

小時風速頻率分佈的曲線擬合——極大似然法

王金慶

中正理工學院應用物理系

(中華民國七十五年一月三十日收件，二月二十四日修正。)

摘 要

本文以二個參數、三個參數、四個參數對數常態分佈函數以及二個參數、三個參數、四個參數韋伯分佈函數作為不考慮風向的風速頻率理論分佈函數，以極大似然法推估頻率分佈函數的參數值，並以桃園中正國際機場地面風速觀測資料作為個案研究，以五種適合度統計量及 Theil 氏不等式係數作為適合度準則，結果顯示四個參數韋伯分佈函數最適於作為桃園濱海地區全年與冬半年地面風速頻率的理論分佈函數，而夏半年則以四個參數對數常態分佈函數較適合。

一、前 言

近年來有些不是專門研究氣象的學者對於風速頻率的預測有很大的興趣。例如，從大氣擴散理論來研究空氣污染問題的學者，有興趣於其研究地區低風速頻率的預測 (Bencala and Seinfeld, 1976 ; Cats and Holtslag, 1980)，因為低風速的發生頻率可以作為發生空氣污染事件的指標之一；從事建築物結構設計的工程師，需要建築物所在地高風速發生頻率的預測資料，以作為設計上的參考 (Kwok, 1977 ; Parkinson, 1971)；石油危機發生以後，風能發電的想法應運而生，風速太小，不足以發電，風速太大，發電機件容易損壞，而且風能與風速的立方成正比，因此欲估算風能場址的發電潛能，需要知道介於切入 (cut in) 與切離 (cut out) 風速之間的風速頻率 (Harder, 1977 ; Hemessey, 1977 ; Hewson, 1975 ; 簡等, 1983)。

統計學上最常用也是最重要的頻率分佈是常態

分佈 (normal distribution)，常態分佈的圖形是對稱的圖形，然而風速頻率分佈通常是正偏 (positive skew) 的圖形 (圖 1)，即風速最小值為零，零風速 (靜風) 的頻率不為零，風速頻率先是隨風速增大而迅速增加，然後隨風速增大而緩慢減小。若用常態分佈作為風速頻率的理論分佈，效果將會很差。因此許多研究者以較具彈性的韋伯分佈 (Weibull distribution) 或對數常態分佈 (lognormal distribution) 作為風速頻率的理論分佈 (Bencala and Seinfeld, 1976 ; Luna and Church, 1974 ; Stewart and Essenwanger, 1978 ; Takle and Brown, 1978 ; 吳, 1982)。對數常態分佈與韋伯分佈預測靜風的頻率為零，然而靜風的頻率通常不為零，為了滿足零風速的頻率不為零的可能性，需要採用三個參數的對數常態分佈或三個參數的韋伯分佈 (Essenwanger, 1976 ; Mage, 1980)。此外，最大風速不能大於音速，若將此條件亦納入考慮，則需要更具適應性的四個參數的對數常態分佈或四個參數的韋伯分佈 (Ess-

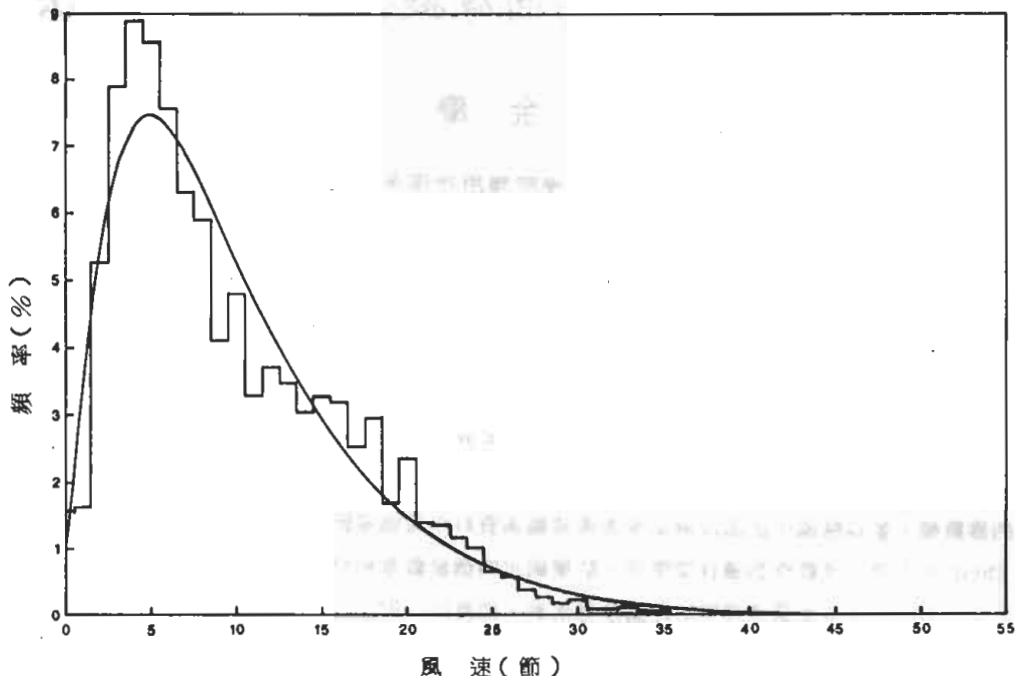


圖 1：桃園濱海地區夏半年地面風速頻率分佈的直方圖及其四個參數對數常態擬合頻率分佈圖。

enwanger, 1976 ; Mage, 1980)。對數常態分佈與韋伯分佈在本文以下分別稱為二個參數的對數常態分佈與二個參數的韋伯分佈。

以往統計模式參數的推估，大都採用數矩法 (method of moment) 與線性最小二乘方法，也有些是由中位數 (median) 和四分位數 (quartiles) 或由平均數和標準差來推測。二個參數韋伯分佈、二個參數及三個參數對數常態分佈函數的參數可以很容易地用上述的方法來推估，對於用上述方法較難推估的三個參數及四個參數韋伯分佈、四個參數對數常態分佈的參數可用極大似然法 (maximum likelihood method) 來推估 (Essenwanger, 1976 ; Justus et al., 1978)。Take 與 Brown (1978) 指出統計模式參數的推估以極大似然法所求得的最為正確。Mage (1980) 以三個參數和四個參數對數常態分佈來模擬風速頻率分佈，Holland 和 Fitz-Simons (1982) 以常態分佈、三個參數對數常態分佈、三個參數伽瑪分佈 (gamma distribution)、三個參數韋伯分佈、四個參數對

