分辨率为 ΔT ,若 $Z_{i,j} - Z_{i,j-r} \geq \Delta T$,则 $r-1 \rightarrow r$,回到③,否则继续;

④ 将 $(i, j-r)$ 点的温度向右插值求出插值距离 d'_L ,若 $d'_L \geq 0.5$,则 $c=1$,否则 $c=0$;令集合 $S \in [(i, j-r), (i, j+c)]$ 为准极值区;

⑤ 为确保区域 S 上的各点温度值相近:找到 $i + \Delta_1$ 和 $i + \Delta_2$,使 $Z_{i+\Delta_1, j-r} = Z_{i+\Delta_2, j+c} = Z_{i,j}$,若 $\max\{\Delta_1, \Delta_2\} > \rho_T$ (默认值取 $0.5 \cdot \Delta T$),则令 $r-1 \rightarrow r, c-1 \rightarrow c$,回到⑤,否则计算此处的脊宽: $s=r+c+1$ 。

(2) 显著性判断

在满足单调递减条件的情况下,将暖脊两侧各外扩 d 个单位,若:

$$\max\{Z_{i,j} - Z_{i,j-r-d}, Z_{i,j} - Z_{i,j+c+d}\} > \alpha_T$$

则宽度为 s 的脊是显著性的,保留显著性特征点、舍弃非显著性特征点(本文设定 $\alpha_T = 1.5^\circ\text{C}$)。

(3) 将脊宽的中心确定为显著性脊的特征点 (i, j_0) ,其中, $j_0 = j + (c-r)/2$ 。

图 6 展现了本文纠偏算法的效果,对比可见,图 2a 中传统断面极值法给出的 19 个坡-陡型峰点的角度及位置偏差全部得以纠正。同时,非坡-陡型峰点依然被正确地提取出来,这是因为,面对坡-坡型或陡-陡型的峰点,算法结果是 $j=0, c=0, s=1, j_0=j$ 。另外,此算法还在一定程度上降低了无效点的数量(52.3%降低到 41.4%)。

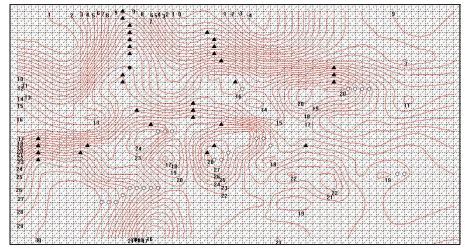


图 6 纠偏算法识别出的暖脊点

Fig. 6 Warm ridge feature points identified by section extreme value correction (SEVC) algorithm

2.2 滤除无效的坡峰点

由暖脊的定义可以看出,在温度值满足极大值阈值的情况下,等值线水平投影需同时满足曲率足够大的条件。由图 4 可知,由坡峰点组成的脊伸展力度不够,应该滤除。在等温线图上,坡峰点附近的等温线曲率小、等温线密集(图 7),温度下降坡度大(伸展度小,见图 4),于是,在坡峰点所在脊宽的外部对满足单调递减条件的 k 个格点数据进行如下运算(以 90° 坡峰点为例):

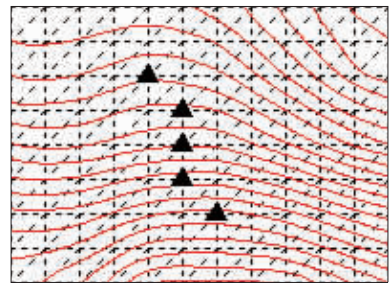


图 7 坡峰点的等值线及其分布

Fig. 7 Contour lines and distribution of slope peak points

求 Δi_{Lk} 和 $\Delta i_{Rk}, k=1, 2, 3$,使 $Z_{i+\Delta i_{Lk}, j_0-r-k} = Z_{i,j}, Z_{i+\Delta i_{Rk}, j_0+c+k} = Z_{i,j}$

$$\Delta i_L = \max\{\Delta i_{Lk}\}, \Delta i_R = \max\{\Delta i_{Rk}\}$$

若 $\min\{\Delta i_L, \Delta i_R\} < 1.5\Delta T$ 且 $\Delta i_L + \Delta i_R < 5\Delta T$,则 (i, j_0) 为无效的坡峰点,将其滤除。

