

MOMS 中尺度模式的多元最优插值 客观分析方法*

沈桐立

温市耕 周慧喜**

(南京气象学院气象系, 南京 210044)

(内蒙古气象局, 呼和浩特 010051)

提 要

数值预报模式的预报结果与初始值关系较为密切, 因此改善初始场的客观分析效果, 是提高预报准确率的一个重要方面。

该文针对 MOMS 中尺度模式的特点, 研究设计了多元最优插值客观分析方法. 采用 T42 模式预报结果作为初估值, 在测站的选择、记录的订正、测站资料补缺等诸方面都考虑了中尺度模式的特点. 该客观分析方案为中尺度数值天气预报及诊断分析提供了较好的初值, 对提高中尺度模式的预报效果起到了较好的作用。

关键词: 中尺度模式; 多元最优插值; 客观分析.

引 言

客观分析是诊断分析和数值天气预报的基础性工作. 30 多年来, 已有了许多比较成熟的方法. 例如有限元法, 权重函数法、变分法、逐步订正法, 以及 1963 年 Gandin 提出的, 但直到最近才逐渐发展起来的最优插值法^{[1][2]}. 最优插值法是从统计意义入手的, 它选取一组使得分析误差达到最小的最优权重系数, 这样可避免一般加权方案的任意性. Gandin 通过数值试验调查了几种客观分析方法的精度后指出, 在观测站网密集的情况下, 各种插值方法的分析精度很接近; 而在站点稀疏时, 最优插值法的分析精度最高, 多项式插值法的分析精度最低. Seaman^[3]用最优插值法、逐步订正法和修正的逐步订正法进行实际资料对均方根误差的比较, 认为最优插值法分析效果最佳.

目前欧洲中期天气预报中心(ECMWF)(Loren^[4])、美国国家气象中心(NMC)、英国等一些国家都在资料同化系统中采用最优插值客观分析方法, 改进了预报质量.

* 本文得到中国气象局强风暴实验室及国家自然科学基金项目 49335060 及 49465013 资助.

** 南京气象学院 90 级研究生.

1994-04-14 收到, 1995-03-09 收到修改稿.

我国气象工作者,在客观分析领域中也做了大量试验研究,并已推广到业务分析工作中^[6].沈桐立等人^[6~9]对客观分析进行了系统性研究,并把模糊数学的概念引入到客观分析中.“七五”期间屠伟铭、杨晓霞、沈桐立^[10,11]等人对有限区域及全球范围业务预报中的最优插值客观分析方法进行了研究,取得了可喜的进展.

本文在上述工作的基础上,对最优插值方法做了进一步研究,参考了欧洲预报中心和美国国家气象中心的全球资料同化系统方案及我国国家气象中心资料库状况.此外,还参考了中尺度气象研究所用于研究最优插值方法的网格区域.其网格原点(区域北界的中点)是在兰勃托投影图上取在 $\lambda_0=115^\circ\text{E}$, $\varphi_0=50.5^\circ\text{N}$;格点数为 21×17 ,格距为150km.此区域内用了125个测站的资料,并加入部分地面测站,进行10个标准等压面及地面上4个气象要素的客观分析,效果比较理想.

1 最优插值法的基本原理

最优插值法是格点上的分析值由格点初估值加上订正值而求得,订正值由周围各测站的观测值与估计值的偏差进行加权求得,其权重系数不是任选的,选择的结果要使格点分析值的误差达到最小.

某一要素的分析值可表示为

$$F_g^a = \hat{F}_g + \sum_{k=1}^m \sum_{i=1}^{n_k} P_{ik} (F_{ik}^o - \hat{F}_{ik}) \quad (1)$$

式中 F_g^a 为格点 g 点上 F 要素的分析值, \hat{F}_g 为此格点上 F 要素的初估值. F_{ik}^o 为第 i 个测站第 k 个要素观测值, \hat{F}_{ik} 为第 i 个测站第 k 个要素的估计值, m 是第 i 个测站要素的总数,包括分析要素, n_k 为具有第 k 个要素的观测站数, P_{ik} 为第 k 个要素的最优插值系数.

观测值 F_{ik}^o 可以看成是由真值 F_{ik} 与观测误差 e_{ik} 两部分组成,即

$$F_{ik}^o = F_{ik} + e_{ik} \quad (2)$$

分析值 F_g^a 的均方差可表示为

$$F_{gk}^2 = \overline{(F_g - F_g^a)^2} = \overline{\{F_g - \hat{F}_g - \sum_{k=1}^m \sum_{i=1}^{n_k} P_{ik} (F_{ik}^o - \hat{F}_{ik})\}^2} \quad (3)$$

式中 F_g 为格点的真值.