數常態分佈和四個參數貝他分佈 (beta distribution) 來模擬臭氧的濃度分佈，他們都用極大似然法來推估統計模式的參數而且效果良好。

本文以二個參數、三個參數、四個參數的對數常態分佈以及二個參數、三個參數、四個參數的韋伯分佈等六個統計模式作為不考慮風向的地面風速頻率的理論分佈，並以桃園中正國際機場民國 68 年至 72 年的地面逐時正點風速觀測資料作為個案研究，以六個適合良度準則來比較六個統計模式的擬合良度。

二、研究方法

假設 X 是連續的隨機變數，它的或然率密度函數 f_x 是由一些待決定的未知參數 $\theta_1, \theta_2, \dots, \theta_k$ 等所決定。對於 X 的任一隨機樣本，若其 N 個觀測值 x_1, x_2, \dots, x_N 未被分組，則似然函數 (likelihood function) 定義為

$$L(\theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^N f(x_i; \theta_1, \theta_2, \dots, \theta_k) \quad (1)$$

若其觀測值被分組，則似然函數定義為

$$L(\theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^m P_i^{n_i} \quad (2)$$

其中， P_i 是一個觀測值落在第 i 組內的或然率， m 為觀測值被分成的組數， n_i 是觀測值落在第 i 組內的次數，而且 $\sum_{i=1}^m n_i = N$ 。參數 $\theta_1, \theta_2, \dots, \theta_k$ 的估計值 $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ 是令似然函數有極值而求得，即

$$\frac{\partial}{\partial \theta_j} L(\theta_1, \theta_2, \dots, \theta_k) = 0, \quad j = 1, 2, \dots, k \quad (3)$$

由於很多或然率函數含有指數項 (exponential terms)，因此，通常較方便的作法是先將似然函數取自然對數成為對數似然函數 (log likelihood function)，即

$$\ln L(\theta_1, \theta_2, \dots, \theta_k) = \sum_{i=1}^m n_i \ln P_i \quad (4)$$

再令對數似然函數有極值，以求得參數的估計值，即

$$\frac{\partial}{\partial \theta_j} \ln L(\theta_1, \theta_2, \dots, \theta_k) = 0, \quad j = 1, 2, \dots, k \quad (5)$$

以極大似然法所求得參數的估計值稱為該參數的極大似然估計值 (maximum likelihood estimates)。雖然對於小樣本，極大似然估計值未必是不偏倚的估計值 (unbiased estimate)，而且其變異值 (variance) 也未必是極小值，但是對於大樣本，極大似然估計值幾乎是不偏倚的，而且其變異值也是極小的。當樣本的大小無限制的增加，極大似然估計值 $\hat{\theta}_i$ 會收斂而趨近於真正值 θ_i 。

三、統計模式

1 兩個參數的對數常態分佈

隨機變數 X (在本文為風速)，若 $Y = \ln X$ ，而且 Y 為常態分佈，則 X 為對數常態分佈。兩個參數的對數常態分佈其或然率密度函數為

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \quad (6)$$

$x > 0, -\infty < \mu < \infty, \sigma > 0$

其中參數 μ 與 σ 分別為隨機變數 Y 的平均值與標準偏差，其極大似然估計值 $\hat{\mu}$ 與 $\hat{\sigma}$ 分別為

$$\hat{\mu} = N^{-1} \sum_{i=1}^m n_i \ln x_i \quad (7)$$

$$\hat{\sigma}^2 = N^{-1} \sum_{i=1}^m n_i (\ln x_i)^2 - \hat{\mu}^2 \quad (8)$$

上式中 N 為觀測總次數。將 x 大於零的觀測資料代入(7)、(8)兩式即可求得 $\hat{\mu}$ 與 $\hat{\sigma}$ 。

2 三個參數的對數常態分佈

三個參數的對數常態分佈其或然率密度函數為

$$f(x; \mu, \sigma, \gamma) = \frac{1}{\sqrt{2\pi} \sigma (x - \gamma)} \exp\left\{-\frac{[\ln(x - \gamma) - \mu]^2}{2\sigma^2}\right\}, -\infty < \gamma < x \quad (9)$$

其中參數 γ 必須小於 x ，其極大似然估計值 $\hat{\gamma}$ 滿足方程式

$$\begin{aligned} & \sum_{i=1}^m \frac{n_i}{x_i - \hat{\gamma}} \ln(x_i - \hat{\gamma}) \\ & + \{N^{-1} \sum_{i=1}^m n_i [\ln(x_i - \hat{\gamma})]^2 \\ & - [N^{-1} \sum_{i=1}^m n_i \ln(x_i - \hat{\gamma})]^2 \\ & - N^{-1} \sum_{i=1}^m n_i \ln(x_i - \hat{\gamma})\} \sum_{i=1}^m \frac{n_i}{x_i - \hat{\gamma}} = 0 \quad (10) \end{aligned}$$

而參數 μ 與 σ 的極大似然估計值 $\hat{\mu}$ 與 $\hat{\sigma}$ 分別為

$$\hat{\mu} = N^{-1} \sum_{i=1}^m n_i \ln(x_i - \hat{\gamma}) \quad (11)$$

$$\hat{\sigma}^2 = N^{-1} \sum_{i=1}^m n_i [\ln(x_i - \hat{\gamma})]^2 - \hat{\mu}^2 \quad (12)$$

