

气象场逐步逼近客观分析的快速算法

陈 创 买

(气象学系)

摘 要

本文对Cressman's气象场客观分析的逐步逼近技术提出了两种快速算法：FASAT法建立以测站为中心进行扫描的计算系统，一般可节省计算量55—80%以上；OFASAT法建立以测站为中心、以同一半径对全部要素和层次进行一次扫描的计算系统，节省计算量达75—99%。

对气象场作客观分析是数值天气预报和现代天气学中各种诊断分析的基本手段。随着数值天气预报的迅速发展，至今已提出几十种方法。逐步逼近技术是气象场客观分析著名的方法之一。自Cressman等人^[1,2]于1959年提出以后，已有不少改进^[3]。该方法由于具有极大的灵活性等优点，许多国家以及北京气象中心已用它来建立客观分析系统。作者在采用这种方法后发现，当网格点较多时计算速度相当慢。原因在于每次按一定影响圆半径扫描时必须大量地重复计算权函数。本文分析了Cressman逐步逼近客观分析技术，指出其算法上存在的缺点，提出了新计算方案，并对可能节省的计算量给出定量的估计和计算实例。为叙述方便，我们约定，以下凡提到计算量和节省计算量都是对权函数的计算而言的。

一、逐步逼近客观分析技术

逐步逼近法或称逐步订正法的物理学和统计学的基础，是气象要素场在水平空间上的统计相关性^[4]。

以下用三个步骤来说明这种方法：

1. 利用计算区域内有限观测站上的标量值 $\Phi_k(x, y)$ ， $k=1, 2, \dots, L$ ，以网格点为中心计算落在各网格点周围半径为 R 的影响圆内($r^2 < R^2$)的标量 Φ (设共 P 个)的平均值作为初始场。

$$\Phi(i, j) = \frac{1}{P} \sum_{k=1}^P \Phi_k(x, y) \quad (i=1, 2, \dots, n, j=1, 2, m) \quad (1)$$

$$r = [(\Delta x)^2 + (\Delta y)^2]^{1/2} \quad (2)$$

本文1983年10月收到

我们称它为一次扫描过程, 而把 R 称为扫描半径。

2. 由网格初始场反算区域内各观测站的标量内插值 $\hat{\Phi}$, 并求其与观测值之差。

观测站的标量内插值, 可用拉格朗日四网格点插值公式求得:

$$\hat{\Phi}_k(x, y) = a + b\Delta x + c\Delta y + d\Delta x\Delta y \quad (3)$$

$$(k=1, 2, \dots, L)$$

其中系数: $a = \Phi(i, j)$

$$b = \Phi(i+1, j) - \Phi(i, j)$$

$$c = \Phi(i, j+1) - \Phi(i, j)$$

$$d = (\Phi(i+1, j+1) - \Phi(i, j+1)) - (\Phi(i+1, j) - \Phi(i, j))$$

观测站的标量内插值, 也可用马克劳林(Maclaurin)九网格点插值公式求出:

$$\hat{\Phi}_k(x, y) = a + b\Delta x + c\Delta y + d\Delta x\Delta y + e(\Delta x)^2 + f(\Delta y)^2 \quad (4)$$

$$(k=1, 2, \dots, L)$$

其中系数: $a = \Phi(2, 2)$

$$b = \frac{1}{2}[\Phi(3, 2) - \Phi(1, 2)]$$

$$c = \frac{1}{2}[\Phi(2, 3) - \Phi(2, 1)]$$

$$d = \frac{1}{4}[(\Phi(1, 1) - \Phi(1, 3)) - (\Phi(3, 3) - \Phi(3, 1))]$$

$$e = \frac{1}{2}[\Phi(3, 2) + \Phi(1, 2) - 2\Phi(2, 2)]$$

$$f = \frac{1}{2}[\Phi(2, 3) + \Phi(2, 1) - 2\Phi(2, 2)]$$

从而内插值与观测值之差(即增量)为:

$$\Delta\Phi_k(x, y) = \hat{\Phi}_k(x, y) - \Phi_k(x, y) \quad (5)$$

$$(k=1, 2, \dots, L)$$

3. 以网格点为中心, 用区域内有限观测站的标量增量对网格初值用权函数 W 作加权订正。

该网格点新的标量值为:

$$\Phi'(i, j) = \Phi(i, j) + C(i, j) \quad (6)$$

其中订正值:

$$C(i, j) = \frac{\sum_{l=1}^p C_l}{\sum_{l=1}^p W_l} = \sum_{l=1}^p W_l \cdot \Delta\Phi_l / \sum_{l=1}^p W_l \quad (7)$$

