

最佳客觀分析法之模擬實驗

林松錦

周思運

摘要

本實驗以實際個案利用數學模擬方式來推訂最佳內插法中相關函數的係數，同時討論內插時所使用測站數對分析結果的影響。實驗結果顯示，一般探空測站的分布密度而言，適當的測站數約為 8 個，但在測站密度較稀或分佈不均時應增加測站數以避免分析結果的偏差。實驗結果也發現相關函數的選訂對分析結果的影響較測站數對分析結果的影響為顯著。

一、前言

客觀分析是以數值方法利用電子計算機將不規則分佈的氣象觀測資料內插到規則分佈的網格點上。將觀測資料內插到網格點上是數值預報必須的步驟，在天氣的診斷分析上也相當有用。隨著計算機的發展，客觀分析在氣象研究及日常作業上愈顯出其功用及重要。就實用上而言，目前較常使用的客觀分析方法有逐次修正法 (successive correction method, Cressman, 1959; Barnes, 1964, 1973) 及最佳內插法 (optimum interpolation method, Gandin, 1963; Belousov 等, 1968)。最佳內插法理論上要較逐次修正法為複雜且費時，但最佳內插法却有逐次修正法所沒有的優點。基本上這二種方法都是先在網格點上假定一初始場，然後利用網格點附近測站的觀測資料來修正初始場。逐次修正法中測站對網格點的權重函數 (weighting function) 是人為所訂的。但在最佳內插法中權重係數 (weighting coefficient) 是根據大氣中氣象場的統計結構來推算的，理論上是較客觀。為期使分析結果較能符合實際大氣，客觀分析時除使用正常的地面及高空探空觀測資料外，常使用一些輔助的資料如氣象衛星資料，飛機報告。尤其是在測站稀少地區或資料

缺失時。對於這些資料來源不同所造成品質不一的問題，最佳內插法有較好，較容易的處理 (Alaka 及 Elvander, 1972)。在測站分佈不均勻的區域，逐次修正法若不加入一測站密度的權重，所分析的结果和實際情形會有偏差。最佳內插法本身却具有隨測站密度而調節權重的功能。因此儘管最佳內插法在計算上較為費時，但因為上述的優點許多先進的氣象中心都已採用最佳內插法來做氣象資料的客觀分析。

在整個最佳內插法的過程中，以利用以往長時期的觀測資料經由統計方法來決定適當的相關函數及計算權重係數時解相關函數矩陣最為費時。當然，相關函數一經決定之後便可長時間地使用。不過由以前的研究顯示，氣象場的結構大致能符合水平均質 (homogeneous) 及均向 (isotropic) 的假設 (Bertoni 及 Lund, 1963)。在這個假設之下氣象場之間的相關函數 (correlation) 僅是距離的函數，因此相關函數似乎可以利用簡單的數學模擬來決定 (Schlatter 等, 1976)。在計算測站對網格點的權重係數時，對每一個網格點須要解一個 $n \times n$ 階相關函數矩陣的反矩陣， n 是網格點附近參與修正的測站數。由於解一個矩陣的計算機時間並不是隨矩陣的階數以線性增加，而是大於線性的增加。因此適當地選擇測站數是控制計算效率

的最大因素。測站數過少會影響分析的结果，但從最佳內插法的特性來看，增加太多的測站數也不見得會使分析的品质有顯著的提高，而只是浪費計算時間。本文將就相關函數及測站數的選訂對分析結果的影響以實際的觀測資料來做一些實驗。

二、最佳內插法的基本理論

最佳內插法及逐次修正法基本上都是先在網格點上設定一初始猜測值 (first guess)，然後在網格點附近選擇一些較近的測站，利用測站上的觀測值來修正猜測值。網格點上的觀測值修正量則是測站觀測值和猜測值之間差值的權重平均，若以 φ 代表一氣象變數則可以下式來表示：

$$\tilde{\varphi}_0 = \bar{\varphi}_0 + \sum_{i=1}^n \alpha_i (\varphi_i - \bar{\varphi}_0) \quad (1)$$

