

利用多时刻资料作低纬地区风场同化的试验*

冯灵敏 王五在 周紫东

(中央气象局气象科学研究所)

提 要

本文讨论了在低纬地区利用多时刻风的观测资料作资料同化的一些试验结果。在资料同化过程中,采用经验正交迭代法。有两种组成资料矩阵的方法。一是把风的两个分量分别作为两个资料矩阵的元素,另一个是把两个分量作为同一资料矩阵的元素。对这两种方法的试验和比较表明,后者比前者有更好的结果。资料同化所得到的风场,可用于客观分析。

一、引 言

低纬地区风场的分析是十分重要的。在风场的客观分析中一个急待解决的问题是如何充分利用从东经 140° 赤道上空静止气象卫星的云图照片所导得的非常规的风的资料问题。由于这类资料的所在位置在空间和时间上的变化很大,使得它在对客观分析的应用上存在一定的困难。针对这种情况,我们试验了一种利用多时刻常规的和非常规的资料作低纬地区风场同化的方法。这就是利用多时刻资料,借助经验正交迭代方法得到某些事先选定的测站或网格点上风的估计值,也称同化值。这个方法的原理是Barnett在处理太平洋信风场的资料时首先提出来的^[1]。可是在他的工作中,在组成风的资料矩阵时,他把风的两个水平方向的分量分别作为两个资料矩阵中的元素,也就是说,风的两个水平分量决定二组不同的特征值和其对应的特征向量;但是风是一个向量,从物理上看,它的两个水平分量应该共同决定着一组特征值和其相应的特征向量。为此,在本文中,我们改进了Barnett的做法,提出了一种更符合问题的物理特点的组成资料矩阵的方法,并得到了较好的效果。

以下就资料的来源与处理、同化、试验结果分析等几个问题进行讨论。

二、资料的来源与处理

我们选用1976年8月每日08时 $80^{\circ}\text{E}-130^{\circ}\text{W}$ 、 $45^{\circ}\text{N}-45^{\circ}\text{S}$ 区域内200 mb海洋天气图上的风场资料作为原始资料(由北京气象中心气象台提供)。在每张图上的风场资料中,有一部分为定时的常规观测资料;有一部分为根据卫星云图资料所导出的风的资料,并经过了简单的同化。由于非常规资料出现的位置随时间具有极大的变动性,因此,我们首先对这部分资料进行一些处理,即在每张图上,将海洋区域内风的非常规资料用简单的插值方法化到以5度为距离单位的经纬线的交点(格点)上,作为该交点(格点)上风的初始资料。

* 本文于1980年1月4日收到,1980年4月21日收到修改稿。

其次,以 5 天的时间间隔把 31 天资料分成 6 组(最后一组为 6 天),确定每一组的站点。在每一组内对至少有 3 天以上风的资料的站点,才选定为同化用的站点,否则就剔除。据此,如在 8 月 1 日到 5 日这个组内共选定 119 个点,11 日到 15 日共选定 138 个点,26 日到 31 日这一组共选定 180 个点。

三、同 化

1. 资料矩阵

既然对某一组,选定了符合要求的点数 n ,于是就可以组成一个 $t \times n$ 的资料矩阵 Z , t 为该组图的张数。矩阵 Z 的元素以 $z_{i,j}$ 表示。在矩阵 Z 中某些对应于资料缺测处的元素是无值的,称为“空元素”。我们的目的在于通过计算得到这些“空元素”的估计值(同化值),变“空元素”为“非空元素”。为资料存放方便起见,在资料矩阵中这些“空元素”(无值元素)先暂以一特殊数值代之,计算前,再按一定的方法给定这些“空元素”的初始估计值,然后通过经验正交迭代,得到它们的同化值。我们感兴趣的是每一组最末一天的资料,即矩阵 Z 的最末一行里的“空元素”。如果用经验正交迭代所得到的估计值比较合理,那就意味着我们可以通过利用初始时刻(指最末一天)及其前几天的多时刻资料,经过用经验正交迭代的方法进行同化,在某种程度上得到一个较好的初始资料。