把式(2)代入式(3)可得

$$F_{gk}^2 = \overline{\{f_g - \sum_{k=1}^m \sum_{i=1}^{n_k} P_{ik} (f_{ik} + e_{ik})\}^2} \quad (4)$$

上式中 $f_g = F_g - \hat{F}_g$, $f_{ik} = F_{ik} - \hat{F}_{ik}$ 分别为格点和测站的初估误差.按照最小二乘法原理,权重系数 P_{ik} 使 E_{gk}^2 达到最小的条件为

$$\frac{\partial E_{gk}^2}{\partial P_{ik}} = 0 \quad (5)$$

由式(5)得如下方程

$$\sum_{l=1}^m \sum_{j=1}^{n_l} (\overline{f_{ik}f_{jl}} + \overline{f_{ik}e_{jl}} + \overline{e_{ik}f_{jl}} + \overline{e_{ik}e_{jl}})P_{jl} = \overline{f_{ik}f_g} + \overline{e_{ik}f_g} \quad (6)$$

式(6)中,如果相关函数 $\overline{f_{ik}f_{jl}}$ 都已知,可以通过求解线性方程求出最优插值系数 P_{ji} .

在求解 P_{ji} 时,运算量很大,为了计算方便,假设观测误差与预报误差无关(实际上两者相关很小),即 $\overline{f_{ik}e_{jl}} = \overline{e_{ik}f_{jl}} = \overline{e_{ik}f_g} = 0$,但此时的计算量仍然很大,为了简便,我们进一步假设 $e_{ik}e_{jl} = 0$,即观测误差的自相关为零.方程(6)简化为

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

对于高度场和风场分量(u, v)的多变量插值分析,式(7)中的相关函数 $f_{ik}f_{jl}$ 可由数学模型和地转关系给出.这样,高度场求解插值系数的方程组可由式(7)改写为

$$\begin{cases} \sum_{j=1}^{n_H} \overline{\delta H_i \delta H_j} a_{jH} + \sum_{j=1}^{n_u} \overline{\delta H_i \delta u_j} a_{ju} + \sum_{j=1}^{n_v} \overline{\delta H_i \delta v_j} a_{jv} = \overline{\delta H_i \delta H_g} \\ \sum_{j=1}^{n_H} \overline{\delta u_i \delta H_j} a_{jH} + \sum_{j=1}^{n_u} \overline{\delta u_i \delta u_j} a_{ju} + \sum_{j=1}^{n_v} \overline{\delta u_i \delta v_j} a_{jv} = \overline{\delta u_i \delta H_g} \\ \sum_{j=1}^{n_H} \overline{\delta v_i \delta H_j} a_{jH} + \sum_{j=1}^{n_u} \overline{\delta v_i \delta u_j} a_{ju} + \sum_{j=1}^{n_v} \overline{\delta v_i \delta v_j} a_{jv} = \overline{\delta v_i \delta H_g} \end{cases} \quad (8)$$

其中 $i = 1, 2, \dots, n_H; j = 1, 2, \dots, n_u; k = 1, 2, \dots, n_v$. n_H, n_u, n_v 分别为影响格点值的具有 H, u 和 v 要素的测站数. a_{jH}, a_{ju}, a_{jv} 为最优插值系数.

对于风场分量 u 和 v 的求解最优插值系数的方程组与 H 场的类似.对于比湿场,自相关函数可用数学模型代替,进行单要素最优插值,其插值系数的方程组可表示为

$$\sum_{j=1}^{n_q} \overline{\delta q_i \delta q_j} d_j = \overline{\delta q_i \delta q_g} \quad (9)$$

式中 $\overline{\delta q_i \delta q_j}$ 为比湿场的自相关函数, d_j 为最优插值系数, n_q 为影响格点的测站个数.

由式(8)和式(9)求出插值系数,在给定初估场后,可由式(1)求出分析值.例如求高度场的分析值公式为

$$H_g^a = \hat{H}_g + \sum_{i=1}^{n_H} a_{iH} (H_i^o - \hat{H}_i) + \sum_{i=1}^{n_u} a_{iu} (u_i^o - \hat{u}_i) + \sum_{i=1}^{n_v} a_{iv} (v_i^o - \hat{v}_i) \quad (10)$$

2 最优插值法的计算方案

从理论上讲,最优插值法比其它客观分析方法具有许多优越性.试验也证明,拟合资料的精度高,数值预报的效果好.但是,用于求解最优插值系数的方程组比较繁琐,计算量较大.为了简单,本方案中作了一些如下的基本假设和技术性处理.