其中 $\tilde{\varphi}_0$ 為經修正後的網格點分析值， φ_i 為測站 i 上的觀測值， $\bar{\varphi}_0$ 及 $\bar{\varphi}_i$ 分別為在網格點及測站 i 上的猜測值， α_i 為測站 i 對網格點的權重係數， n 為測站數。測站上的猜測值可以由網格上的猜測值內插得到。假定 φ_0 為網格點上的真正觀測值 (例如把一測站當做網格點，利用附近測站的觀測值來算該測站的分析值)，則最佳內插法的分析誤差 (即 $(\tilde{\varphi}_0 - \varphi_0)$ 的均方根 (root-mean-square) 來表示，即

$$E = \frac{(\tilde{\varphi}_0 - \varphi_0)^2}{n} = \frac{[(\tilde{\varphi}_0 - \bar{\varphi}_0) - (\bar{\varphi}_0 - \varphi_0)]^2}{n} = \frac{(\tilde{\varphi}_0' - \varphi_0')^2}{n} = \frac{(\sum_{i=1}^n \alpha_i \varphi_i' - \varphi_0')^2}{n} \quad (2)$$

其中 $\varphi_i' = \varphi_i - \bar{\varphi}_0$ 。(1) 式中的權重係數是在要求分析誤差為極小值的條件下計算得到的，即

$$\frac{\partial E}{\partial \alpha_i} = 0 \quad (i = 1, 2, \dots, n) \quad (3)$$

將 (2) 式代入 (3) 式可導致

$$\sum_{j=1}^n \alpha_j \overline{\varphi_i' \varphi_j'} = \overline{\varphi_i' \varphi_0'} \quad (i = 1, 2, \dots, n) \quad (4)$$

(4) 式中 $\overline{\varphi_i' \varphi_j'}$ 及 $\overline{\varphi_i' \varphi_0'}$ 分別為二測站之間及測站和網格點之間氣象變數 φ 的互變異數 (covariance)。此時若將 φ_i 認為是在測站 i 上變數 φ 的長時間氣候平均值，則 φ_i' 為瞬時觀測值和氣候平均值的差。從氣象變數的統計結構顯示，互變異數在水平空間符合均質及均向性所涵蓋的範圍太小。在實際應用上可假設氣象變數的變異數 (variance) 是均質的，而將 (4) 式中的互變異數標準化 (normalized) 以相關數表示

$$\sum_{j=1}^n \alpha_j \mu_{ij} = \mu_{i0} \quad (i = 1, 2, \dots, n) \quad (5)$$

其中 $\mu_{ij} = \overline{\varphi_i' \varphi_j'} / \sqrt{\overline{\varphi_i'^2} \overline{\varphi_j'^2}}$ 為二測站之間氣象變數 φ 的相關數。Bertoni 及 Lund (1963) 的研究結果指出大氣中氣象變數的相關數具有水平均質及均向性，因此可以假設氣象變數的相關數只是測站之間距離的函數。只要知道相關函數，(1) 式中的權重係數便可由 (5) 式計算得到。相同地，(2) 式也可以相關數來表示

$$E = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mu_{ij} - 2 \sum_{i=1}^n \alpha_i \mu_{i0} + 1 \quad (6)$$

將 (5) 式代入 (6) 式可得

$$E = 1 - \sum_{i=1}^n \alpha_i \mu_{i0} \quad (7)$$

三、個案資料及計算方式

本文以 1983 年 5 月 14 日 1200 GMT 在中國東北地區的 500 mb 高度場做為分析實驗的個案。時間與地點的選擇並無特殊的天氣特性，只是這個地區測站分佈尚稱均勻且剛好有一低壓中心經過。客觀分析的區域是東經 $110^\circ \sim 130^\circ$ ，北緯 $40^\circ \sim 55^\circ$ 範圍內。為避免邊界的問題，參與分析的測站範圍向四方各擴展 10 個經緯度包括東經 $100^\circ \sim 140^\circ$ ，北緯 $30^\circ \sim 65^\circ$ 。這個區域內共有 66 個測站，測站的平均距離約為 400 公里。測站的分佈及當時日本氣象廳主觀分析的 500 mb 高度場如圖 1 所示，低氣壓的中心高度約略低於 5340 公尺。