2. 经验正交迭代

我们以某种方法给出矩阵 Z 中诸“空元素” $z_{r,s}$ 的初值 $z_{r,s}^{(0)}$,并记矩阵 Z 为 $Z^{(0)}$ 。求矩阵

$$R^{(0)} = Z^{(0)*} Z^{(0)} \quad (Z^{(0)*} \text{ 为 } Z^{(0)} \text{ 的转置矩阵})$$

的最大特征值及其相应的特征向量,于是把资料 Z 场进行展开,只取与最大特征值相应的有关的项,然后倒算得到诸 $z_{r,s}$ 上高一级的近似值 $z_{r,s}^{(1)}$,以代替 $z_{r,s}^{(0)}$ 。记含有元素 $z_{r,s}^{(1)}$ 的矩阵为 $Z^{(1)}$,再求矩阵

$$R^{(1)} = Z^{(1)*} Z^{(1)}$$

的最大的及次大的特征值,并按它们的特征向量展开 $Z^{(1)}$ 场,可倒算得到比 $z_{r,s}^{(1)}$ 更高级近似的 $z_{r,s}^{(2)}$,以此代替 $z_{r,s}^{(1)}$ 。重复这个过程,最终就可以得到一个比较合理的“空元素”的添补值,我们称之为同化值。

矩阵 R 的全部特征值和特征向量可以用 Jacobi 方法同时求出来。

如果 $t \ll n$, R 至多有 t 个非零特征值,而特征值为零的至少有 $n-t$ 重,为了节省计算工作量可以采用“时空转换法”^[2]。即求出矩阵

$$\hat{R} = Z Z^*$$

的特征值 \hat{D} (对角矩阵)及特征向量 \hat{W} (矩阵)。可证

$$W = Z^* \hat{W} \hat{D}^{-\frac{1}{2}}$$

$$T = \hat{W} \hat{D}^{\frac{1}{2}}$$

其 W 为 R 的特征向量矩阵, \hat{W} 为 \hat{R} 的特征向量矩阵, T 为时间系数矩阵。于是 Z 可展为

$$Z = T W^*$$

3. 风场资料矩阵的组成方法

在组成资料矩阵时，考虑到风是一个向量，它所造成的资料矩阵只能确定一组特征值及其相应的特征向量这样一个物理特点，我们把风的两个水平分量作为同一个资料矩阵中的元素。这样，对于具有 n 个测站和网格点，用 t 次多时刻资料所组成的资料矩阵是 $t \times 2n$ 的实矩阵。对此，以后简称为 B_2 法。

在 Barnett 的工作中^[1]，风的两个水平分量分别作为两个不同的资料矩阵的元素。每个矩阵都是 $t \times n$ 的矩阵。以后简称 Barnett 的做法为 B_1 法。

就计算工作量来说，显然改进的 B_2 法要比 B_1 法节省很多的工作量。

四、试验结果分析

1. 效果分析

分析效果采用了两种方法。其一是与根据较多资料所确定的原流场图进行比较。由于在缺测处没有实测值，同化值无法与实测值比较以确定其效果优劣。但考虑到分析原流场图时所用的测值要比我们同化时用到的值多，所以可以把原流场作为比较的标准。并且规定：凡被同化的点的风向（指计算得的风向值）若与原流场的流线相切或其交角不大于 45° 均算正确，否则就算错误。其二是制造人为的缺测点，即去掉一定数量的有测值的点上的测值，再用经验正交迭代方法进行添补（同化），最后将得到的同化值与去掉的原测值比较以确定其效果好坏。