2.1 相关模型

我们所做的插值分析是在预先规定好的标准等压面上进行的.要素场的真值是未知量,高度场与风场的相关比较复杂,要考虑动力协调性.式(7)中的相关函数用数学模型代替.假设高度场的自相关函数为高斯分布

$$\overline{\delta H_i \delta H_j} = \exp(-\mu r^2) \quad (11)$$

假设比湿场的自相关函数也为高斯分布

$$\overline{\delta q_i \delta q_j} = \exp(-\mu r^2) \quad (12)$$

式中 r 为两点间的水平距离, μ 可按经验选取. Bergman(1979) 取 $\mu = 0.98 \times 10^{-5}(\text{km})^2$. 其它的相关可由地转关系求得, 例如

$$\overline{\delta H_i \delta u_j} = \frac{Gg}{f} \frac{\partial}{\partial y_i} \overline{\delta H_i \delta H_j} = -2\mu \frac{Gg}{f} (y_i - y_j) e^{-\mu^2} \quad (13)$$

$$\overline{\delta H_i \delta v_j} = \frac{Gg}{f} \frac{\partial}{\partial x_i} \overline{\delta H_i \delta H_j} = 2\mu \frac{Gg}{f} (x_i - x_j) e^{-\mu^2} \quad (14)$$

$$\overline{\delta v_i \delta u_j} = -\frac{G^2 g^2}{f^2} \frac{\partial^2}{\partial x_i \partial y_j} \overline{\delta H_i \delta H_j} = -4\mu^2 \frac{G^2 g^2}{f^2} (x_i - x_j)(y_i - y_j) e^{-\mu^2} \quad (15)$$

$$\overline{\delta u_i \delta u_j} = \frac{G^2 g^2}{f^2} \frac{\partial^2}{\partial y_i \partial x_j} \overline{\delta H_i \delta H_j} = -2\mu^2 \frac{G^2 g^2}{f^2} [1 - 2\mu(y_i - y_j)^2] e^{-\mu^2} \quad (16)$$

$$\overline{\delta v_i \delta v_j} = \frac{G^2 g^2}{f^2} \frac{\partial^2}{\partial x_i \partial x_j} \overline{\delta H_i \delta H_j} = 2\mu^2 \frac{G^2 g^2}{f^2} [1 - 2\mu(x_i - x_j)^2] e^{-\mu^2} \quad (17)$$

式(13)~(17)中, $G = |v|/|v_g| = 1 - \exp(-0.05|\varphi|)$, φ 是纬度, 以度为单位, 而 $f = 2\Omega \sin\varphi$ 是地转参数, g 是重力加速度.

2.2 方程组的求解

由相关函数组成的方程组可以写成矩阵形式, 求高度场插值系数的方程组为

$$\begin{bmatrix} \overline{\delta H_i \delta H_j} & \overline{\delta H_i \delta u_j} & \overline{\delta H_i \delta v_j} \\ \overline{\delta u_i \delta H_j} & \overline{\delta u_i \delta u_j} & \overline{\delta u_i \delta v_j} \\ \overline{\delta v_i \delta H_j} & \overline{\delta v_i \delta u_j} & \overline{\delta v_i \delta v_j} \end{bmatrix} \begin{bmatrix} a_{iH} \\ a_{iu} \\ a_{iv} \end{bmatrix} = \begin{bmatrix} \overline{\delta H_i \delta H_g} \\ \overline{\delta u_i \delta H_g} \\ \overline{\delta v_i \delta H_g} \end{bmatrix} \quad (18)$$

求 u 场插值系数的方程组为

$$\begin{bmatrix} \overline{\delta H_i \delta H_j} & \overline{\delta H_i \delta u_j} & \overline{\delta H_i \delta v_j} \\ \overline{\delta u_i \delta H_j} & \overline{\delta u_i \delta u_j} & \overline{\delta u_i \delta v_j} \\ \overline{\delta v_i \delta H_j} & \overline{\delta v_i \delta u_j} & \overline{\delta v_i \delta v_j} \end{bmatrix} \begin{bmatrix} b_{iH} \\ b_{iu} \\ b_{iv} \end{bmatrix} = \begin{bmatrix} \overline{\delta H_i \delta u_g} \\ \overline{\delta u_i \delta u_g} \\ \overline{\delta v_i \delta u_g} \end{bmatrix} \quad (19)$$