2.3 传统断面极值法所得特征点的有条件保留

前两小节不论是显著性特征点的识别,还是坡峰点的滤除,均是针对 90° 和 0° 脊线而言的。考虑到 45° 暖脊线不是主流方向,首先借助传统断面极值法将它们检出,再附加判断条件,将错误的坡-陡

点滤除,即只保留满足附加条件的 45°特征点。

错误的 45°特征点是其两侧梯度变化不均匀所致(即坡-陡峰点),为此,本文将待考核点视为均值,利用方差概念构建约束量 β

$$\beta = \arctan(f'_y/f'_x) \quad (5)$$

式中, f'_y 、 f'_x 分别是八邻域格点数据纵向和横向方差,即

$$\begin{cases} f'_y = \frac{\sum_{k=2,5,8} [(Z_k - Z_{k-1})^2 + (Z_k - Z_{k+1})^2]}{6} \\ f'_x = \frac{\sum_{k=4,5,6} [(Z_k - Z_{k-3})^2 + (Z_k - Z_{k+3})^2]}{6} \end{cases} \quad (6)$$

显然,候选点纵横向方差相同时, $\beta=45^\circ$,两者相差越大, β 越趋向 0° 或越趋向 90° 。将式(6)作为保留 45°特征点的附加条件,即若 $\theta'_0(i, j)$ 为 45° 且 $Z_{i, j}$ 为 135° 方向断面上的极值,则进一步计算 β ,若 $\beta \in [30^\circ, 60^\circ]$,则认为 (i, j) 点的 45° 方向为真,保留此点,否则舍弃之。

图 8 是经过以上各步的暖脊特征点提取结果。与图 2a 相比,错误的 45°点全部被纠正、有效点由 47.7%(42/88)增加到 91.9%(34/37),无效点所占比例由 52.3%下降到 8.1%。

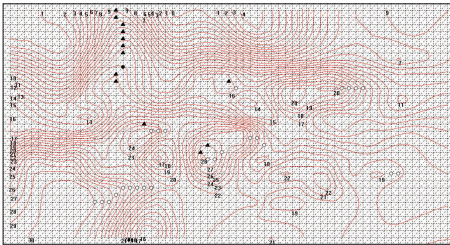


图 8 纠偏-过滤-有条件保留后的暖脊特征点

Fig. 8 Warm ridge feature points after section extreme value correction (SEVC) algorithm-filtration-conditional retention

2.4 暖脊特征点识别流程

设某高度的温度格点数据规模为 $N \times M$,对格点数据 $Z_{i, j}(i=2, \dots, N, j=1, \dots, M)$ 进入图 9 所示流程。

3 暖脊线的自动绘制

在识别出暖脊特征点的基础上,绘制暖脊线的

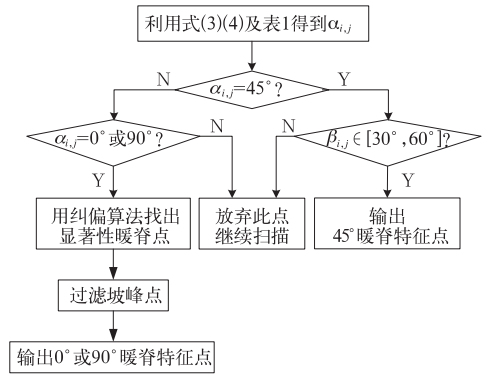


图 9 暖脊特征点识别流程图

Fig. 9 Flowchart of warm ridge feature points recognition

工作分为“定向搜索-形成折线”和“光滑中轴-拟合暖脊线”两个步骤。

3.1 定向搜索形成折线

鉴于暖脊线的方向限制在 $[0^\circ, 90^\circ]$ 范围内,按照从右到左、从上到下的顺序搜索特征点时,首先遇到的特征点一定是暖脊线的低值(尾端)。搜索算法设计如下:

