

最优插值客观分析方法

杨晓霞 沈桐立 徐文金 周军 闵锦忠

(山东省气象台)

(南京气象学院)

摘要 本文对最优插值客观分析方法中相关函数的参数选取、测站分布与插值系数的关系、求解插值系数的方法以及初估场对插值的影响均作了试验研究,建立了12层、4个要素的可变区域、可变网格的单要素内插和多要素内插方案。该方案中分析要素和分析要素层次都是可变的。在多要素最优插值方案中,高度和风分量(u 、 v)采用多元最优插值方法,而相对湿度则应用单要素最优插值方法。

关键词 客观分析, 气象要素场

自从1955年客观分析方法首先在美国使用以来,经过三十多年的发展,现已有了一套比较成熟的解决客观分析问题的方法^{[1][2]}。大量的分析试验表明,最优插值法的效果最佳,而多要素最优内插的效果又优于单要素最优内插的效果^[3-7]。本文对最优插值客观分析方法进行了较系统的研究。为了比较,我们编制了单要素最优插值方案和多要素最优插值方案。在多要素插值方案中,用准地转高度值代替等压面高度的观测值效果较好。为了确定最佳方案,我们对相关函数中参数的选取、测站分布与插值系数的关系、求解最优插值系数方程组的解法以及初估场对插值结果的影响均作了试验分析。为了适应广大气象台站的业务,方案中采用了可变区域、可变网格的分析方法^[8],分析要素和分析要素层都取为可变的。本文分析试验了12个标准等压面上的4个要素,进行了均方误差的比较,并与手工分析图作了对比分析。

1 基本原理和计算方案

最优插值法中,格点上的分析值是由格点的初估值(又称估计值)加上订正值而求得的。其订正值由周围各测站的观测值与初估值的偏差加权求得。其权重系数(即最优插值系数)应使得格点分析值的误差达到最小。分析值可表示为

$$F_g^a = F_g^G + \sum_{k=1}^m \sum_{i=1}^{n_k} P_{ik} (F_{ik}^0 - F_{ik}^G) \quad (1)$$

式中 F_g^a 为格点的分析值, F_g^G 为格点的初估值, P_{ik} 为第 i 个测站第 k 个要素的插值系数, F_{ik}^0 为第 i 个测站上第 k 个要素的观测值, F_{ik}^G 为第 i 个测站上第 k 个要素的初估

值。 m 为参加分析的要素数目, n_k 为参加某格点分析的第 k 个要素所用的测站数。根据最小二乘法原理可得到求解最优插值系数的方程组

$$\sum_{l=1}^m \sum_{j=1}^{n_l} (\overline{f_{ik}f_{jl}} + \overline{f_{ik}f_{jl}} + \overline{e_{ik}f_{jl}} + \overline{e_{ik}e_{jl}}) P_{jl} = \overline{f_{ik}f_g} + \overline{e_{ik}f_g} \quad (2)$$

式中“—”代表平均运算, 下标 j 、 l 的意义分别与 i 、 k 的意义相同, 当 i 与 j 取值不同及 k 与 l 不同时, 代表的是不同测站不同要素间的关系。式中 f 和 e 分别代表初估误差和观测误差。其为

$$f_{ik} = F_{ik} - F_{ik}^G \quad (3)$$

$$e_{ik} = F_{ik}^0 - F_{ik} \quad (4)$$

其中 F_{ik} 为第 i 个测站第 k 个要素的真值。

用(2)式求解 P_{jl} 时运算量很大, 为计算方便, 假设观测误差与初估误差无关(实际上两者关系是很小的), 即

$$\overline{f_{ik}e_{jl}} = \overline{e_{ik}f_{jl}} = \overline{e_{ik}f_g} = 0 \quad (5)$$

即使这样, (2)式的计算量仍然很大。为此, 进一步假设

$$\overline{e_{ik}e_{jl}} = 0 \quad (6)$$

即各站要素之间观测误差相关为零, (2)式简化为

$$\sum_{l=1}^m \sum_{j=1}^{n_l} \overline{f_{ik}f_{jl}} P_{jl} = \overline{f_{ik}f_g} \quad (7)$$

式中的相关函数项 $\overline{f_{ik}f_{jl}}$ 、 $\overline{f_{ik}f_g}$ 一般是用统计方法求出, 但也可以用数学模型或物理模型给出, 后者计算量小得多。