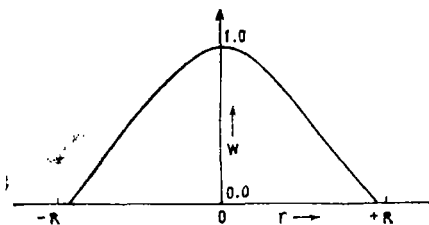


图1 权函数 W 对距离 r 的关系曲线

$$C_l = W_l \Delta\Phi_l(x, y)$$

$$W = \begin{cases} (R^2 - r^2)/(R^2 + r^2) & \text{当 } r < R \\ 0 & \text{当 } r > R \end{cases} \quad (9)$$

这个权函数的特性如图1, 显然 $W \geq 0$ 。这也称为一次扫描过程。第2、3两步可反复进行, 但所取扫描半径应逐次减少。其结果, 差值 $\Delta\Phi(x, y)$ 也将逐次减少。订正何时结束, 有两种

办法: a). 固定扫描次数, 但扫描多少次为好, 应由试验决定^[2]。b). 满足下式时:

$$|\Delta\Phi_k(x,y)| < \epsilon \quad (k=1,2,\dots,L) \tag{10}$$

其中 ϵ 为正实数。

二、快速计算的FASAT方案

从以上各步骤可看出，第1和第3步的计算，对每一个网格点，每次都以自己为中心，以 R 为半径，对全部测站进行扫描。但真正落入影响圆内的通常只有少数几个测站。当网格点数很多时，为确定那些落在影响圆内的测站，机器就需用很多时间重复地去计算权函数，这就使计算变得十分缓慢。

为提高计算速度，我们摒弃上述第1和第3步的计算方法。建立以测站为中心，以 R 为半径划测站影响方块（见下文），分散地计算方块中各网格点按方程(8)的部分加权订正值。最后计算每个网格点在该次扫描中的总订正的计算系统。图2为这种新计算方法的示意图。图中小圆点为网格点，实折线 $ABCD$ 为以第 M 测站为中心、 R 为内圆半径的正方形，可称为测站影响方块（以下简称方块），粗圆点为在方块内需要分散计算部分加权订正值的网格点。虚折线 $EFGH$ 代表以第 N 测站为中心、 R 为半径所作的同样的方块。三角形代表该方块中需要检查和计算部分加权订正值的网格点。实三角形代表为第