求 v 场插值系数的方程组为

$$\begin{bmatrix} \overline{\delta H_i \delta H_j} & \overline{\delta H_i \delta u_j} & \overline{\delta H_i \delta v_j} \\ \overline{\delta u_i \delta H_j} & \overline{\delta u_i \delta u_j} & \overline{\delta u_i \delta v_j} \\ \overline{\delta v_i \delta H_j} & \overline{\delta v_i \delta u_j} & \overline{\delta v_i \delta v_j} \end{bmatrix} \begin{bmatrix} c_{iH} \\ c_{iu} \\ c_{iv} \end{bmatrix} = \begin{bmatrix} \overline{\delta H_i \delta v_g} \\ \overline{\delta u_i \delta v_g} \\ \overline{\delta v_i \delta v_g} \end{bmatrix} \quad (20)$$

式(18)~(20)中, 左边第一个大的〔〕为相关矩阵, 其内小的〔〕也代表相关矩阵; 第二个大的〔〕及其内小的〔〕为插值系数列向量; 等式右边大的及其内小的〔〕均代表站与格点间相关函数的列向量.

求解上述线性方程组时, 存在两个问题, 一是在观测站过于靠近的情况下方程组可能产生病态, 方程系数的任何一个小的变化都会对内插系数影响很大. 内插系数是由方程系数的总体确定的, 改变这些系数中的任何一个数, 一般说来对方程组整体的影响不会很大, 因而对最优插值系数也没有多大影响. 可知出现病态是不合理的. 二是由相关函数组成的系数矩阵趋于零时, 求逆矩阵解方程组会导致计算不稳定. 为了避免这两种问题, 方程组求解最好不用求逆矩阵的方法, 而采用迭代法去解.

2.3 资料处理

经过报文处理和译码后得到的气象电报资料, 不能直接用来进行客观分析, 因为接收

的电报可能有错,在观测、发报和接收过程中也有可能出现随机性错误,必须对资料进行误差审查才能应用.大致步骤是,通过计算机把接收到的气象电报进行信息筛选,滤去无用信息,保留有用信息,对资料进行质量订正及检查,并进行补站处理.

误差的审查是根据各气象要素在垂直方向分布的相互关联进行的,把每一个站按要素分布规律和物理相关特性进行极值检查、静力学关系审查.

2.4 建立字典

通过以上检验所得资料,可用于插值分析.在进行插值分析时,要挑选格点附近最佳的影响测站.为节省寻找测站所用时间,我们建立了格点带测站和测站带格点字典,把格点周围的最佳影响测站或测站周围的最佳影响格点,按一定规则存放在一起.进行插值分析时,直接从字典中提取,而不需反复在资料文件中寻找.

2.5 建立初估场

初估场的质量关系到插值结果的质量,我们设计了 3 种初估场方案:①用权重函数法直接由观测值内插出初估场;②用逐次订正法形成初估场;③用国家气象中心 T42 谱模式预报值作为初估场.

前两种方法大家较为熟悉,这里不作介绍.而用国家气象中心 T42 谱模式预报值作初估场时,因为该模式的网格点与本方案中的网格点不一致,需内插到所分析的网格点上,采用二元等距双抛物线插值法进行内插,可以得到较满意的效果.

3 试验个例分析

本文用中国气象局强风暴实验室提供的 3 次观测资料进行了实例计算.现仅以 1991 年 7 月 21 日 500 hPa 的试验结果为例来加以说明.

图 1 给出 7 月 21 日 08 时 500 hPa 的形势场.我国东北地区为一脊;河套以西为一低压槽,槽线为东北—西南走向;华东地区受副热带高压控制,有一个 592 hPa 的高压中心位于东海上空.

我们研究了 T42 模式预报 7 月 21 日 08 时 500 hPa 的预报场(图略).东北地区的脊预报值太弱;河套附近的槽也报得偏弱且位置偏西;其东南方副高区位置及强度与实况较符合.由此可知,用 T42 的预报值直接作为中尺度模式的初始场并不适宜.