最优插值法就是对各格点利用(7)式计算出权重系数 P_{jl} , 而后利用公式(1)计算出格点的值。这种方法所得结果的效果较好, 但它需要对于每一个格点都求解(7)式, 如果对每一个格点使用的测站及要素个数较多, 其计算量也是很大的。

本文作了一些基本假设和技术性处理, 使计算量大为减小, 并由此建立插值方案。

1.1 单要素插值方案

插值分析式由(1)式简化为

$$F_g^* = F_g^G + \sum_{i=1}^n P_i (F_i^0 - F_i^G) \quad (8)$$

式中符号含义与(1)式相同。求解最优插值系数的方程组由(7)式变为

$$\sum_{j=1}^n \overline{f_i f_j} P_j = \overline{f_i f_g} \quad (9)$$

式中 $\overline{f_i f_j}$ 为要素场初估误差的相关函数。观测事实及其分析研究表明, 其常随测站 i 与

j 之间的距离 r_{ij} 呈指数衰减。故 $\overline{f_i f_j}$ 可以表示为

$$\overline{f_i f_j} = e^{-\mu r_{ij}} \quad (10)$$

式中 μ 为参数, 可根据经验选取。解方程组(9)求出 P_i , 再由(8)式可求得分析值。

1.2 简化的多要素内插方案

一般在多元插值中, 求解插值系数的方程组比较繁琐, 计算量较大。为了减小计算量并保证客观分析的精度, 我们设计了一种简化的多元插值方案, 是用地转关系来协调高度场与风场之间的关系。

现假定风场和高度场满足地转关系, 由高度场的全微分形式可得

$$dH = \frac{f}{g} (v_g dx - u_g dy) \quad (11)$$

对上式从 j 测站到 i 测站积分, 可得 i 测站的高度值

$$H_i = H_j + \frac{f}{2g} [(v_i + v_j)(x_i - x_j) - (u_i + u_j)(y_i - y_j)] \quad (12)$$

式中 H_j , v_i , v_j , u_i , u_j 分别用 i 和 j 测站的观测值代替。在 i 测站周围选 4 个最近的测站, 每个测站到 i 测站求积分, 得到 4 个积分值, 然后再把 4 个积分值加权平均作为 i 测站的协调高度值(即准地转高度值)^[3]。即

$$H_i = \sum_{j=1}^4 w_j H_{ji} / \sum_{j=1}^4 w_j \quad (13)$$

式中 w_j 为权重函数

$$w_j = e^{-\mu r_{ij}^2} \quad (14)$$

r_{ij} 为两站间的距离, μ 可根据经验选取。

1.3 建立客观分析字典

进行插值分析时, 要挑选格点附近最佳影响测站。为了节省寻找测站的时间并取得较好的效果, 本文建立了格点与测站之间固定的信息场, 称之为字典。Gandin 指出: 当测站数大于 6 时, 相对均方差变化很小; 而测站数增多, 则计算量迅速增大。因此, 为计算快捷并保证一定的精度, 仅选取 4 个影响测站。

在求解测站的初估值时, 要用到测站周围的影响格点。因此, 我们选取测站周围的 4 个最近格点, 建立了“站带格字典”。类似地, 还建立了“站带站”字典, 供多元最优插值中计算协调高度值时使用。

1.4 线性方程组的求解

求解最优插值系数的线性方程组时, 遇到两个问题: 一是观测站过于靠近时, 方程组可能产生病态; 二是方程组系数矩阵趋于零时, 求逆矩阵解方程组会导致计算不稳定。为了避免上述问题, 最好用迭代法求解。但是, 迭代法求解要耗费大量的机时。为

了避免方程组出现病态, 我们剔除了相距很近的测站。另一方面, 相关函数不仅与距离有关, 还与参数 μ 有关, 选取适当的 μ 值可以避免方程组系数矩阵趋于零。

我们建立了精确法求解和迭代法求解二个方案, 一般情况下用精确法求解。

1.5 初估场的建立

初估值的好坏直接影响插值分析的质量。我们设计了3种形成初估值的方案: 用权重函数法直接由观测值内插出初估值; 用逐次订正法由观测值形成初估值; 用欧洲中心T₄₂截谱模式预报值作为初估值。现简述如下。

1.5.1 权重函数法初估值

网格点上的值可用周围一些测站的观测值加权平均求得

$$F_g^G = \sum_{i=1}^N w_i F_i^0 / \sum_{i=1}^N w_i \quad (15)$$

式中 F_g^G 为权重法网格点上的分析值, 将来作为最优插值法的初估值; F_i^0 为第 i 个测站的观测值; w_i 为第 i 个测站的权重, (14)式给出了它的关系; N 为测站数。