由(10)式可求得 $\hat{\gamma}$ ，然後將之代入(11)與(12)式即可得 $\hat{\mu}$ 與 $\hat{\sigma}$ 。

3 四個參數的對數常態分佈

四個參數的對數常態分佈其或然率密度函數為

$$f(x; \mu, \sigma, \gamma, \theta) = \frac{\theta - \gamma}{\sqrt{2\pi} \sigma (x - \gamma) (\theta - x)}$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[\ln \left(\frac{x-\gamma}{\theta-x} \right) - \mu \right]^2 \right\}$$

$$-\infty < \gamma < x < \theta < \infty \quad (13)$$

其中參數 θ 必須大於 x ，其極大似然估計值 $\hat{\theta}$ 滿足方程式

$$\sum_{i=1}^m \frac{n_i}{\hat{\theta}-x_i} \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right)$$

$$+ \left\{ N^{-1} \sum_{i=1}^m n_i \left[\ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right]^2 \right.$$

$$\left. - \left[N^{-1} \sum_{i=1}^m n_i \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right]^2 \right\}$$

$$\times \left[\frac{N}{\hat{\theta}-\hat{\gamma}} - \sum_{i=1}^m \frac{n_i}{\hat{\theta}-x_i} \right]$$

$$- \left[N^{-1} \sum_{i=1}^m n_i \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right] \sum_{i=1}^m \frac{n_i}{\hat{\theta}-x_i}$$

$$= 0 \quad (14)$$

參數 γ 的極大似然估計值 $\hat{\gamma}$ 滿足方程式

$$\sum_{i=1}^m \frac{n_i}{x_i-\hat{\gamma}} \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right)$$

$$- \left\{ N^{-1} \sum_{i=1}^m n_i \left[\ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right]^2 \right.$$

$$\left. - \left[N^{-1} \sum_{i=1}^m n_i \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right]^2 \right\}$$

$$\times \left[\frac{N}{\hat{\theta}-\hat{\gamma}} - \sum_{i=1}^m \frac{n_i}{x_i-\hat{\gamma}} \right]$$

$$- \left[N^{-1} \sum_{i=1}^m n_i \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right] \sum_{i=1}^m \frac{n_i}{x_i-\hat{\gamma}}$$

$$= 0 \quad (15)$$

而參數 μ 與 σ 的極大似然估計值 $\hat{\mu}$ 與 $\hat{\sigma}$ 分別為

$$\hat{\mu} = N^{-1} \sum_{i=1}^m n_i \ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \quad (16)$$

$$\hat{\sigma}^2 = N^{-1} \sum_{i=1}^m n_i \left[\ln \left(\frac{x_i-\hat{\gamma}}{\hat{\theta}-x_i} \right) \right]^2 - \hat{\mu}^2 \quad (17)$$

給(14)式中的 $\hat{\gamma}$ 一個猜測值 $\hat{\gamma}_0$ ，即令(14)式中的 $\hat{\gamma} = \hat{\gamma}_0$ 。以求得 $\hat{\theta}$ 的暫時值 $\hat{\theta}_1$ ，令(15)式中的 $\hat{\theta} = \hat{\theta}_1$ 求 $\hat{\gamma}$ 的暫時值 $\hat{\gamma}_1$ ，然後令(14)式中的 $\hat{\gamma} = \hat{\gamma}_1$ 以求得 $\hat{\theta}$ 的新暫時值 $\hat{\theta}_2$ ，如此反覆計算直到 $\hat{\gamma}_k$ 與 $\hat{\theta}_k$ 同時都能滿足(14)與(15)兩式為止。最後將 $\hat{\gamma}_k$ 與 $\hat{\theta}_k$ 代入(16)與(17)式即可

得 $\hat{\mu}$ 與 $\hat{\sigma}$ 。上述的反覆計算過程是為疊代法 (iterative process)。

4. 兩個參數的韋伯分佈

兩個參數的韋伯分佈其或然率密度函數為

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha} \right)^{\beta-1} \exp \left[-\left(\frac{x}{\alpha} \right)^\beta \right],$$

$$x \geq 0, \alpha > 0, \beta > 0 \quad (18)$$

其中參數 β 為形狀因子 (shape factor) 是無因次的量，其極大似然估計值 $\hat{\beta}$ 滿足方程式

$$\hat{\beta}^{-1} + N^{-1} \sum_{i=1}^m n_i \ln x_i$$

$$- \left(\sum_{i=1}^m n_i x_i^{\hat{\beta}} \ln x_i \right) \left(\sum_{i=1}^m n_i x_i^{\hat{\beta}} \right)^{-1}$$

$$= 0 \quad (19)$$

參數 α 為尺度因子 (scale factor)，其單位與 X 相同，數值等於累積頻率為 63.2% 時的 X 值，其極大似然估計值 $\hat{\alpha}$ 為

$$\hat{\alpha} = \left[N^{-1} \sum_{i=1}^m n_i x_i^{\hat{\beta}} \right]^{1/\hat{\beta}} \quad (20)$$

由(19)式求得 $\hat{\beta}$ ，然後將之代入(20)式求 $\hat{\alpha}$ 。

5. 三個參數的韋伯分佈

三個參數的韋伯分佈其或然率密度函數為

$$f(x; \alpha, \beta, \gamma) = \frac{\beta}{\alpha} \left(\frac{x-\gamma}{\alpha} \right)^{\beta-1}$$

$$\exp \left[-\left(\frac{x-\gamma}{\alpha} \right)^\beta \right], -\infty < \gamma < x \quad (21)$$