图 2、图 3 分别为逐次订正法与多要素最优插值法的分析结果.从逐次订正法的分析结果看,槽脊的基本位置正确,强度也较适宜.但由于脊附近有几个站纪录偏低,出现了不太合理的小系统;其东部分析值也明显偏低.而图 3 多要素最优插值法分析的结果可看出上述缺点得到明显克服.虽然低压槽的强度不如前者,但其用于模式,作为预报初值将优于前者.

图 4 给出单要素最优插值法分析结果,与图 3 相比,不仅等高线小波动较多,在蒙古地区还分析出一个小系统.显然这样的分析结果也不十分理想.

另外,我们还研究了以逐次订正法的结果作为初估场的多要素最优插值法.它的结果不如图 3 的结果好,对副高附近及蒙古地区的分析也有不妥之处(图略).

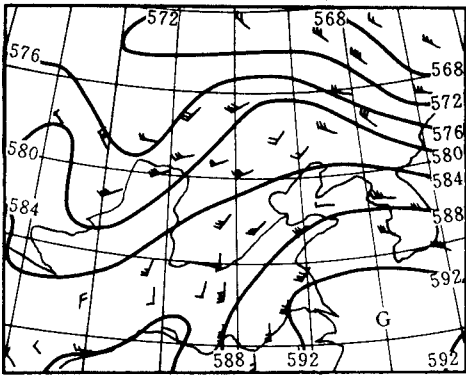


图1 1991年7月21日0000 UTC 500 hPa 实况场(手工分析)

Fig. 1 500 hPa observational field (manual analysis) at 0000 UTC 21 July, 1991

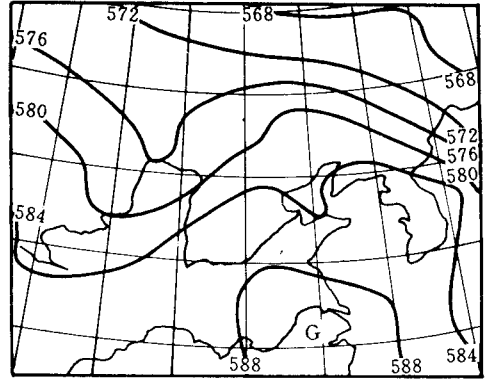


图2 1991年7月21日0000 UTC 500 hPa 逐次订正法分析场

Fig. 2 As in Fig. 1, but for analysis field of successive correlation

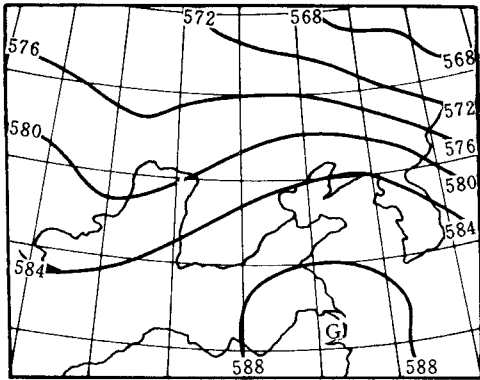


图3 1991年7月21日0000 UTC 500 hPa 多要素最优插值法分析场(用 T42 作为初估场)

Fig. 3 500 hPa analysis field of multivariate optimum interpolation at 0000 UTC 21 July, 1991 (first guess field: T42)

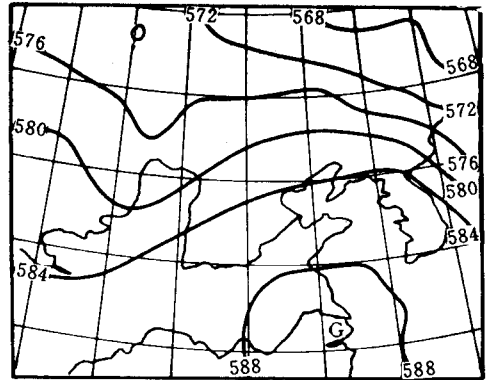


图4 1991年7月21日0000 UTC 500 hPa 单要素最优插值法分析场(用 T42 作为初估场)

Fig. 4 As in Fig. 3, but for simple optimum interpolation

综上所述,各种方案的分析结果,以使用 T42 预报值作为初估场的多要素最优插值法效果最佳,图形也较光滑.