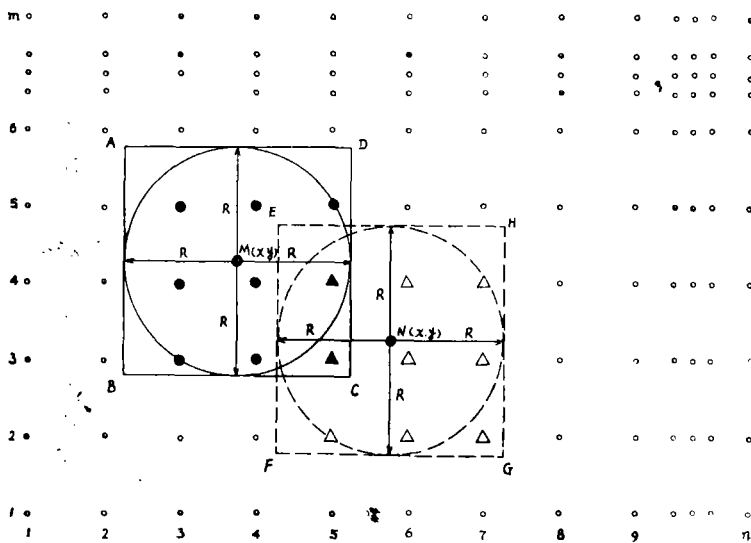


图2 $R = 1.5$ 网格长度时，以测站为中心的影响方块示意图

M 测站方块和第 N 测站方块所重叠复盖的网格点等等。这样，对任一网格点 $Q(i,j)$ 在该次扫描过程中的总订正值，由加权平均求出

$$C(i,j) = (W_M \Delta\Phi_M + W_N \Delta\Phi_N + \dots) / B(i,j) \tag{11}$$

$$(i=1,2,\dots,n; \quad j=1,2,\dots,m)$$

其中 $B(i,j) = W_M + W_N + \dots$ (12)

代表落在该网格点的影响圆内的诸测站的权函数之和。网格点上的新标量值 $\Phi'(i,j)$ 仍

按(6)式计算。并且在边界上,仅当方块部分地复盖计算区域(网格)时,才计算被复盖的网格点的部分加权订正值。若整个方块落在计算区域之外,则不对任何网格点计算部分加权订正值。

这样,就将贯穿该方法的以网格点为中心的三个步骤,都变成以测站为中心的客观分析方法,不过第1步和第3步为按不同半径 R 进行扫描,对网格点分别给出初始场和订正场。第2步才是对观测站求内插值。这样做并不改变任何计算结果,但却大幅度地节省了计算量。我们把这种方法称为气象场逐步逼近客观分析技术的快速算法,简称FASAT方案,意思是: The Fast Algorithm of Successive Approximation Technique for Objective Analysis of the Meteorological Field. 而将Cressman的算法称为原始方案。

三、计算量的估计

1. 对原始方案计算量的估计

设有 V 个要素,每个要素有 L 层等压面,每层有 $m \times n$ 个网格点,区域内共有 S 个观测站。则每个网格点需要计算 S 次权函数。若计算一个权函数需作 G 次四则运算,则一次扫描过程的运算次数为:

$$T_i = V \times L \times S \times G \times m \times n \quad (13)$$

根据方程(1)和(9),要计算一个权函数 W ,需作 $G=10$ 次四则运算。若总共作 Z 次扫描过程,则原始方案的机器运算总次数为:

$$\begin{aligned} T_{s(h)} &= \sum_{i=h}^Z T_i = (Z-h+1)T \\ &= (Z-h+1) \times V \times L \times S \times G \times m \times n \end{aligned} \quad (14)$$

式中当需形成初始场时 $h=1$, 否则 $h=2$ (下同)。

2. 对FASAT方案计算量的估计

设方块可能复盖的网格点数(以下简称方块可复盖数)为 N_g ,又设第 i 次扫描方块平均可复盖数为 N_{g_i} ,则在相同的情况下,一次扫描过程的运算次数估计为

$$T_i^* = V \times L \times S \times G \times N_a \times N_{g_i} \quad (15)$$

其中 $N_a=8$ 为每次确定方块范围需作的四则运算次数。设选用不同半径 R 作 Z 次扫描过程,则

$$\begin{aligned} T_{s(h)}^* &= \sum_{i=h}^Z T_i^* = V \times L \times S \times G \times N_a \left(\sum_{i=h}^Z N_{g_i} \right) \\ &= (Z-h+1) V \times L \times S \times G \times N_a \times \bar{N}_{g(h)} \end{aligned} \quad (16)$$

$$\text{其中} \quad \bar{N}_{g(h)} = \frac{1}{(Z-h+1)} \sum_{i=h}^Z N_{g_i} \quad (17)$$

表示对方块平均可复盖数的总平均。研究表明, \bar{N}_{g_i} 值有随扫描半径 R 的增大而迅速增大的特性。并且还可分为 R a)。取为网格长度的整数倍和 b)。取为网格长度的非整数倍这两种截然不同的情况。限于篇幅, 此处仅列出两种情况下 \bar{N}_{g_i} 的表达式:

$$a) \quad \bar{N}_{g_i} = \lceil (2R_i)^2 + (2R_i)(2R_i + 1) \rceil / 2 \quad (18)$$