其中參數 γ 必須小於 x ，其極大似然估計值 $\hat{\gamma}$ 滿足方程式

$$N^{-1} (\hat{\beta}-1) \sum_{i=1}^m n_i (x_i-\hat{\gamma})$$

$$- \left[\beta \sum_{i=1}^m n_i (x_i-\hat{\gamma})^{\hat{\beta}-1} \right] \left[\sum_{i=1}^m n_i (x_i-\hat{\gamma})^{\hat{\beta}} \right]^{-1}$$

$$= 0 \quad (22)$$

參數 β 的極大似然估計值 $\hat{\beta}$ 滿足方程式

$$\hat{\beta}^{-1} + N^{-1} \sum_{i=1}^m n_i \ln (x_i-\hat{\gamma})$$

$$- \left[\sum_{i=1}^m n_i (x_i-\hat{\gamma})^{\hat{\beta}} \ln (x_i-\hat{\gamma}) \right]$$

$$\left[\sum_{i=1}^m n_i (x_i - \hat{\gamma})^{\hat{\beta}} \right]^{-1} = 0 \quad (23)$$

而參數 α 的極大似然估計值 $\hat{\alpha}$ 為

$$\hat{\alpha} = \left[N^{-1} \sum_{i=1}^m n_i (x_i - \hat{\gamma})^{\hat{\beta}} \right]^{1/\hat{\beta}} \quad (24)$$

用疊代法由(22)與(23)式求 $\hat{\gamma}$ 與 $\hat{\beta}$ ，然後將之代入(24)式求 $\hat{\alpha}$ 。

6. 四個參數的韋伯分佈

四個參數的韋伯分佈其或然率密度函數為

$$f(x; \alpha, \beta, \gamma, \theta) = \frac{\theta - \gamma}{(\theta - x)^2} \frac{\beta}{\alpha} \left(\frac{1}{\alpha} \frac{x - \gamma}{\theta - x} \right)^{\beta - 1} \exp \left[- \left(\frac{1}{\alpha} \frac{x - \gamma}{\theta - x} \right)^\beta \right],$$

$$-\infty < \gamma < x < \theta < \infty \quad (25)$$

其中參數 θ 必須大於 x ，其極大似然估計值 $\hat{\theta}$ 滿足方程式

$$\frac{\hat{\beta} + 1}{N} \sum_{i=1}^m \frac{n_i}{\hat{\theta} - x_i} - \left[\hat{\beta} \sum_{i=1}^m \frac{n_i}{\hat{\theta} - x_i} \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \right] \left[\sum_{i=1}^m n_i \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \right]^{-1} - \frac{1}{\hat{\theta} - \hat{\gamma}} = 0 \quad (26)$$

參數 γ 與 β 的極大似然估計值 $\hat{\gamma}$ 與 $\hat{\beta}$ 分別滿足下列兩個方程式

$$\frac{\hat{\beta} - 1}{N} \sum_{i=1}^m \frac{n_i}{x_i - \hat{\gamma}} - \left[\hat{\beta} \sum_{i=1}^m \frac{n_i}{x_i - \hat{\gamma}} \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \right] \left[\sum_{i=1}^m n_i \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \right]^{-1} + \frac{1}{\hat{\theta} - \hat{\gamma}} = 0 \quad (27)$$

$$\hat{\beta}^{-1} + N^{-1} \sum_{i=1}^m n_i \ln \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right) - \left[\sum_{i=1}^m n_i \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \ln \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right) \right] \left[\sum_{i=1}^m n_i \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \right]^{-1} = 0 \quad (28)$$

而參數 α 的極大似然估計值 $\hat{\alpha}$ 為

$$\hat{\alpha} = \left[N^{-1} \sum_{i=1}^m n_i \left(\frac{x_i - \hat{\gamma}}{\hat{\theta} - x_i} \right)^{\hat{\beta}} \right]^{1/\hat{\beta}} \quad (29)$$

給(28)式中的 $\hat{\beta}$ 與 $\hat{\gamma}$ 各一個猜測值 $\hat{\beta}_0$ 與 $\hat{\gamma}_0$ ，即令(28)式中的 $\hat{\beta} = \hat{\beta}_0$ ， $\hat{\gamma} = \hat{\gamma}_0$ 以求得 $\hat{\theta}$ 的暫時值 $\hat{\theta}_1$ ，接着令(28)式中的 $\hat{\beta} = \hat{\beta}_0$ ， $\hat{\theta} = \hat{\theta}_1$ 以求得 $\hat{\gamma}$ 的暫時值 $\hat{\gamma}_1$ ，令(28)式中的 $\hat{\gamma} = \hat{\gamma}_1$ ， $\hat{\theta} = \hat{\theta}_1$ 以求得 $\hat{\beta}$ 的暫時值 $\hat{\beta}_1$ ，然後令(28)式中的 $\hat{\beta} = \hat{\beta}_1$ ， $\hat{\gamma} = \hat{\gamma}_1$ 以求得 $\hat{\theta}$ 的新暫時值 $\hat{\theta}_2$ ，如此反覆計算直到 $\hat{\beta}_k$ 、 $\hat{\gamma}_k$ 與 $\hat{\theta}_k$ 同時都能滿足(26)、(27)與(28)式。最後將 $\hat{\beta}_k$ 、 $\hat{\gamma}_k$ 與 $\hat{\theta}_k$ 代入(29)式求 $\hat{\alpha}$ 。

四、適合度準則

本文以五種適合度統計量 (goodness of fit statistic) (Holland and Fitz-Simons, 1982) 以及一個非統計量作為統計模式的適合度準則 (goodness of fit criteria) 以比較上述六個統計模式的擬合良度 (goodness of fit)。六種適合度準則分別為絕對偏差 (absolute deviations, AD)，權重絕對偏差 (weighted absolute deviations, WAD)，卡方檢定 (chi-square test, CS)，Kolmogorov-Smirnov type statistic (KS)，Cramer-von Mises-Smirnov statistic (CMS)，以及非統計量 Theil 氏不等式係數 (Theil's inequality coefficient, TIC) (Kheir & Holmes, 1978)。其定義分述如下：