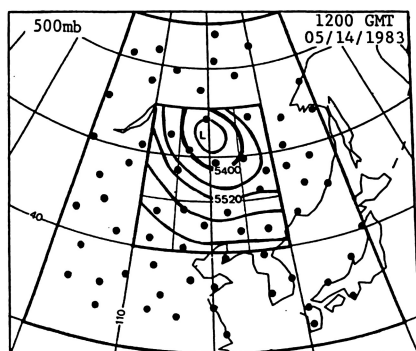


圖 1：測站分佈及日本氣象廳的主觀分析，等高線的間距為 60 公尺。

分析區域內的網格點是直接以經緯度線的交點做為網格點位置。網格點的間距是 1 度，因此共有 21×16 個網格點。因為是經緯度網格點也就不必考慮地圖投影的問題。二點之間的實際距離（ r ）則由二點位置的經緯度利用下式計算

$$r = [a^2(\lambda_1 - \lambda_2)^2 \cos^2(\frac{\phi_1 + \phi_2}{2}) + a^2(\phi_1 - \phi_2)^2]^{1/2} \quad (8)$$

其中 a 為地球半徑， λ 為經度， ϕ 為緯度。最佳內插法所使用的初始猜測場通常是氣候平均場或是預報場，在這裡為方便起見 500mb 高度場的初始猜測值是採在經向為均勻而只隨緯度變化的函數，若 z 表初始高度場，則

$$z = 5700 - 20(\phi - 40) \quad (9)$$

也就是說初始高度場的等值線是平行於緯度線。解權重係數的聯立方程式是以高斯消去法求相關數矩陣的反矩陣。

四、實驗與結果討論

前面已經提過在均質及均向的假設之下，相關函數只是測站之間距離的函數，距離愈近相關數愈大，愈遠則愈小。在這實驗裡所採用的是高斯形式的相關函數

$$\mu(r) = A \exp(-Br^2) \quad (10)$$

其中 A 及 B 為待決定的係數。係數 A 代表的是當二個測站非常接近時（ $r \rightarrow 0$ ）的相關數，理論上 A 應趨近於 1，但由於實際觀測資料通常都帶有少許的誤差（random error），即使二個測站幾乎重疊所得到的相關數仍不會等於 1，因此在這實驗裡將 A 訂為 0.95 且保持不變。係數 B 是決定相關數隨距離增加而遞減的斜率， B 值愈大相關數遞減得愈快，愈小則愈慢。以綜觀尺度的大氣運動（1000 km ~ 2000 km）為準，我們採用了 5 種不同的 B 值（由 0.62 到 3.10）來做模擬分析。不同 B 值時相關數隨距離的變化如圖 2 所示，其中距離 r

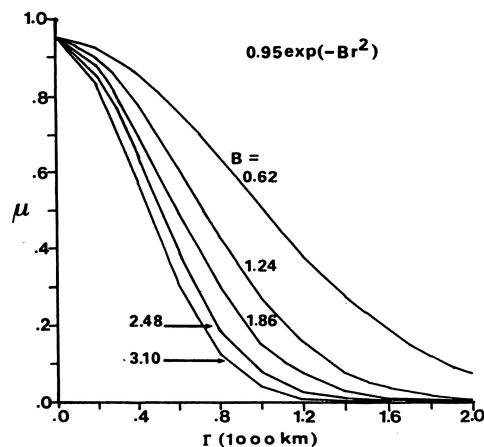


圖 2：不同 B 值時相關函數隨距離的變化

是以 1000 公里做為單位。分析時所使用的測站數不但影響計算的效率同時也會影響分析的結果。為了避免將相關函數及測站數對分析結果的影響混淆不清，實驗中先以固定的測站數來討論不同相關函數對分析結果的影響，由此選定適當的相關函數，再以此相關函數來討論測站數對分析結果的影響。

Meleshko 及 Prigodich (1964) 的結論指出適當的測站數約為 6 ~ 8 個。實驗先固定以 8 個測站利用不同的相關函數來做分析，所得的結果如圖 3。由圖 3 的結果可以看出 B 值愈小高度場的等