试验结果如下：

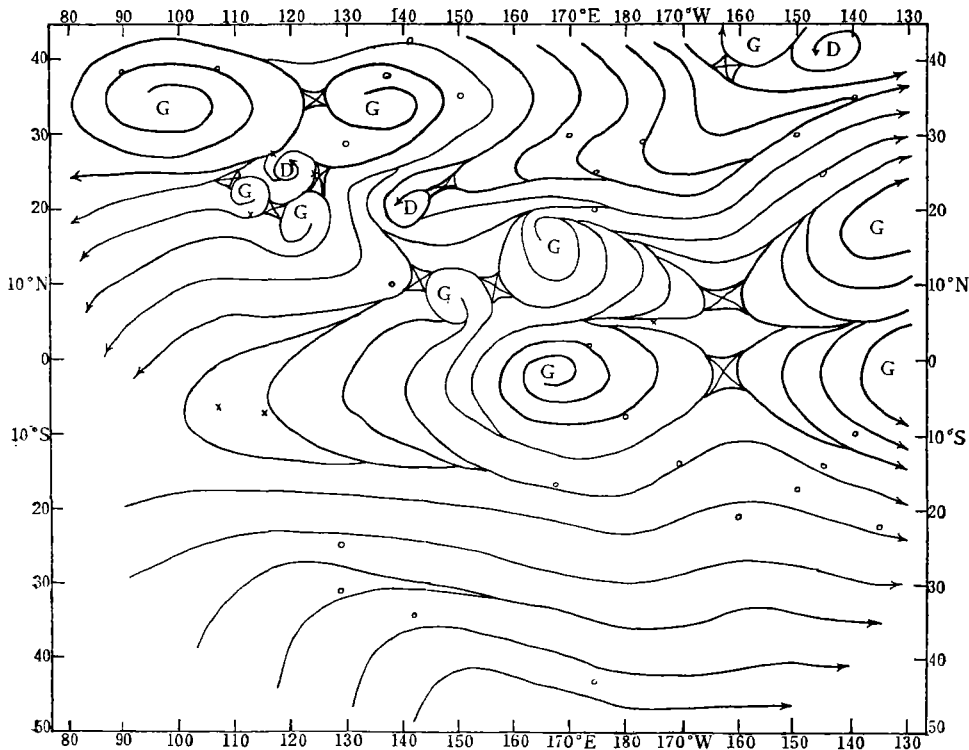


图 1a 1976 年 8 月 1 日 08 时原流场图(实况图)