1.5.2 逐次订正法初估值

用逐次订正法求得网格点的值, 作为最优插值法的初估值。为节约时间, 只做了两次订正。

1.5.3 预报值作为初估值

我们采用欧洲中心T₄₂截谱模式的预报值作为最优插值法的初估值。当模式的网格点与客观分析格点不一致时, 用二元二次插值法把预报值内插到分析的网格点上, 作为初估值。

2 试验结果分析

2.1 测站分布对插值的影响

由(9)式可知, 插值系数不仅与格点到测站的距离有关, 而且还与测站和测站之间的距离有关。为了进一步了解格点周围测站分布对插值的影响, 我们对4种不同测站分布做了试验分析。结果指出: 彼此较近的两站, 其插值系数较小; 孤立测站的插值系数最大; 所有测站的插值系数之和近似为1。由此可知, 最优内插法能够考虑测站的分布情况, 并增大测站稀疏地区测值的作用。

2.2 求解插值系数方程组方法对插值的影响

对比试验分析表明: 用两种方法求最优插值系数所求得的插值结果无明显差异, 但是, 迭代法所用时间是精确法的5倍。由此可见, 为了节省机时, 只要采取一些措施, 避免方程组出现病态和计算不稳定, 用精确法求解方程组也是可能的。

2.3 初估值对插值的影响

对前述的3种初估值方案的试验分析结果表明: 用欧洲中心T₄₂截谱模式的预报值

作为初估值的分析效果最好(图1)。它与手工分析图(图2)上的形势比较一致,且等高线比较光滑;用观测值直接形成的初估值不理想(图略),仅在欧亚区域分析效果较好,在洋面上和非洲大陆效果较差,等高线不够光滑,原因是后者地区测站稀疏。

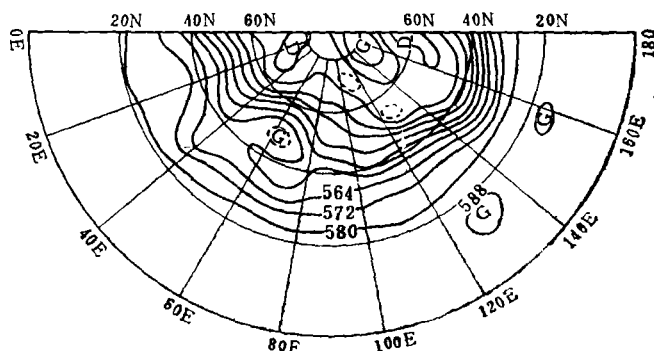


图1 1990年1月28日00Z 500hPa欧洲中心T₄₂预报模式的高度场

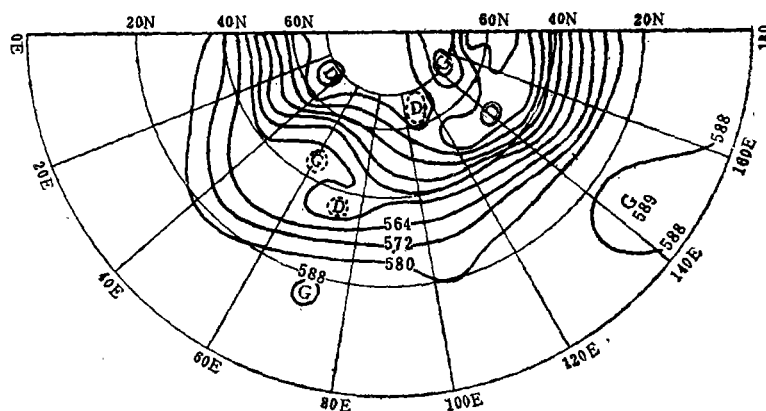


图2 同图1, 手工分析图

2.4 单要素插值分析试验结果

用单要素插值方案对12层标准等压面上的4个要素(H、u、v、g)进行最优内插分析(参见图3),分析区域为东半球(0—90°N, 0—180°E), 3.75°×3.75°的经纬度