1. 絕對偏差 (AD)

$$AD = \sum_{i=1}^m |Po_i - Pe_i| N \quad (30)$$

上式 Po_i = 數據落在第 i 組內的觀測或然率，
 Pe_i = 數據落在第 i 組內的期望或然率，
 N = 總觀測次數。

2. 權重絕對偏差 (WAD)

$$WAD = \sum_{i=1}^m |Po_i - Pe_i| Pe_i N \quad (31)$$

3. 卡方檢定 (CS)

$$CS = \sum_{i=1}^m (Po_i - Pe_i)^2 N / Pe_i \quad (32)$$

自由度為 $m-1-k$ ， k 為統計模式中待推估參數的數目。

4 Kolmogorov-Smirnov type statistic (KS)

$$KS = N \max_i |Co_i - Ce_i| \quad (33)$$

其中 Co_i = 數據小於第 i 組代表值的觀測累積頻率，

Ce_i = 數據小於第 i 組代表值的期望累積頻率。

5 Cramer-von Mises-Smirnov statistic

(CMS)

$$CMS = \sum_{i=1}^m (Co_i - Ce_i)^2 Pe_i N \quad (34)$$

6 Theil 氏不等式係數 (TIC)

$$TIC = \frac{\sqrt{\sum_{i=1}^m (Po_i - Pe_i)^2}}{\left(\sqrt{\sum_{i=1}^m Po_i^2} + \sqrt{\sum_{i=1}^m Pe_i^2} \right)} \quad (35)$$

很明顯地， $0 \leq TIC \leq 1$ 。如果 TIC 的值接近於零，則表示觀測頻率與期望頻率的分布是幾乎相同的，TIC 的值愈大表示兩者的差異愈大。通常當 $TIC < 0.3$ 時，觀測頻率與期望頻率的分布就可被認為是相當一致的。

五、結果與討論

本文以桃園濱海地區的地面風速作為個案研究，而桃園濱海地區的地面風速則以桃園中正國際機場民國 68 年至 72 年的地面風速逐時正點觀測資料作為代表。中正機場氣象觀測報告中風速單位為節 (knot, 1 節 = 1 哩 / 小時 = 0.5144 米 / 秒)，風速報告值為整數，即觀測時已將風速值分組 (例如，介於 4.5 ~ 5.5 節的風速值皆紀錄為 5 節)，因此本文中不再將風速報告值另行分組。推在計算卡方值 (chisquare value) 時將風速值分為六組 (0 ~ 2 節、3 ~ 7 節、8 ~ 12 節、13 ~ 20 節、21 ~ 30 節、30 節以上)。五年期間共有 43824 個風速觀測資料，其中最大的風速值為 55 節，發生於民國 71 年 7 月 29 日上午 7 時，當日適值安迪颶風由東部登陸侵襲本省。

為了探討桃園濱海地區冬天與夏天風速頻率的理論分布是否有差異，本文將上述風速觀測資料分為冬半年 (10、11、12、1、2、3 月) 與夏半年 (4、5、6、7、8、9 月)，上述五年期間冬半年與夏半年各有 21864 與 21960 個風速資料，其中冬半年最大的風速值為 37 節，夏半年最大的風速值為 55 節。

桃園濱海地區地面風速的一般特性列於表 1，平均風速以冬半年的較大，標準偏差以冬半年的較小，偏度係數都是正的，以夏半年的較大，峯度係數也以夏半年的較大，亦即夏半年的風速以低風速為主，而且低風速佔絕對優勢 (圖 1)。

本文以掃描法逐漸縮小所欲求參數估計值的可能範圍以求解 (10)、(14)、(15)、(19)、(22)、(23)、(26)、(27) 及 (28) 等式，精確度可自行選擇，本文參數估計值的數值算至小數點後面第四位為止。

桃園濱海地區風速頻率的六種理論分布函數，其參數的極大似然估計值列於表 2。由於桃園濱海地區全年、冬半年與夏半年零風速 (靜風) 的頻率不為零，所以表 2 中的 γ 值都小於零。

桃園濱海地區全年、冬半年及夏半年地面風速頻率理論分布函數的擬合良度值分別列於表 3a、3b 及 3c，並標出擬合良度的序數。表 3 中的 TIC 值都小於 0.3，這表示表 3 中的六種頻率分布函數都可以被接受作為桃園濱海地區全年、冬半年與夏半年地面風速頻率的理論分布函數，其中最適合作為全年風速頻率的理論分布函數是四個參數韋伯分布函數，其次是三個參數韋伯分布函數，再其次是四個參數對數常態分布函數。最適合作為冬半年風速頻率的理論分布函數是四個參數韋伯分布函數，其次是四個參數對數常態分布函數。而夏半年則以四個參數對數常態分布函數較適合。

表 2 中夏半年四個參數韋伯分布函數的參數 θ ，其極大似然估計值高達 1229.6 節，這表示夏半年的風速頻率數據以四個參數或三個參數的韋伯分布函數來擬合的效果差不多，因為四個參數頻率分布函數的第四個參數值若趨近於無限大，則四個參數頻率分布函數蜕化為三個參數頻率分布函數。因

此表 3c 中三個參數與四個參數的韋伯分佈，其六種擬合良度值都很相近。

圖 2a 與 2b 分別為繪於韋伯底圖 (Weibull paper) (Takle and Brown, 1977) 上的全年與冬半年的風速觀測數據及其四個參數韋伯分佈擬合的累積頻率圖。圖 2c 為繪於對數常態底圖上的夏半年風速觀測數據及其四個參數對數常態分佈擬合的累積頻率圖。上述三個圖顯示觀測數據都非常接近理論分佈的直線，雖然就卡方檢定而言，表 3 中的六種頻率分佈函數在 0.001 的顯著水準 (level of significance) 是被拒絕作為桃園濱海地區地面風速頻率的理論分佈函數，但這是由於卡方檢定對於大樣本相當靈敏的關係 (Holland and Fitz-Simons, 1982)。因此就工程應用上而言，四個參數的韋伯分佈函數及四個參數的對數常態分佈函數對桃園濱海地區全年、冬半年及夏半年地面風速頻