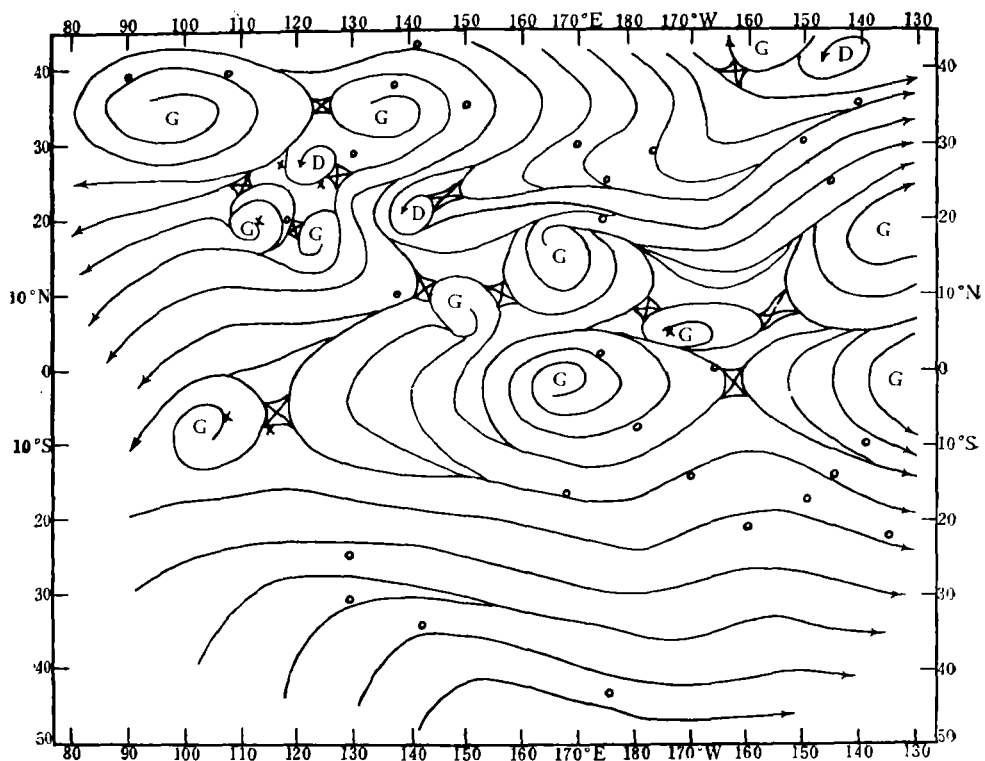


图 1b 1976年8月1日08时, B_1 法算得的同化流场图

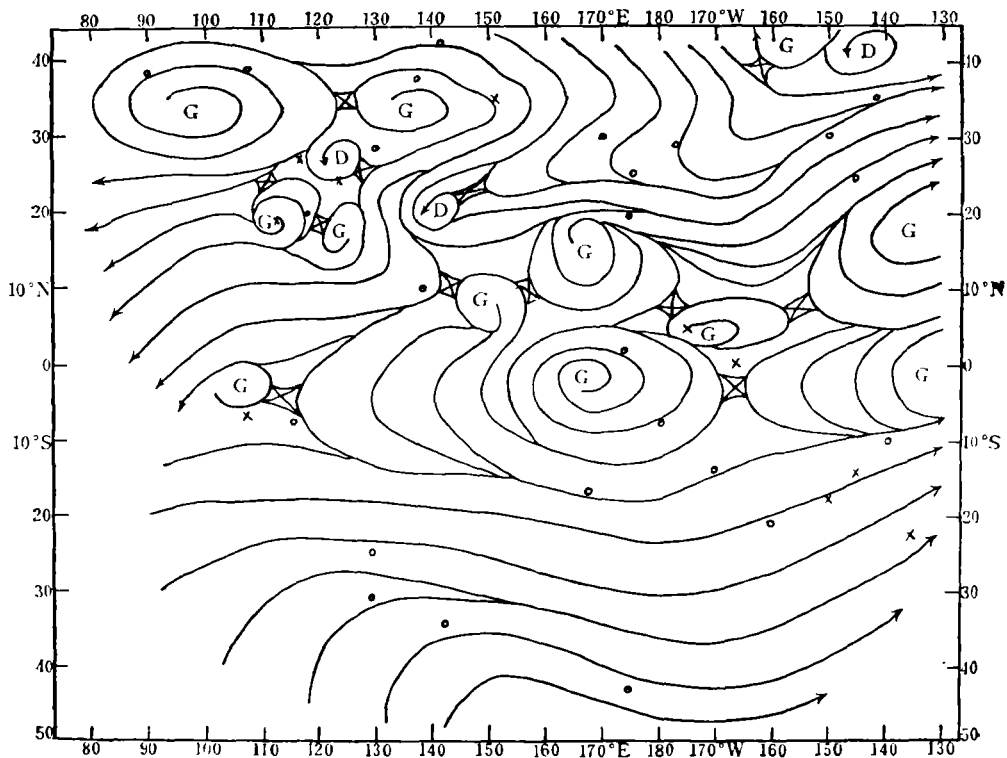


图 1c 1976年8月1日08时, B_2 法算得的同化流场图

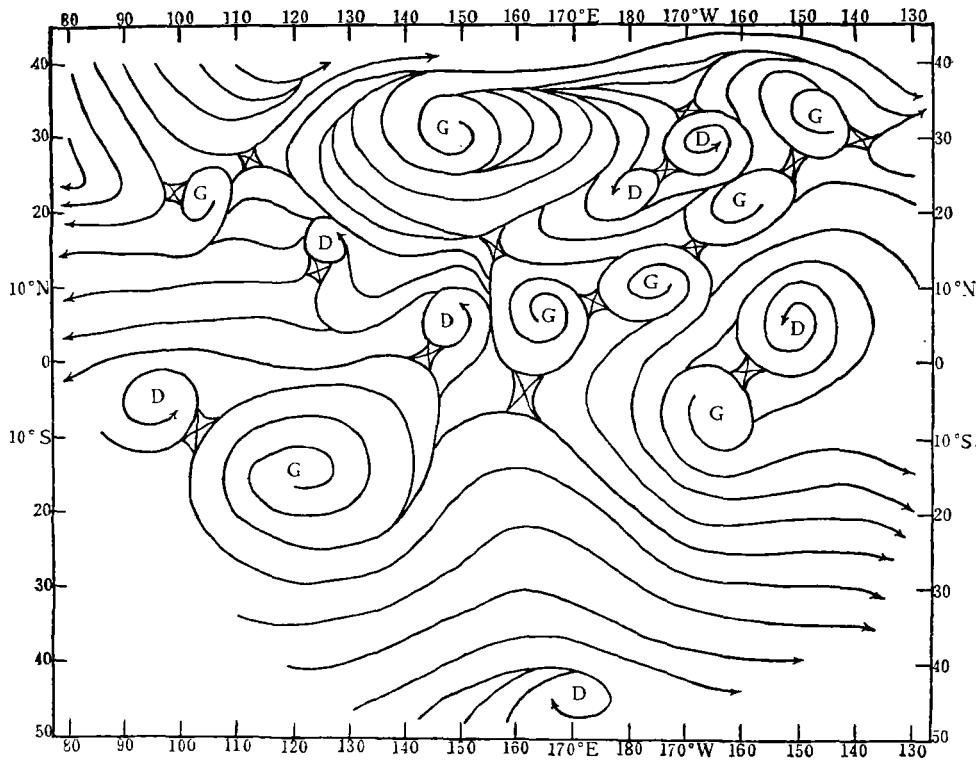


图 2a 1976年8月15日08时原流场图