例如: 当 R 分别为 $R = 4, 3, 2, 1$ 个网格长度时, 可得 $\bar{N}_{g(1)} = 39.6$, $\bar{N}_{g(2)} = 32.5$ 个网格。

$$b) \quad \bar{N}_{g_i} = \lceil 2R_i \rceil \lceil 2R_i + 1 \rceil \quad (19)$$

其中 $\lceil x \rceil$ 表示取不大于 x 的最大整数。若 R 分别取为 $R = 4.75, 3.60, 2.20, 1.80$ 个网格长度^[2], 则 $\bar{N}_{g(1)} = 53.6$, $\bar{N}_{g(2)} = 44.5$ 个网格。

3. 对FASAT方案可节省计算量的估计

先引入三个参量: 1) 定义 $DT_{s(h)}^*$ 为可节省计算量的绝对值。令, $A = Na \cdot \bar{N}_{g(h)}$, 则

$$DT_{s(h)}^* = T_{s(h)} - T_{s(h)}^* = (z - h + 1) \times V \times L \times S \times G \times (m \cdot n - A) \quad (20)$$

2) 定义 $K_{s(h)}^*$ 为FASAT方案和原始方案总计算量的比值:

$$K_{s(h)}^* = \frac{T_{s(h)}^*}{T_{s(h)}} = \frac{A}{m \cdot n} \quad (21)$$

3) 定义 $E_{s(h)}^*$ 为FASAT方案可节省的计算量:

$$\begin{aligned} E_{s(h)}^* &= \frac{DT_{s(h)}^*}{T_{s(h)}} \times 100\% = [1 - K_{s(h)}^*] \times 100\% \\ &= [1 - \frac{A}{m \cdot n}] \times 100\%, \quad (m \cdot n > A) \end{aligned} \quad (22)$$

由上述显见, FASAT方案可节省的计算量只与网格点总数 $m \cdot n$ 成正比, 而与所取测站数无关, 因为 A 是常数。例如对 $R = 4, 3, 2, 1$ 个网格长度的情形, $A = 316.8$ (对 $\bar{N}_{g(1)}$) 或 260 (对 $\bar{N}_{g(2)}$)。

四、快速计算的OFASAT方案

在上述分析中, 我们还可看到, 不但以网格点为中心的扫描方法, 许多权函数的计算是重复的, 而且即使以测站为中心, 当对不同要素和层次扫描时, 对权函数的计算也是重复的。为省去这种重复计算, 我们建立一种在一次扫描过程中以测站为中心、以同一半径 R 对全部要素和层次进行扫描的计算法。它将是最节省计算量的, 因而定名为 Optimun FASAT, 简称 OFASAT 方案。这种方案, 当以测站为中心, 用同一半径 R 对不同要素和层次进行扫描时, 可以断定, 某一定测站的方块可复盖数及具体被复盖的网格点都是固定的, 因而测站对网格点的权函数值也是固定的。若将第一个被扫描的层

次的各测站方块所复盖的网格点和相应的权函数这两种信息存储下来, 则对其余层次的扫描都可不必再重复计算权函数。为节省内存单元, 我们采用将网格编号并和权函数一齐记鼓的办法。于是一次扫描过程的运算次数可近似表为

$$T_i^{**} = S \times G \times Na \times \bar{N}_{gi} + (V \cdot L - 1) \times S \times G^* \times \bar{N}_{gi}^* \quad (24)$$

(对第一个被扫描层次) (对其余全部被扫描层次)

其中: \bar{N}_{gi}^* 为测站影响圆可实际复盖的网格点数, 平均在测站方块复盖数的八成以下。 G^* 为将一个网格点编号还原成二维坐标及取出权函数所需的运算次数共 $G^* = 10$ 。因而可取 $\bar{N}_{gi}^* = 0.8\bar{N}_{gi}$, $G^* = G$, 将它们代入(24)式, 并假定用不同半径作 Z 次扫描过程, 则总运算次数为

$$T_{s(h)}^{**} = (Z - h + 1) \times S \times G \times \bar{N}_{g(h)} \times (Na + 0.8(V \cdot L - 1)) \quad (25)$$