- (1) 在暖脊特征点集内找到未被标记过的最低点,将该点作为当前点并作标记(黄培之,2001);
- (2) 从当前点出发,根据其梯度方向确定搜索范围 $(3+\Delta x) \times (3+\Delta y)$ 如图 10 所示,需要根据暖脊线的方向确定 Δx 和 Δy 的取值;
- (3) 重复步骤(2),直到搜索范围内未标记点空;
- (4) 重复步骤(1)~(3),将连续搜索到的暖脊特征点顺序连线,并记录起点 P_1 和终点 P_2 ,计算暖脊长度 $L = \|P_1 - P_2\|$;
- (5) 遍历暖脊线,若某条暖脊线的终点和另一条暖脊线的起点距离小于等于 Δ ,且两者走向基本一致,则首尾相接,合并成一条新的暖脊线。

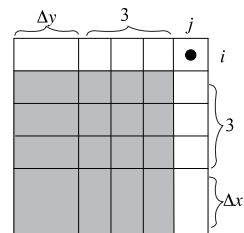


图 10 暖脊点搜索范围

Fig. 10 Searching scope of warm ridge points

(6) 遍历各条暖脊线,若 $L \leq 3$,则表明此区域暖脊强度较弱,删除此暖脊线。

对图 8 执行上述算法的 4 条折线如图 11 所示。

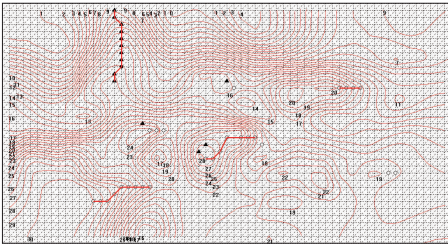


图 11 暖脊折线效果图

Fig. 11 Effect of broken warm ridge lines

本文以暖脊线上多数特征点方向作为暖脊方向,以暖脊线的长度表示其强度。

3.2 光滑中轴拟合算法

人们通常使用最小二乘法用已知数据点拟合出一条光滑的曲线(倪慧等,2011),这时,需要设定曲线的函数形式,而暖脊线因其多变性和不定性,试图建立统一有效的数学模型是不客观的,本文基于“找到某条光滑曲线,使其即反映数据点预示的变化趋势,又使光滑曲线对数据点拟合的误差平方和最小”的最小二乘设计思想,构建光滑中轴算法,其中,中轴概念如图 12 所示。

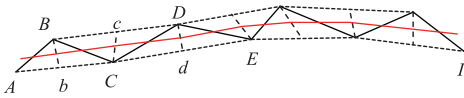


图 12 离散序列点的中轴

Fig. 12 Medial axis of discrete sequence points

图中,设 A, B, \dots, I 为定向跟踪算法得到的属于一条暖脊线的暖脊特征点,虚线为辅助线, b, c 分别为 AC 和 BD 的中点,以此类推,顺势连接 Bb 中点、 Cc 中点、 Dd 中点……并将端点适当延长,即可得到均匀地穿过离散数据点的中轴折线,再用二次 B 样条对中轴折线做进一步平滑处理,得到光滑的中轴。经此算法处理得到的平滑缓变的暖脊线如图 13 所示。

4 测试及分析

为了验证本文算法的正确性和有效性,随机采集 2005—2013 年间 30 组温度场数据(850 hPa)进

行测试。图 14 是 2005 年 7 月 31 日 08 时的数据样本。图 14a 中的四条暖脊线由气象专家在 MICAPS 系统上手工绘制而成,图 14b 的背景是以三角网格法得到的等温线(戴泽军等,2003;张利等,2009),基于本文算法的全部暖脊线识别结果自动绘于其上。

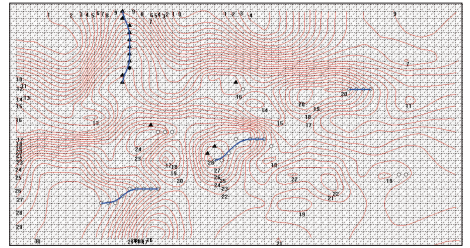


图 13 光滑中轴算法得到的暖脊线效果图

Fig. 13 Effect of warm ridge lines using smooth medial axis fitting (SMAF) algorithm