4 小 结

本文对最优插值客观分析方法进行了研究,建立了单要素和多要素插值分析系统,其特点及有关问题讨论如下:

(1)根据试验结果,我们建立的最优插值方案的特点是:

①根据距离最近测站优先选取的原则,建立格点带站和站带格点、站带站字典,大大节省了搜寻测站和格点的时间,无记录的测站,进行补站,加快了计算速度。

②由观测资料,用逐步订正法直接形成初估场,在无 T42 资料情况下,也是一条可行之路,但效果并不十分理想。

③建立了两种求解线性方程组的方案,为了节省机时,对每一格点我们选取 6 个最近的测站进行分析,这是根据 Gandin 的研究选取的。他的实验表明,当测站数小于 6 时,相对均方差随测站的增多而迅速减小。当测站数大于 6 时,相对均方差的变化不大。测站愈多,相关函数愈不确定,求解方程组时出现病态几率亦加大。

对方程组一般用列主元高斯消去法求解,当方程组出现病态或计算不稳定时,用迭代法求解,既可省机时又可避免客观分析失败而导致业务工作无法进行的弊病。

④在应用多元插值理论时,考虑地转关系可使预报效果更好。

(2)相关模型不仅与两点间的距离有关,还与参数 μ ($0 < \mu < 1$) 有关, μ 的选取决定着有效相关距离,当测站稀疏时, μ 应取得小些,当测站密集时,应选取较大的 μ 值。适当选取 μ ,既可以提高插值精度又可以避免线性方程组求解时出现不稳定。

(3)最优插值系数不仅与格点距离有关,还与测站的分布有关,测站密集处,虽然距离较近的两测站,但其插值系数也可能较小,孤立测站的插值系数较大。

(4)初估场对最优插值起着重要作用,在观测资料稀少区域,初估值的作用尤为重要,可用较好的数值模式预报结果来作为初估值。

(5)在插值分析时,应采用经过误差检验和订正的资料,因为观测资料的误差会使客观分析效果不好。

总之,我们所建立的这套方案,既有一定的理论意义,又有较高的实用价值。使用方便、灵活,插值效果较好。但本方案还有待进一步改进与完善。

参 考 文 献

- 1 张玉玲,吴辉碁,王晓林. 数值天气预报. 北京: 科学出版社,1986. 380~433.
- 2 廖洞贤,王两铭. 数值天气预报原理及应用. 北京: 气象出版社,1986. 356~364.
- 3 Seaman R S. Some read date tests of the interpolation accuracy of Bratscch's successive method. *Tellus*. 1988, 40A : 173~176.
- 4 Lorenc. A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, 1981, 109 : 701~721.
- 5 上海市气象局研究所数值预报组. 高空图的业务客观分析. *大气科学*, 1979, 3(2) : 184~189.
- 6 沈桐立,董纯朴. 一种可变网格的客观分析方案. *南京气象学院学报*, 1982, 5(2) : 217~222.
- 7 沈桐立,高峰. 可变网格的客观分析方法. *南京气象学院学报*, 1983, 6(1) : 25~32.
- 8 沈桐立,李喜仓. 非均匀网格的客观分析. *南京气象学院学报*, 1988, 11(4) : 394~403.
- 9 沈桐立,吴诚鸿,梁佳兴,等. 模糊数学在客观分析中的应用. *气象科学*, 1987, 10(1) : 34~41.
- 10 杨晓霞,沈桐立,徐文金,等. 最优插值客观分析方法. *南京气象学院学报*, 1991, 14(4) : 566~574.
- 11 沈桐立,陈宏尧,陈春丽. 适用于中尺度数值模式的最优插值客观分析方法的研究. *空军气象学院学报*, 1994, 15(3) : 191~201.

AN OBJECTIVE ANALYSIS METHOD OF MULTIVARIATE OPTIMUM INTERPOLATION IN MESOSCALE MODEL

Shen Tongli

(*Nanjing Institute of Meteorology, Nanjing 210044*)

Wen Shigeng Zhou Huixi

(*Meteorological Bureau of Inner Mongolia, Hohhot 010051*)

Abstract

The forecast results from numerical weather prediction models are closely related to their initial values. Thus, the improvement of the objective analysis would result in the increase of forecast accuracy.

In this study, based on the features of the MOMS mesoscale model, an objective analysis method with multivariate optimum interpolation is advanced. The forecast results of T42 model is used as the first guess and the features of mesoscale model have been taken into consideration in the selection of observational stations, correction of the records, and filling up a vacancy of observation data.

The objective analysis schemes have provided good initial values for mesoscale numerical weather prediction model and diagnostic analysis, which is favourable for the great improvement of the forecast accuracy of mesoscale model.

Key words: Mesoscal model; Multivariate optimum interpolation; Objective analysis.