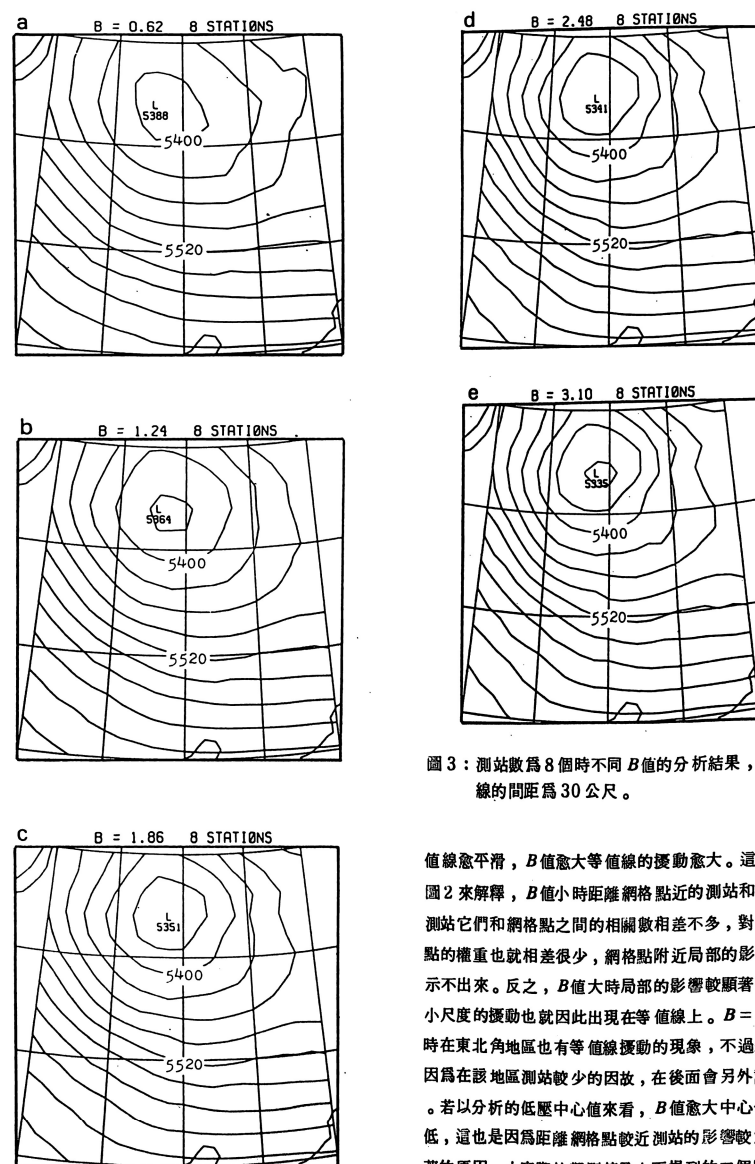


圖 3：測站數為 8 個時不同 B 值的分析結果，等高線的間距為 30 公尺。

值線愈平滑， B 值愈大等值線的擾動愈大。這可由圖 2 來解釋， B 值小時距離網格點近的測站和遠的測站它們和網格點之間的相關數相差不多，對網格點的權重也就相差很少，網格點附近局部的影響顯示不出來。反之， B 值大時局部的影響較顯著，較小尺度的擾動也就因此出現在等值線上。 $B = 0.62$ 時在東北角地區也有等值線擾動的現象，不過這是因為在該地區測站較少的因故，在後面會另外討論。若以分析的低壓中心值來看， B 值愈大中心值愈低，這也是因為距離網格點較近測站的影響較為顯著的原因。由實際的觀測值及上面提到的二個因素

來考慮，就本個案實驗而言(1)式中的係數 B 應取 1.86 較為適當。雖然 $B = 2.48$ 時低壓中心的值 (5341 公尺) 和實際觀測值相當接近，但就客觀分析的性質來說並不見得很理想，因為資料內涵本身更含有些微的修勻作用，因此較實際低壓中心值稍高一些應屬正常的。此外，在同一地區對不同的個案 $B = 2.48$ 有可能得到低於實際觀測值的情形。

圖 4 是 $B = 1.86$ 時使用不同的測站數所得到的分析圖。從這一系列的分析圖可以看出，不論對等值線的平滑或低壓中心值，測站數愈多分析的品質愈好。測站太少時(如圖 4.a 五個測站)等值線的擾動較明顯且低壓中心的位置也稍有偏差。本實驗個案區域內測站的密度及分佈尚稱均勻，測站太少時所造成的影響不是很明顯。如果測站較稀疏而