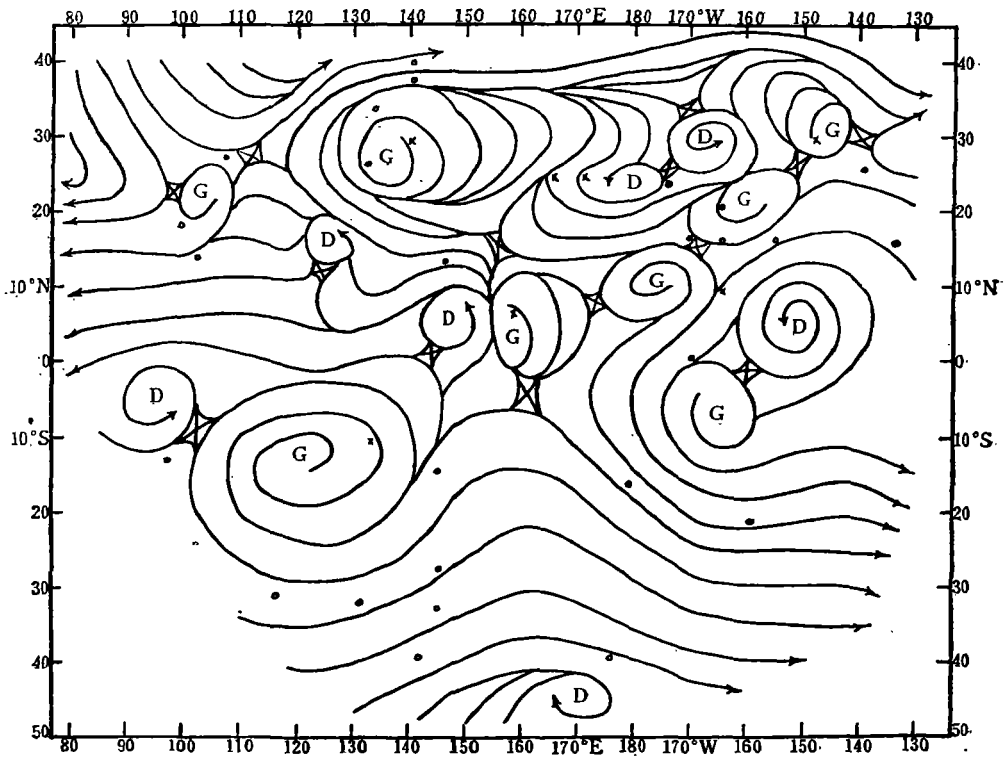


图 2b 1976年8月15日08时, B_2 法算得的同化流场图

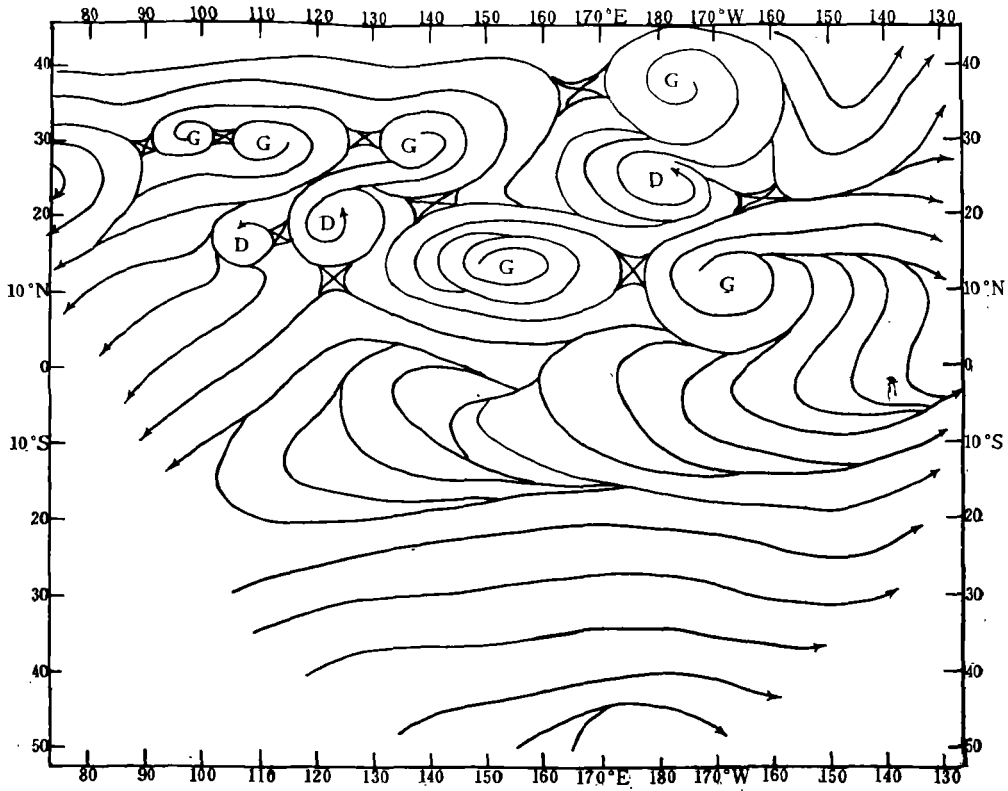


图 3 a 1976年8月31日08时原流场图

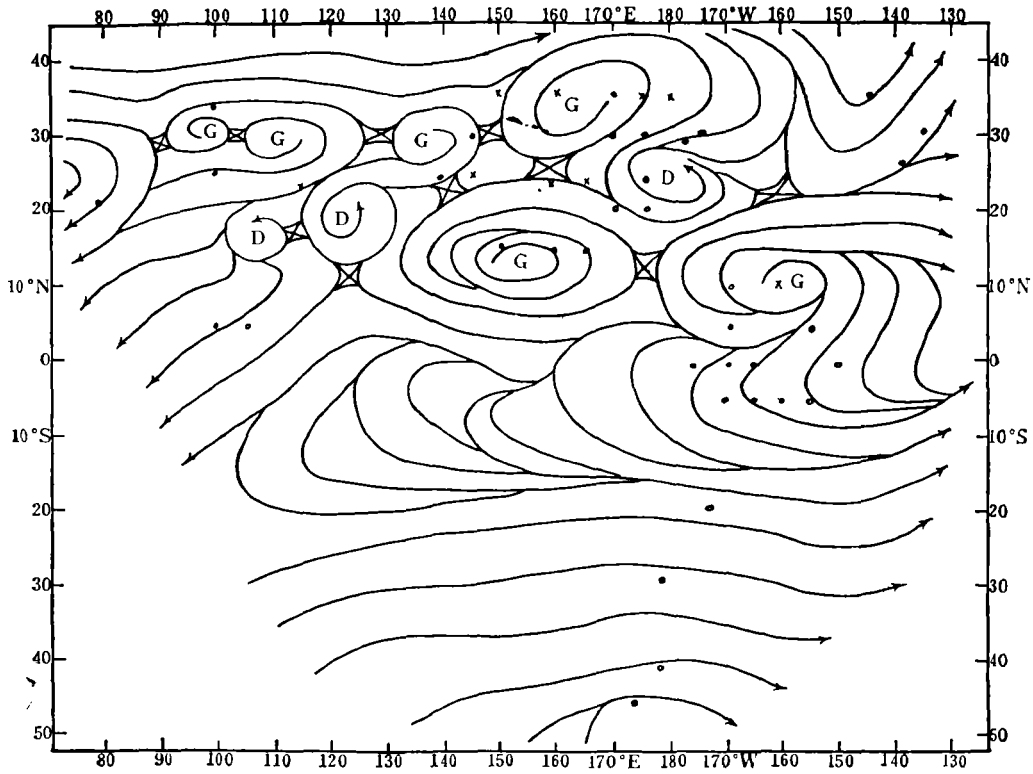


图 3 b 1976年8月31日08时, B_2 法算得的同化流场图

其一, 与原流场比较。

(1) 考察用 B_2 法所得到的风场的同化效果。我们分别给出 1976 年 8 月 1 日, 15 日, 31 日的流场分析图(它们是分别利用第 1, 3, 6 组的资料计算得到的), 见图 1b, 图 2b, 图 3b; 并且为了对比, 同时给出了相应的原流场图, 见图 1a, 图 2a, 图 3a。此外, 在同化的流场图上, 还同时给出了缺测点的分布位置; 并以“°”表示同化值是正确的, “×”则表示同化值误差较大。