同样我们定义以下三个参量:

1) 定义 $DT_{s(h)}^{**}$ 为 OFASAT 方案可节省计算量的绝对值:

$$\begin{aligned} DT_{s(h)}^{**} &= T_{s(h)} - T_{s(h)}^{**} \\ &= (Z - h + 1) \times V \times L \times S \times G \times [m \times n - (\frac{0.9}{V \cdot L} + 0.1)A] \end{aligned} \quad (26)$$

2) 定义 $K_{s(h)}^{**}$ 为 OFASAT 方案与原始方案计算量的比值:

$$\begin{aligned} K_{s(h)}^{**} &= \frac{T_{s(h)}^{**}}{T_{s(h)}} = K_{s(h)}^{**'} \times K_{s(h)}^* \\ &= (\frac{0.9}{V \cdot L} + \frac{0.8}{Na}) \frac{A}{m \cdot n} \end{aligned} \quad (27)$$

其中:

$$K_{s(h)}^{**'} = \frac{T_{s(h)}^{**}}{T_{s(h)}^*} = (\frac{0.90}{V \cdot L} + 0.1) \quad (28)$$

则为 OFASAT 方案与 FASAT 方案计算量的比值。

3) 定义 $E_{s(h)}^{**}$ 为 OFASAT 方案的可节省量:

$$\begin{aligned} E_{s(h)}^{**} &= \frac{DT_{s(h)}^{**}}{T_{s(h)}} = [1 - K_{s(h)}^{**}] \times 100\% \\ &= [1 - K_{s(h)}^{**'} \cdot K_{s(h)}^*] \times 100\% \\ &= [1 - (\frac{0.90}{V \cdot L} + 0.1) \frac{A}{m \cdot n}] \times 100\% \end{aligned} \quad (29)$$

$$m \cdot n > \left(\frac{0.90}{V \cdot L} + 0.1 \right) A \quad (30)$$

和(23)式相比, OFASAT方案已将 $m \cdot n$ 的限制从FASAT方案的水平上降低了 $(1 - K_{s(h)}^{**}) \times 100\%$ 。在 $m \cdot n > A$ 的情况下, 只要 $V \cdot L > 0.90$, 则条件(30)总是成立。表明只要做两个层次的客观分析, 便可把计算量从FASAT方案的水平上减少一半。分析的层次越多, 则计算量减少越多。说明OFASAT方案确实比FASAT优越得多。在特殊情况下, 当 $V = L = 1$, 则 $[0.90/(V \cdot L) + 0.1] = 1$, 表明如果只对一个层次作客观分析, 则OFASAT方案将恰与FASAT方案等价。

五、计算实例和分析

为了阐明本方法的优越性, 下面将用一个区域性客观分析实例和一个全球性客观分析设想, 对前述三种客观分析方案的计算量作详细的估计和分析。

1. 对一个区域气象场的客观分析

用逐步逼近法对一个包含有 40×28 个网格点的区域上的 H, U, V 三个变量的气象场作客观分析。每个变量各有 $P = 1000, 850, 500, 200mb$ 四层等压面。经马吉溥^[5]方法处理, 将区域内的测站资料变成600个人工测站资料。

a) 用原始方案计算。按(14)式, 机器需作 $T_{s(1)} = 4.032 \times 10^8$ 次计算。设把这个计算量用一台速度为 5.5 〔万次/秒〕(下同)的计算机来完成, 则估计需时122.2分钟。

b) 用FASAT方案计算。按(16)式, 取 $\bar{N}_{g(1)} = 39.6$, 则机器需作 $T_{s(1)}^* = 1.14048 \times 10^8$ 次运算。若不计初始场, 则取 $\bar{N}_{g(2)} = 32.5$, 机器需作 $T_{s(2)}^* = 7.488 \times 10^7$ 次运算, 由(20)~(22)式分别求得 $DT_{s(1)}^* = 2.89152 \times 10^8$ 次, $K_{s(1)}^* = 0.283$, $E_{s(1)}^* = 71.7\%$ 。这就是说, 在所述情况下, FASAT方案的计算量仅及原始方案的0.28, 节省72%, 达2.89亿次, 约87分钟的运算。