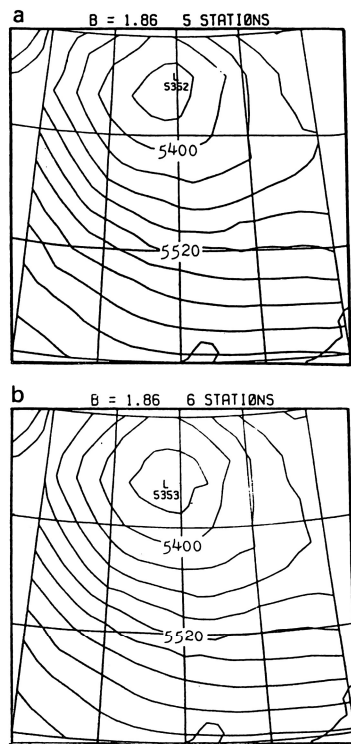


圖 4： $B = 1.86$ 時不同測站數的分析結果，等高線的間距為 30 公尺。

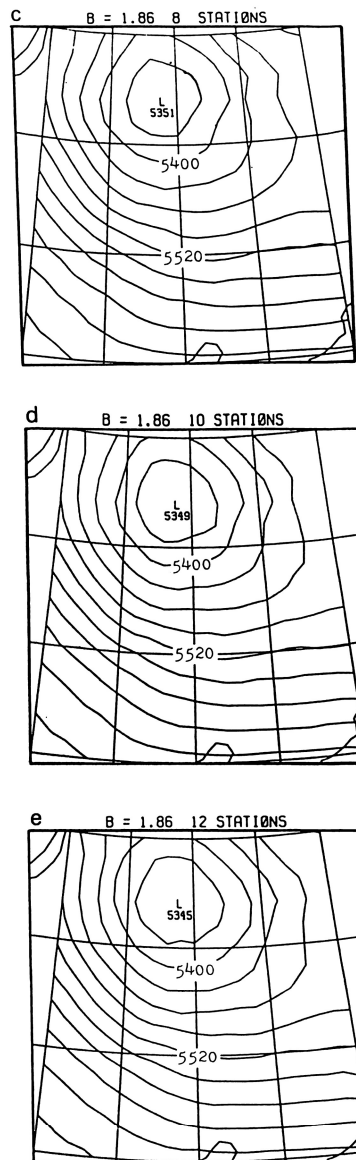


圖 5：不同 B 值時分析誤差隨測站數的變化

使用的測站數又少，較近的測站可能偏於某一個方向因而造成分析結果偏離實際觀測。如圖 4 內的東北角因為測站較少，只使用 5 個測站時便造成低壓中心偏離的情形，測站數增加後中心的位置便回到正確的位置。由(7)式及圖 5 顯示，分析誤差隨著測站數的增加而減少。不過從圖 4 的結果來看，測站數從 8 增加到 10 甚至 12 所得到的結果除了在東北角測站稀少區稍好外，對於其他地區並沒有很顯著的改變。

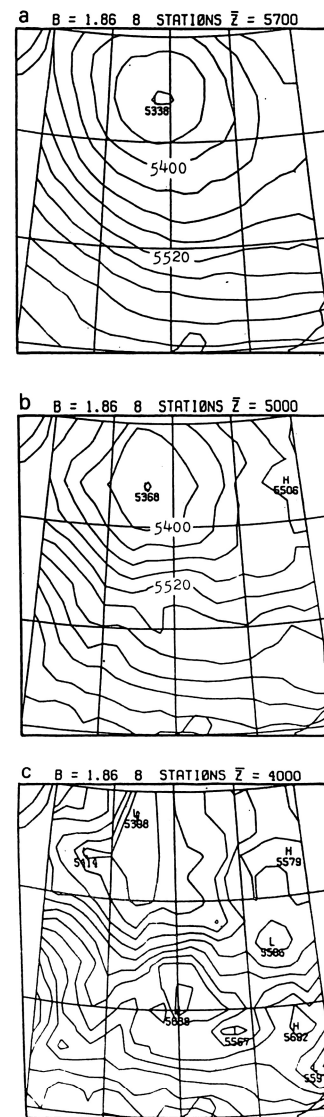


圖 6： $B = 1.86$ ，測站數為 8 個時使用不同的常數初始猜測場的分析結果。

前面提過，最佳內插法一般是以氣候平均場或預報場來做為初始猜測場。本個案中 500 mb 高度場的等值線約略平行於緯度線(中、高對流層綜觀尺度的高度場大致都有此趨勢)，因此實驗中假設初始高度場只是緯度的函數(9)式。不過這種假設的初始猜測場對於其他的氣象場或較小尺度運動的分析有時會造成錯誤或較差的分析結果。初始猜測場對分析結果的影響可由圖 6 的分析看出。圖 6 中假設三種常數的初始猜測場 ($\bar{z} = 5700$ 公尺，5000 公尺，4000 公尺)，當常數初始場和實際的觀測值相差不多時 ($\bar{z} = 5700$ 公尺)，分析的結果還大致不離譜，但當 $\bar{z} = 4000$ 公尺時所得到的結果便完全無法辨認了。這個原因可以由(1)式看出，(1)式中測站對網格點的權重的總和，理論上應