从这些图上, 可以看出:

首先, 同化的流场与原流场十分一致。

在 8 月 1 日的例子中, 在 170°W 以东、赤道与 25°S 之间的比较平直的气流中, 由于缺测点上同化值比较正确, 使得流场分析比较合理。在 8 月 15 日的例子中, 不仅在 140°E 以西的南半球区域的比较平直的气流中同化的流场比较合理, 而且在 140°E 以西的北半球区域内对两个大小尺度不同的高压环流, 同化的流场分析得也比较好。在 8 月 31 日的例子中, 缺测点主要分布在 135°E 以东的北半球区域内, 这一区域有 5 个高低气压环流, 在同化的流场上, 除对其中一个高压环流的分析位置与原流场图差别较大以外, 对其余四个系统的分析也都比较好。

其次, 对那些缺测点分布在尺度比较小的系统及由小尺度系统所构成的鞍型区域附近时, 由于在这些地区资料过于稀少, 同化值的误差可能较大, 因而使得同化流场上对系统的描写能力显得较差。如在 8 月 1 日的同化流场图上, 由于同化值的误差, 使得在赤道附近 105°E 及 175°W 附近出现了二个虚假的尺度较小的高压系统。

(2) 考察资料矩阵不同取法对结果的影响

我们给出了由 B_1 法得到的同化流场图, 见图 1c。

与图 1a、图 1b 比较, 显而易见, 对分布在 170°W 以东、赤道与 25°S 之间的那些缺测点, 由于 B_1 法得到的同化值的误差较大, 使得该地区盛行西南气流, 与原流场图上的西北偏西的气流差别较大。对于 B_2 法的情况, 如前所述, 它对该地区的流场分析则是比较合理的。

(3) 为了进一步考察资料矩阵不同取法对同化效果的影响, 与原流场比较了 B_2 法和 B_1 法同化值的统计结果。

两种方法对一些个例进行计算的结果统计见表 1。

表 1 B_2 法与 B_1 法结果的比较

日 期	选取点数	缺测点数	同化值错误个数		错误比率(错误点数/缺测点数)	
			B_2 法	B_1 法	B_2 法	B_1 法
1	119	35	6	10	17%	29%
2	119	25	2	5	8%	20%
3	119	19	3	4	16%	21%
4	119	27	5	6	19%	22%
5	119	15	3	4	20%	27%
15	138	33	7	9	21%	27%
31	180	45	9	12	20%	27%

在7个个例的对比试验中, B_2 法所得同化值错误的个数以及错误比率都要比 B_1 法的少。 B_2 法在每张图上所得到的缺测点上的同化值80%以上是正确的, 但相应的 B_1 法的准确率平均只有75%左右。

综上所述, B_2 法所得的同化流场比 B_1 法所得的同化流场更为合理, 除个别尺度较小的系统外, 它与原流场十分一致。

其二, 去掉一些点上的已知测值, 再通过计算得到同化值与去掉的原测值比较。

从表2可见, 在4个个例的对比试验中, B_2 法的效果显然要比 B_1 法好, 采用 B_2 法后, 其错误比率比 B_1 法平均降低了5%。这种与实况比较得到的结论和与原流场比较

表2 B_1 法与 B_2 法的一些个例的计算结果统计

日期	选取点数	缺测点数	同化值错误点数		错误比率	
			B_2 法	B_1 法	B_2 法	B_1 法
1	119	15	3	4	20%	27%
5	119	12	3	4	25%	33%
15	138	15	1	2	7%	13%
31	180	20	2	2	10%	10%

得到的结论是一致的。

2. 问题讨论

(1) 缺测处的初值对计算结果的影响

对应于风场的缺测处的资料矩阵的“空元素”(无值元素), 计算前须按一定办法给定初始值。我们分别采用了两种办法:

一是以同一个“时空平均值”作为所有缺测处的初值, 即设 \bar{u} 、 \bar{v} 分别表示 u 、 v 分量的“时空平均值”, 则

$$\left. \begin{aligned} \bar{u} &= \frac{1}{t \times n - s} \sum_{i=1}^{t \times n} u_i \delta_i \\ \bar{v} &= \frac{1}{t \times n - s} \sum_{i=1}^{t \times n} v_i \delta_i \end{aligned} \right\} \begin{aligned} \delta_i &= 0, \text{ 在缺测处,} \\ \delta_i &= 1, \text{ 在非缺测处.} \end{aligned}$$