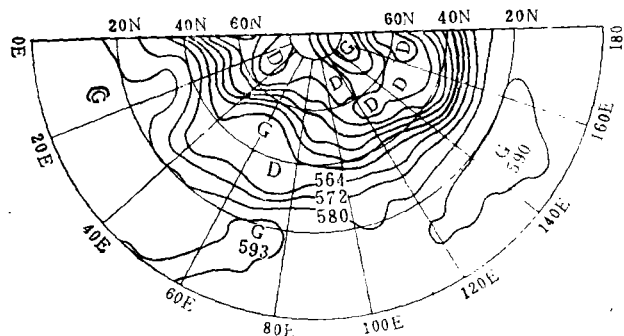


图3 同图1, 以预报值为初估值的单要素客观分析

网格距。以预报值作为初估值, 用精确求解法解线性方程组。所用观测资料进行了规范化处理。

图 3 和图 2 相比较可以看出, 500hPa 高度场客观分析图与手工分析图非常接近, 高低压中心和槽脊位置相同, 且强度差别很小。副高位置和形状与手工分析非常一致。从风场环流(图 4)来看, 风场和高度场配合较好, 符合风压关系。例如: 威海区域为明显的反气旋环流, 太平洋大槽区的风场有明显的气旋性切变, 副高区的反气旋环流非常明显。

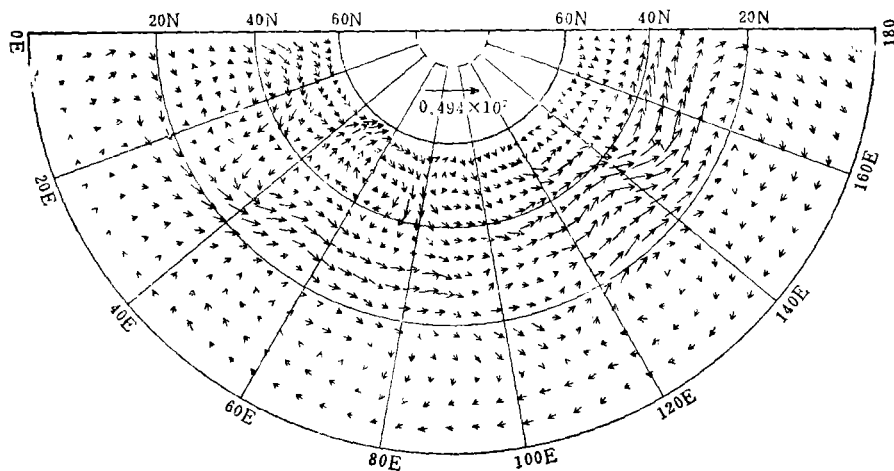


图 4 500hPa 风速分量单要素最优插值合成的风矢量

2.5 多要素插值分析试验

对 1988 年 6 月 28 日 00Z 的观测资料分别进行了单要素内插和多要素内插分析, 分析区域为 (0° — 67.5° N, 60° — 142.5° E) 欧亚区域, 分析网格距仍为 $3.75^{\circ} \times 3.75^{\circ}$, 用权重函数法由观测值直接形成初估值。比较图 5—图 7 看出: 单要素内插较接近手工分析; 而多要素内插其分析场较平滑, 滤去了高原上和低纬度地区的一些小系统, 这可