率分佈分別提供了一個非常好的近似分佈。桃園濱海地區夏半年地面風速頻率分佈的直方圖以及四個參數對數常態擬合分佈示於圖 1。

表 1 桃園濱海地區地面風速的一般特性

	全年	冬半年	夏半年
觀測總次數	43824	21864	21960
最大風速 (節)	55	37	55
最小風速 (節)	0	0	0
平均風速 (節)	11.249	12.730	9.775
標準偏差 (節)	6.886	6.672	6.779
偏度係數	0.506	0.056	1.025
峯度係數	2.708	2.102	4.071

表 2 桃園濱海地區地面風速頻率分佈函數，其參數的極大似然估計值

	全年	冬半年	夏半年
二個參數 對數常態分佈	$\mu=2.2089$ $\sigma=0.7326$	$\mu=2.3688$ $\sigma=0.6876$	$\mu=2.0490$ $\sigma=0.7413$
三個參數 對數常態分佈	$\mu=2.8686$ $\sigma=0.3740$ $\gamma=-7.6060$	$\mu=5.0452$ $\sigma=0.0429$ $\gamma=-142.69$	$\mu=2.3737$ $\sigma=0.5495$ $\gamma=-2.6531$
四個參數 對數常態分佈	$\mu=-1.3034$ $\sigma=0.7185$ $\gamma=-2.4788$ $\theta=55.672$	$\mu=-0.4344$ $\sigma=0.7466$ $\gamma=-3.9964$ $\theta=37.201$	$\mu=-1.6810$ $\sigma=0.7912$ $\gamma=-1.3115$ $\theta=59.318$
二個參數 韋伯分佈	$\alpha=12.801$ $\beta=1.7235$	$\alpha=14.493$ $\beta=2.0312$	$\alpha=11.093$ $\beta=1.5518$
三個參數 韋伯分佈	$\alpha=13.153$ $\beta=1.7404$ $\gamma=-0.4761$	$\alpha=17.187$ $\beta=2.4786$ $\gamma=-2.4976$	$\alpha=11.063$ $\beta=1.5074$ $\gamma=-1.9757$
四個參數 韋伯分佈	$\alpha=0.2704$ $\beta=1.2693$ $\gamma=-0.0530$ $\theta=61.586$	$\alpha=0.5958$ $\beta=1.2936$ $\gamma=-0.1576$ $\theta=40.986$	$\alpha=0.0091$ $\beta=1.4795$ $\gamma=-0.1760$ $\theta=1229.6$

表 3a 桃園濱海地區全年地面風速頻率分佈函數的擬合良度

	AD	WAD	CS	KS	CMS	TIC
二個參數 對數常態分佈	14759.5 (6)	480.869 (5)	6327.25 (6)	3506.62 (6)	97.2272 (6)	0.155298 (6)
三個參數 對數常態分佈	12864.8 (5)	502.684 (6)	2363.96 (5)	2458.37 (4)	57.9756 (5)	0.147376 (5)
四個參數 對數常態分佈	10729.5 (3)	410.497 (3)	1569.06 (3)	2292.75 (3)	38.3093 (3)	0.122206 (2)
二個參數 韋伯分佈	11118.0 (4)	435.698 (4)	1591.40 (4)	2559.85 (5)	50.1028 (4)	0.130632 (4)
三個參數 韋伯分佈	10337.9 (2)	408.536 (2)	1378.67 (2)	2208.23 (2)	36.5830 (2)	0.122893 (3)
四個參數 韋伯分佈	8224.9 (1)	324.425 (1)	1154.95 (1)	1421.77 (1)	14.8290 (1)	0.110545 (1)

表 3b 桃園濱海地區多半年地面風速頻率分佈函數的擬合良度

	AD	WAD	CS	KS	CMS	TIC
二個參數 對數常態分佈	9600.4 (6)	330.498 (6)	5437.64 (6)	2430.00 (6)	103.995 (6)	0.221475 (6)
三個參數 對數常態分佈	5404.3 (4)	204.493 (3)	1006.42 (4)	1614.31 (5)	27.867 (4)	0.138516 (4)
四個參數 對數常態分佈	4867.7 (2)	188.936 (2)	504.68 (2)	1009.12 (2)	16.865 (2)	0.118679 (2)
二個參數 韋伯分佈	6312.3 (5)	241.840 (5)	1411.91 (5)	1348.20 (4)	33.401 (5)	0.146330 (5)
三個參數 韋伯分佈	5351.3 (3)	207.338 (4)	750.88 (3)	1173.91 (3)	22.519 (3)	0.130908 (3)
四個參數 韋伯分佈	3173.5 (1)	122.882 (1)	257.77 (1)	652.94 (1)	5.335 (1)	0.084443 (1)

表 3c 桃園濱海地區夏半年地面風速頻率分佈函數的擬合良度

	AD	WAD	CS	KS	CMS	TIC
二個參數 對數常態分佈	4272.9 (2)	131.194 (1)	1246.44 (6)	1001.29 (5)	13.998 (5)	0.089850 (1)
三個參數 對數常態分佈	4534.3 (5)	185.029 (3)	730.89 (5)	917.02 (4)	12.422 (4)	0.100012 (3)
四個參數 對數常態分佈	3868.5 (1)	166.166 (2)	442.79 (1)	668.56 (1)	7.406 (1)	0.093197 (2)
二個參數 韋伯分佈	4879.6 (6)	221.446 (6)	616.48 (2)	1472.84 (6)	29.372 (6)	0.126434 (6)
三個參數 韋伯分佈	4395.8 (4)	207.851 (5)	710.88 (3)	786.69 (3)	11.175 (3)	0.119248 (4)
四個參數 韋伯分佈	4384.6 (3)	207.087 (4)	724.83 (4)	733.48 (2)	10.597 (2)	0.120802 (5)