其中 s 为 t 张图的全部缺测点数。在本文中, 对所有个例, 在计算前, “空元素”的初值都是这样给定的。

二是在各缺测点处补以各该点的“时间平均值”为其初值, 即对第 j 点, 有

$$\left. \begin{aligned} \bar{u}_j &= \frac{\sum_{i=1}^t u_{i,j} \delta_i}{\sum_{i=1}^t \delta_i} \\ \bar{v}_j &= \frac{\sum_{i=1}^t v_{i,j} \delta_i}{\sum_{i=1}^t \delta_i} \end{aligned} \right\} \begin{aligned} &\text{在缺测处, } \delta_i = 0, \\ &\text{在非缺测处, } \delta_i = 1. \end{aligned}$$

用这两种初值分别对一些个例进行计算的结果表明: 这两种不同的初值取法对最终的同化值是有影响的。如8月1日, 在 $115^\circ \text{E} - 125^\circ \text{E}$ 、 $20^\circ \text{N} - 30^\circ \text{N}$ 这个区域用第一种办法同化值错了的两个点, 用第二种办法则未产生差错, 另一个同化值有错误的点差

错也比原来小；当然，有的点产生了新的差错（对第一种办法而言）。又如 8 月 31 日，用第二种办法比用第一种办法同化值的差错率有所减少。

因此，如何合理地给定缺测处的初值需要作进一步的研究。

(2) 测值的疏密对计算结果的影响

按本文选定站点的办法，如缺测点分布在资料过于稀少的地区，则缺测点的同化值易于产生较大的误差。由于主要补初始时刻的资料，所以在选定站点时以初始时刻那一天（一组资料的末一天）为准比较好，因为这样可使初始时刻所有的资料都能保留。

(3) 缺测点的多少对结果的影响

在一张图上，在全部被选取的站点中，缺测点数所占的比例越大，同化值的误差就越大。缺测点数所占被选点数的比例以多大为宜？这是一个尚待解决的问题。

(4) 关于迭代次数

在我们所计算的个例中， $t \ll n$ ，因此，非零特征值至多为 t 个，故迭代次数最多为 t 次，即 5 次或 6 次（26 号—31 号的一组）。计算表明：迭代计算到最高次数是不必要的。因为迭代到第 3 次或第 4 次以后，在我们所用的资料和要求精度范围内（风速精确到 1 米/秒，风向精确到 1 度），不计小数部分，计算所得结果即已稳定不变，在非缺测处算得的值则也已与原测值相同。特征值（按大小排列）递减得越快，所得结果稳定得越快。因此，根据我们的计算经验，迭代次数只要取 3、4 次就可以了。

五、结 语

分析所作的一些例子可以看出，这个方法基本上是有有效的和可行的。在对资料进行合理的处理、合理地给定缺测处的初值的情况下，可以得到较为理想的同化结果。

考虑到风是一个向量，因而可以用一个复数来表示它，资料矩阵就变成了复矩阵，而相关矩阵则是 Hermitian 矩阵；求 Hermitian 矩阵的特征值和特征向量问题可以用一般的求复矩阵的特征值和特征向量的方法，如减模 Jacobi 型方法^[3]求出来。这是我们即将进行试验的。

在工作过程中，杜行远、廖洞贤同志给予热忱的帮助，在此深为感谢。

参 考 文 献

- [1] Barnett, T. P., The principal time and space scales of the Pacific trade wind fields, *J. Atmos. sci.*, 34, 221—235, 1977.
- [2] 郑庆林、杜行远, «使用多时刻观测资料的数值预报新模式», 数值预报和数理统计预报论文集, 1973.
- [3] 计算机数学和应用数学, 重庆情报所编, 1978, 9.

SOME EXPERIMENTS OF DATA ASSIMILATION UTILIZING MULTIPLE-INSTANT WIND DATA

Feng Ling-min Wang Wa-zai Zhou Zi-dong

(Academy of meteorology, Central Meteorological Service)

Abstract

By using multiple-instant wind data, some experimental results of low latitudes data assimilation are discussed. The empirical orthogonal iteration method is used in data assimilation. Two methods in constructing data matrices are tested and compared. In the first method, the wind components are treated separately in two different matrices. In the second method the wind components are treated as two elements in one matrix. Results show that the latter is better than the former. It is indicated that the obtained assimilation values are useful informations for objective analysis.