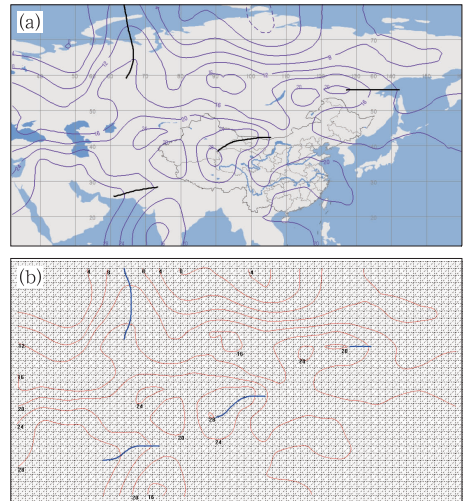


图 14 暖脊线

(a)气象专家绘制,(b)本文算法绘制

Fig. 14 Warm ridge lines

(a) warm ridge lines drawn by weather expert, (b) warm ridge lines identified by proposed algorithm

表 2 是对 30 组样本检验后的统计结果。可以看出,针对测试样本,本文算法的暖脊线击中率高达 96.91%,且未出现误识,凡击中的暖脊线,在一定程度上比人工分析更精细,且方便提供量化参数;漏识原因分析:本文算法基于局部思想而设计,识别暖脊线时欠缺顾全整体,在特征点稀疏处,容易出现一个暖脊系统被分裂为两个及以上的情况,如果分暖脊线长度较短,会因其强度不够被删除。

表2 对本文算法的测试结果

Table 2 Testing result of proposed algorithm

样本数 (等温线图)	真实暖脊条数	击中率 /%	误识率 /%	漏识率 /%
30	97	96.91	0	3.09

5 结论和讨论

针对天气图分析的实际需求和人工绘制暖脊线效率低、质量参差不齐的问题,本文进行了暖脊线自动识别方法的研究。将地形中山脊线的识别方法进一步改进用于暖脊识别,改进措施是根据对流层温度场的特点进行了改进,提出的“纠偏算法”将极值点放宽到准极值区,有效解决了面对坡-陡峰特征点时传统断面极值法出现的方向偏差和位置偏差问题。在函数模型未知的情况下,根据最小平方误差的思想,提出基于离散数据点进行曲线拟合的“光滑中轴算法”,使自动绘制的暖脊线符合气象业务规范与气象专家手工绘制的暖脊线吻合较高。对于测试样本,在未出现误识的前提下获得了接近97%的击中率。

暖脊自动识别的实现一则可以协助业务员快速准确找出格点数据所反映出来的具有气象意义的特

征,二则特别适合于从多年的气象资料中,挖掘出与暖脊相关的客观气象信息。再就是,将该方法做适当改进,可以实现高度场上脊线的自动识别,以及副高、南亚高压等系统脊线的自动识别。

参考文献

- 陈楠,王钦敏,汤国安. 2007. 基于DEM的坡向提取算法对比分析. 遥感信息,16(1):70-75.
- 戴泽军,苗春生,禹伟,等. 2003. 一种绘制地面天气图及要素场等值线方法. 南京气象学院学报,26(1):130-135.
- 黄培之. 2001. 提取山脊线和山谷线的一种新方法. 武汉大学学报,26(3):247-252.
- 黄培之,刘泽慧. 2005. 基于地形梯度方向的山脊线和山谷线的提取. 武汉大学学报,30(5):396-399.
- 李月安,曹莉,高嵩. 2010. MICAPS 预报业务平台现状与发展. 气象,36(7):50-55.
- 倪慧,李重,宋红星,等. 2011. 带插值条件的移动最小二乘曲线拟合. 浙江理工大学学报,28(1):135-139.
- 张杰. 2006. 中小尺度天气学. 北京:气象出版社,pp. 281.
- 张利,王俊彪,张贤杰. 2009. 基于矩形网格追踪法的曲面主曲率等值线生成算法. 计算机应用研究,26(8):3179-3181.
- 张小玲,张涛,刘鑫华,等. 2010. 中尺度天气的高空地面综合图分析. 气象,36(7):143-150.
- 朱乾根. 2007. 天气学原理和方法(第四版). 北京:气象出版社,pp. 649.