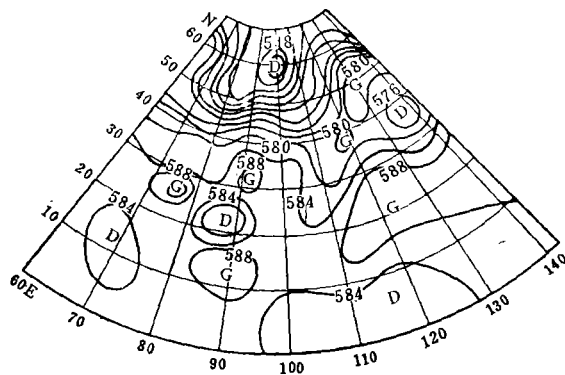


图 5 欧亚区域 500hPa 高度场单要素客观分析

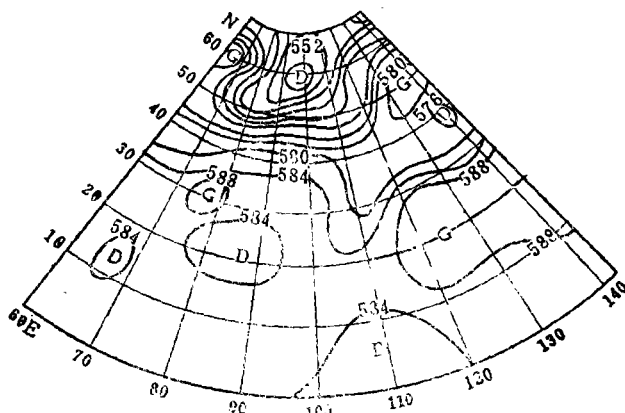


图6 欧亚区域的500hPa高度多要素客观分析

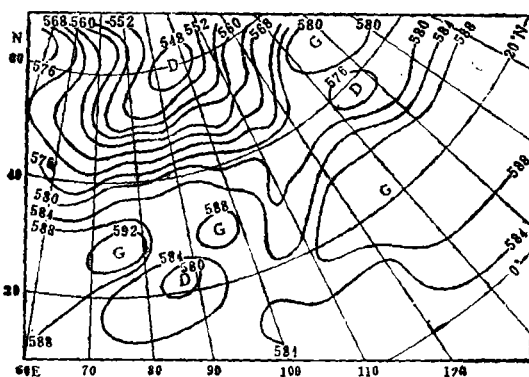


图7 欧亚区域500hPa高度场手工分析图

能与应用(13)式求协调高度值时滤去了非地转部分有关。由此可见,多要素内插可以为大尺度数值预报模式提供较好的初估场资料。

2.6 误差分析

对分析值进行误差分析,均方误差由下式求得

$$E = 1 - \sum_{i=1}^4 P_i \bar{f}_i \bar{f}_i \quad (16)$$

式中符号与(9)式中相同。

我们对东北半球的单要素内插方案,求出了均方误差分布图(图略)。其指出,洋面上和非洲大陆都为误差高值区。因这些区域观测资料稀少,初估值起着重要作用,而用预报值作初估值又比用观测资料作初估值效果好得多。

对于欧亚区域的单要素插值和多要素插值,我们分别给出初估值和最优插值的均方误差无偏估计。各种分析的500hPa高度场的均方差如下:单要素权重函数法为1.95十位势米,单要素最优内插法为1.35十位势米,多要素权重函数法为1.38十位势米,多要素最优内插法为0.88十位势米。可见,单要素权重函数法的均方差最大,多要素最优内

插的均方差最小。

3 总 结

本文对最优插值客观分析方法进行了试验研究, 建立了单要素和多要素插值分析方案。本方案的特点如下: (1)分析区域、分析网格、分析要素和分析层都是可变的; (2)可以用精确法, 也可以用迭代法求解最优插值系数的方程组; (3)初估场可以自动形成, 也可以人为给定; (4)建立了字典, 大大节省了机时; (5)可以进行单要素内插或多要素内插。

试验分析结果表明, 本方案灵活性强, 使用方便, 插值效果比较好。

参 考 文 献

- 1 张玉玲, 吴辉旋, 王晓林. 数值天气预报, 北京: 科学出版社, 1987; 392—421
- 2 廖洞贤, 王两铭. 数值天气预报原理及其应用, 北京: 气象出版社, 1986; 361—364
- 3 Seaman, Tellus, 1988; 40A, 173—176
- 4 Thomas W. Mon Wea Rev, 1975; 103; 246—257
- 5 Schlatter, Branstator, Thiel. Mon Wea Rev, 1976; 104; 765—783
- 6 陈学中, 张程道等. 气象学报, 1982; 40(3); 371—375
- 7 张程道, 邢书辰. 气象科学, 1982; 1(2); 56—64
- 8 沈桐立, 高峰. 南京气象学院学报, 1983; 1; 25—31
- 9 马吉涛, 柯小麒, 朱云桂, 狄杨波. 大气科学, 1981; 5(1); 43—49

AN OBJECTIVE ANALYSIS METHOD OF THE OPTIMUM INTERPOLATION

Yang Xiaoxia

(Shandong Provincial Meteorological Observatory)

Shen Tongli Xu Wenjin Zhou Jun Min Jinzhong

(Nanjing Institute of Meteorology)

Abstract An experimental study is made on the selection of parameters of the correlation function, the relationship between the observational station distribution and the interpolation coefficients, the method of determining interpolation coefficients, and the influence of the first guess field on the interpolation in the objective analysis method of the optimum interpolation. A scheme is thus established for the mono-element and multi-element interpolation in the 12-layer variable area and grid of four elements. The analysis of the elements and that of the element layers in the scheme are both variable. In the multi-element scheme, the multivariate optimum interpolation is used for the height and wind component (u , v), whereas the mono-element interpolation is used for the relative humidity.

Key words objective analysis, meteorological element field