c) 用OFASAT方案计算, 按(25)式, $T_{s(1)}^{**} = 1.99584 \times 10^7$ 次。由(26)~(29)式分别求得 $DT_{s(1)}^{**} = 3.83241 \times 10^8$ 次, $K_{s(1)}^{**} = 0.175$, $K_{s(1)}^* = 0.0495$, $E_{s(1)}^* = 95\%$ 。可见在相同情况下, OFASAT方案的计算量仅及原始方案5%。这相当于在FASAT方案的基础上再削减82.5%, 使计算量相对节省达95.0%, 相当于3.83亿次, 约116分钟的运算。

d) 初步试验结果

对上述区域气象场在DJS—220机上作了试验计算。按原始方案, 包括形成初始场和各种其它订正计算, 共需2小时15分钟。按FASAT方案, 不计初始场, 但包括各种

其它订正计算, 仅需约15分钟, 图3.a 为1983年7月4日08时500mb等压面上流场的客观分析结果。图3.b为同一时次同一等压面上的流场实况。由图可见, 计算结果和实况是基本一致的, 特别是主要的气旋性中心和反气旋性中心几乎一一对应。

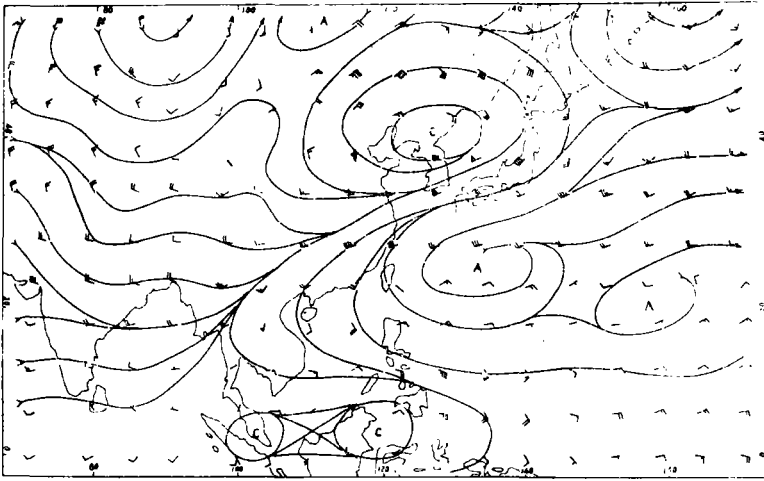


图3 a. 1983年7月4日08时500mb流场的客观分析结果

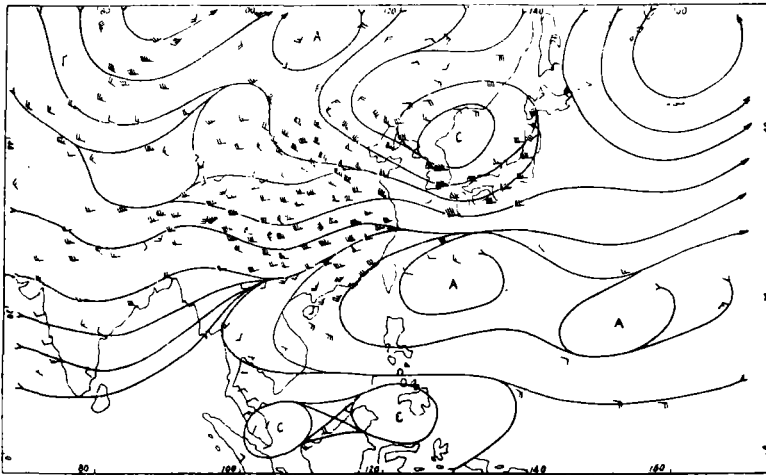


图3 b. 1983年7月4日08时500mb流场实况

2. 对全球气象场的客观分析

现在讨论用FASAT方案和OFASAT方案对全球气象场作客观分析的一个假设例。设将全球按每2.5个经纬距作网格, 则总共有 144×72 个网格点。若需分析 U 、 V 、 T 和 Td 四个要素。每个要素各有 $P=1000, 850, 700, 500, 300, 200, 100\text{mb}$ 等共7层等压面。设收集到比较均匀地分布于全球的1000个测站的数据, 又取扫描半径为 $R=4, 3, 2, 1$ 个网格距, 且需形成初始场。