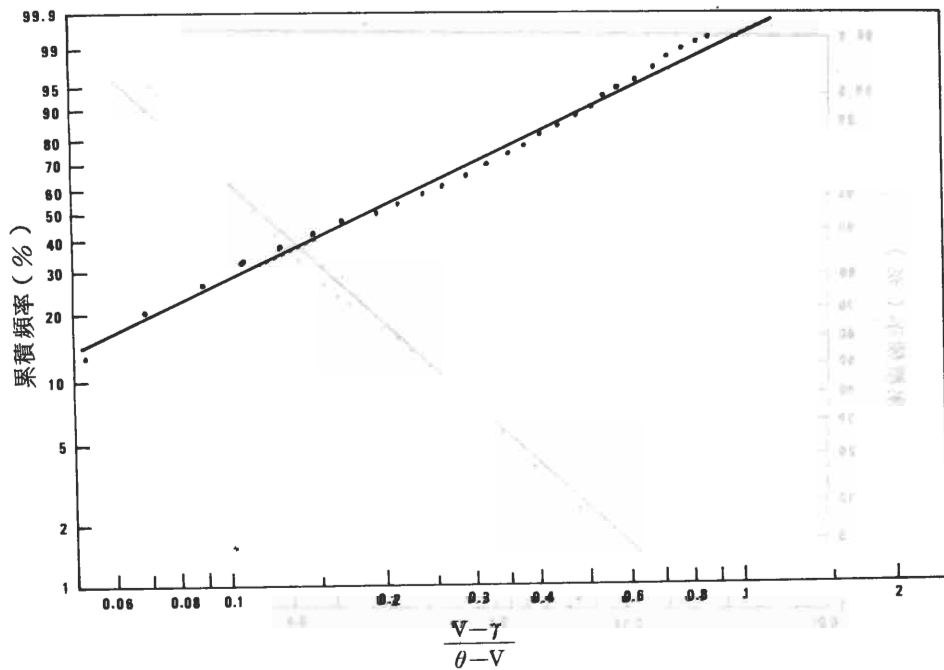


圖 2a: 桃園濱海地區全年地面風速觀測的累積頻率圖 (•) 及其四個參數韋伯擬合分佈的累積頻率圖 (—)。

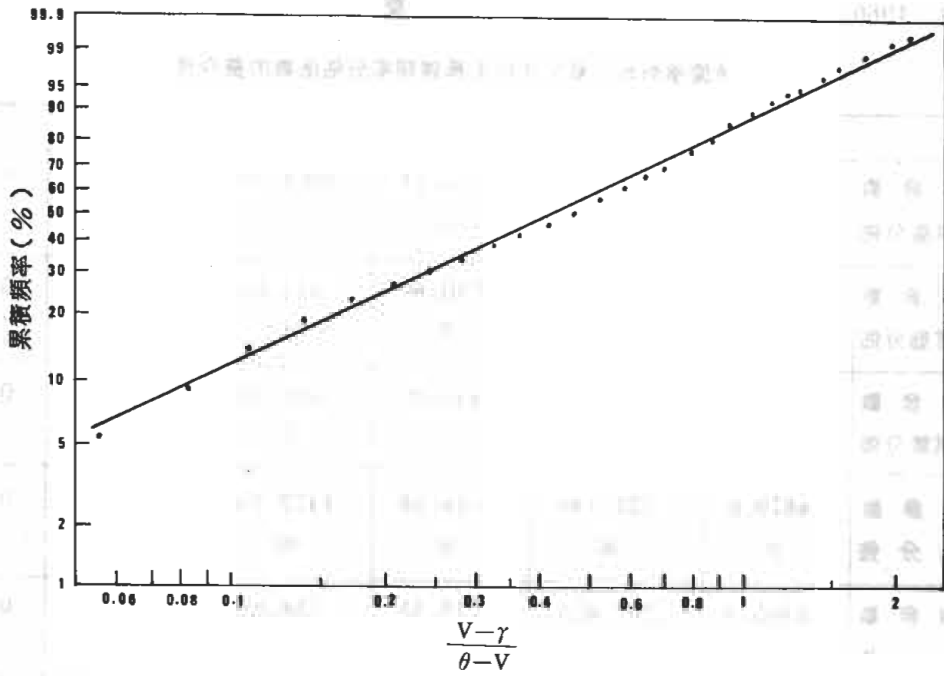


圖2b: 與圖2b. 相同, 但為冬半年。

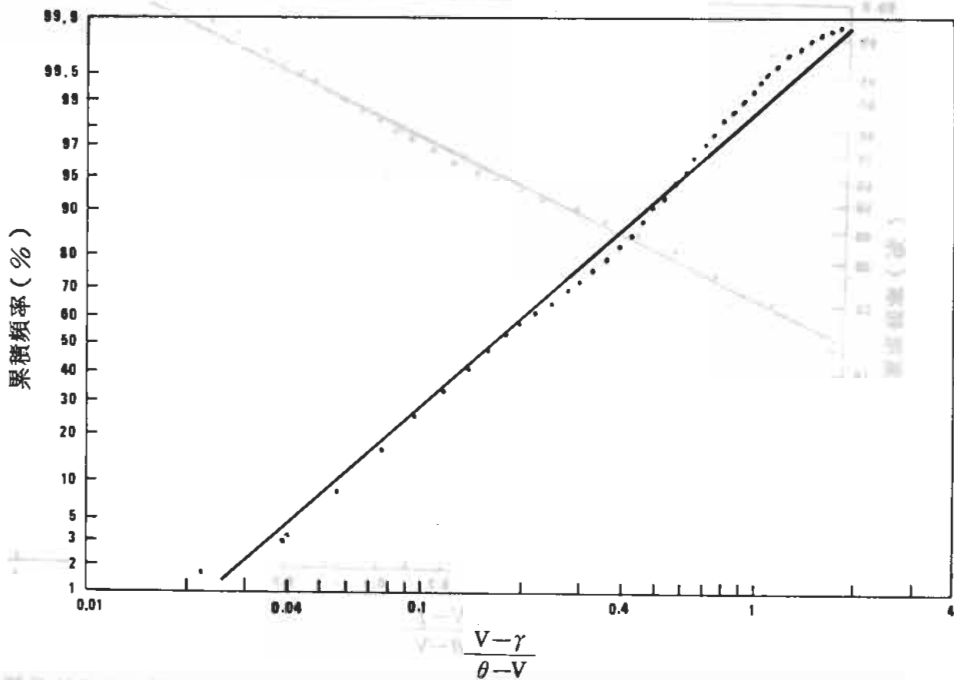


圖2c: 桃園濱海地區夏半年地面風速觀測的累積頻率圖(·)及其四個參數對數常態擬合分佈的累積頻率圖(—)。

六、結 論

本文以二個參數、三個參數、四個參數對數常態分佈函數以及二個參數、三個參數、四個參數韋伯分佈函數作為不考慮風向的風速頻率理論分佈函數，以極大似然法來推估頻率分佈函數的參數值，並以桃園中正國際機場地面風速觀測資料作為個案研究，以五種適合度統計量及 Theil 氏不等式係數作為適合度準則，結果顯示四個參數韋伯分佈函數最適合作為桃園濱海地區全年與冬半年地面風速頻率的理論分佈函數，而夏半年則以四個參數對數常態分佈函數較適合。

雖然用極大似然法推估統計模式的參數必須用疊代法反覆計算較為繁複，但是對於用其他方法很難推估參數值的三個參數及四個參數的統計模式，其參數值的推估，極大似然法確是一個可行而且效果良好的方法。

誌 謝

本文承蒙交通部民航局氣象中心曾憲瑗主任提供桃園中正國際機場地面氣象觀測資料，蒲金標技正及楊進賢先生幫忙錄製氣象資料，在此特致由衷的謝忱。

參考文獻

- 吳清吉，1982：瞬時風速的一些統計分析。大氣科學，9，43—54。
- 簡來成、梁文傑與張能復，1983：台灣北部西海岸地區風力潛能之評估。中央研究院物理研究所研究報告 FM 51，164 PP。
- Bencala, K. E., and J. H. Seinfeld, 1976: On frequency distributions of air pollutant concentrations. Atmos. Environ., 10, 941—950。
- Cats, G. J., and A. A. M. Holtslag, 1980: Prediction of air pollution frequency distribution — Part I. the lognormal model. Atmos. Environ., 14, 255—258.
- Essenwanger, O. M., 1976: Applied Statistics in Atmospheric Science, Part A, Frequencies and Curve Fitting. Elsevier, 424 PP。
- Harder, E. L., 1977: Specific output of windmills — A discovery. Proc. IEEE, 65, 1623—1625。
- Hennessey, J. P., 1977: Some aspects of wind power statistics. J. Appl. Meteor., 16, 119—128。
- Hewson, E. E., 1975: Generation of power from the wind. Bull. Amer. Meteor. Soc., 56, 660—675。
- Holland, D. M., and T. Fitz-Simons, 1982: Fitting statistical distribution to air quality data by the maximum likelihood method. Atmos. Environ., 16, 1071—1076。
- Justus, C. G., W. R. Hargraves, A. Mikhail and D. Graber, 1978: Methods for estimating wind speed frequency distributions. J. Appl. Meteor., 17, 350—353。