等於 $1 (\sum_{i=1}^n \alpha_i = 1)$ ，但實際計算的結果在每個網格點上都稍有些微的誤差，有些網格點上約略大於 1，而有些則約略小於 1，如果 $(\varphi_i - \bar{\varphi}_i)$ 的値又很大的話，整個分析結果便會顯得零亂不堪。因此如果要採用常數的初始場，應採用接近於觀測值的常數（如觀測值的區域平均），而且應將(1)式中的權重係數標準化，亦即將(1)式改成

$$\bar{\varphi}_s = \bar{\varphi}_0 + \frac{\sum_{i=1}^n \alpha_i (\varphi_i - \bar{\varphi}_i)}{\sum_{i=1}^n \alpha_i} \dots\dots\dots(1)$$

由以上的結果大致可以得到如下的結論：

(一)以(10)式高斯形式的相關函數來做綜觀尺度區域性的最佳內插，式中適當的 B 值約在 1.86 左右。 B 值太小會造成中心值過度的平滑，而 B 值太大又會造成過深的中心值。

(二)就一般探空測站分佈的密度而言，在計算經濟的考慮之下，使用的測站數約為 8 個。

(三)若測站密度較稀或分佈不均時，應增加測站數以避免參與分析的測站偏向某一方向造成分析結果的偏差。

四由圖 3 及圖 4 的結果顯示，相關函數對分析結果的影響較測站數對分析結果的影響為顯著。事實上，相關函數本身便已決定測站的影響半徑。如圖 2 中 $B = 1.86$ 時，距離網格點 1200 公里以外的測站其相關數均小於 0.1，這些測站對網格點的權重已相當小，對網格點上的分析值幾乎沒有影響。因此當增加測站時亦應注意所加入的測站是否在影響半徑之內，否則只有徒增計算時間而已。

致 謝

本實驗承中央大學電算中心及遙測中心提供計算機的使用，在此一併致謝。

參考文獻

- Alaka, M. A., and R. C. Elvander, 1972: Optimum interpolation from observations of mixed quality. Mon. Wea. Rev., 100, 612-624

- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. J. Appl. Meteor., 3, 396-409.
- _____, 1973: Mesoscale objective map analysis using weighted time-series observations. NOAA Tech. Memo., ERL NSSL-62, 60pp.
- Belousov, S. L., L. S. Gandin, and S. A. Mashkovich, 1968: Computer processing of meteorological data. 210pp. Translated from Russian, 1971.
- Bertoni, E., and I. A. Lund, 1963: Space correlations of the height of constant pressure surfaces. J. Appl. Meteor., 2, 539-545.
- Cressman, G. P., 1959: An operational objective analysis system. Mon. Wea. Rev., 87, 367-374.
- Gandin, L. S., 1963: Objective analysis of meteorological fields. 242pp. Translated from Russian, 1965.
- Meleshko, V. P., and A. E. Prigodich, 1964: Ob'ektivnyi analiz vlazhnosti i temperatury (Objective analysis of humidity and temperature). Trudy simpoziuma po chislennym metodam prognoza. Leningrad, Gidrometeoizdat.
- Schlatter, T. W., G. W. Branstator, and L. G. Thiel, 1976: Testing a global multivariate statistical objective analysis scheme with observed data. Mon. Wea. Rev., 104, 765-783.

Some Experiments on the Optimum Interpolation Method

Song-Chin Lin Shin-Yun Chou

ABSTRACT

This experiment uses observed data to determine the coefficients of the correlation function, which is required in the optimum interpolation method, through mathematical simulations. The impacts of the number of stations on the objective analysis are also discussed. Experimental results indicate that, for the average distribution density of the rawinsonde stations, the appropriate number of station is 8, but this number should be increased in the region where the station density is sparse or stations are not evenly distributed. Results also depict that the impacts of the selected correlation function on the objective analysis are more significant than those caused by the number of stations.