a) 用原始方案计算。按(14)式, 可求得 $T_{s(1)} = 1.45152 \times 10^{10}$ 次运算。

b) 按照FASAT方案计算。由(16)式, 取 $\bar{N}_{g(1)} = 39.6$, 则 $T_{s(1)}^* = 4.4352 \times 10^8$ 次运算。再由(20)~(22)式可分别求得: $DT_{s(1)}^* = 1.40716 \times 10^{10}$ 次, $K_{s(1)}^* = 0.031$,

$E_{s(1)}^* = 96.9\%$ 。这表明, 利用FASAT方案对全球气象场作客观分析, 其计算量可比原始方案节省96.9%, 达140.7亿次的运算。当用不同速度的计算机运算时, 所需时间如表1。由表1可见, 要用原始方案作逐日全球气象场的客观分析, 没有速度达20〔万次/秒〕的计算机是困难的。即使是100〔万次/秒〕的计算机, 也需4小时。但FASAT方案几乎对任何(假定它们都有足够容量)的计算机都能适应。即使速度仅有5〔万次/秒〕的计算机也仅需2.5小时。由此, 我们可看出这种快速算法的实用价值。

表1 各种速度计算机作全球客观分析需时估计

速度 [万次/秒]	原始方案		FASAT方案		节省量	
	$T_{s(1)} = 1.45152 \times 10^{10}$ 次		$T_{s(1)}^* = 4.4352 \times 10^8$ 次		$DT_{s(1)}^* = 1.40716 \times 10^{10}$ 次	
	分	小时	分	小时	分	小时
5	4838.4	80.64	147.8	2.46	4690.5	78.18
10	2419.2	40.32	73.9	1.23	2345.3	39.09
20	1209.6	20.16	37.0	0.62	1172.6	19.54
50	483.8	8.06	14.8	0.25	469.5	7.82
100	241.9	4.03	7.4	0.12	234.5	3.91

c) 按照OFASAT方案计算。由(25)~(29)式可分别算出 $T_{s(1)}^{**} = 5.8608 \times 10^7$ 次运算, $DT_{s(1)}^{**} = 1.44565 \times 10^{10}$ 次, $K_{s(1)}^{**} = 0.132$, $K_{s(1)}^* = 0.00403$, $E_{s(1)}^{**} = 99.6\%$ 。可见利用OFASAT方案作全球客观分析, 又可将FASAT方案所需计算量再削减86.8%, 从而使实际计算量仅及原始方案的0.40%, 约0.6亿次。节省出99.6%, 达144.6亿次的运算量。

参 考 文 献

- [1] Bergthorson, P., B.R.Döös, *Tellus*, 7 (1955), 3, 320-340.
[2] Cressman, G. P., *Monthly Weather Review*, 87 (1959), 367-374.
[3] Ramanathan, Y., Padma Kulkani, D.R. Sikka, *J. Appl. Meteor*, 12 (1973),
5, 977-983.
[4] Bertoni, E., T.A. Lund, *J. Appl. Meteor*, 2 (1963), 539-545.
[5] 马吉溥, 大气科学, 5 (1981), 1, 43-49.

**Rapid Computational Algorithm of Successive Approximation
for Objective Analysis of Meteorological Fields**

Chen Chuangmai

Abstract

This paper presents two time-saving computational schemes of Cressman's successive approximation for objective analysis of meteorological fields. The first scheme (FASAT) sets up a computing system taking an observational station as the scanning centre. The computation volume saved may amount to 55-80%. The second scheme (OFASAT) sets up another computing system which scans all the meteorological variables on each level within a specified radius centred at a station. The saved computational volume attains to 75-99%.