- Kheir, N. A., and W. M. Holmes, 1978: On validation simulation models of missile systems. Simulation, April, 117—128。
- Kwok, K. C. S., 1977: Cross-wind response of structures due to displacement dependent excitations. Ph. D thesis, Dept. of Mech. Engg, Monash University, Australia.
- Luna, R. E., and H. W. Church, 1974: Estimation of long-term concentration using a "universal" wind speed distribution. J. Appl. Meteor., 13, 910—916。

- Mage, D. T., 1980: Frequency distribution of hourly wind speed measurements. Atmos. Environ., 14, 367-374 °
- Parkinson, G. V., 1971: Wind-induced instabilities of structures. Phil. Trans. Royal Society of London, Series A, 269, 395-409 °
- Stewart, D. A., and O. M. Essenwanger, 1978: Frequency distribution of wind speed near the surface. J. Appl. Meteor., 17, 1633-1642 °
- Takle, E. S., and J. M. Brown, 1978: Note on the use of Weibull statistics to characterize wind-speed data. J. Appl. Meteor., 17, 556-559 °

Curve Fitting Frequency Distribution of Hourly Wind Speed Data by the Maximum Likelihood Method

Jin-Chihng Wang
Department of Applied Physics
Chung Cheng Institute of Technology

A computer program has been developed for fitting frequency distribution of hourly wind speed data. The program fits the following continuous distribution: two-parameter, three-parameter, four-parameter lognormal and two-parameter, three-parameter, four-parameter Weibull. The parameters of each distribution are estimated by the maximum likelihood method. Data from Tao-Yuan/Chung Cheng Airport are used to demonstrate the use of the above models to correlate wind speed data. Although the chi-square test leads to the rejection of the hypothesis that wind data are lognormal or Weibull distributed, the four-parameter lognormal and four-parameter Weibull appear to fit the data well enough for engineering calculations.