

水位數據品管檢核之研究

Study on Accuracy Verifying Procedure of Water Level Data

國立成功大學
水利及海洋工程系
副教授

顏沛華

Pei-Hwa Yen

國立成功大學
水利及海洋工程系
博士研究生

蔡宗旻*

Tsung-Min Tsai

經濟部水利署
河川海岸組
簡任正工程師

李友平

Youe-Ping Lee

摘要

對於現場蒐錄之水位數據，本研究提出觀測資料必須符合資料合理性、連續性及關聯性三項檢驗原則之品管程序，以 90% 信賴區間法濾除水位數據之異常值，對於缺漏或經品管檢核遭剔除的資料，則採用時序列卡門濾波法進行補遺；以儀器規格及探頭(感測器)高程作為合理性之比對基準，以馬可夫鏈模式來作連續性品管，以不同物理量間的關係作為關聯性檢測的依據。對於如颱風等異常氣候造成資料無法通過檢核時，則改採人工檢核，令現場量測之水位數據更為準確可靠。本研究以灣內橋及達德安攔砂壩三測站之實測值作資料合理性、連續性及關聯性三項檢驗之測試，據分析結果顯示，總缺遺筆數不應超過總檢核筆數之 7.32% (8% 以下)，否則經補遺之水位數據與實測值間可能有超過 10% 的誤差。

關鍵詞：卡門濾波，馬可夫鏈，信賴區間，品管程序。

ABSTRACT

Accuracy data verifying procedures such as observation rationality, data continuity and phenomena relativity tests are proposed in this paper. Water level data being collected including abnormal value in data bank should be screened out by statistical consideration of 90% confidence interval and the Kalman filtering algorithm was adopted here as the supplement technique for data missing. Limitations of device specifications and sensor elevation were regarded as the basis of observation rationality test. The Markov chain model (AR model) was used for data continuity verification. Physical relation between two parameters may be used as the phenomena relativity test. As to extreme events (say,

*通訊作者，國立成功大學水利及海洋工程所博士研究生，70101 台南市東區大學路 1 號，n8893106@ccmail.ncku.edu.tw

typhoon or earthquake) resulted in failure testing through the accuracy data verification procedure, the manual inspection should be running as an option to ensure data accuracy and avoid fault applications. In the aspect of data verification procedure, testing data collected from Wan-Ray Bridge and Dar-Ter-An check dam stations were used for observation rationality, data continuity and phenomena relativity testing and analyzing result revealed that the supplementary data should not exceed the total amount data by 8%, else more than 10% error may be expected in between measuring and supplementary data values.

Keywords: Kalman Filtering, Markov Chain, Confidence interval, Accuracy verifying procedure.

一、前言

水位觀測主要目的在於提供長期且正確之水位資料供各界運用，欲達成此目的除儀器需正常運作外，仍需資料的品管檢核與處理。然水位資料經由觀測、編碼、通訊傳遞、接收解碼、儲存的複雜過程中，可能因為系統的誤差、儀器的老舊、外在環境的衝擊，或其他不明原因的影響，造成資料的漏失或異常值，這些未經品管程序之原始數據(raw data)必須經過嚴格的品管過程、檢核與補遺，以避免後續計算、統計分析時所產生的錯誤資訊，而誤導水科學之學術研究結果，或造成水資源政策或工程應用、規劃、設計上難以彌補的重大疏失。

為此，本研究提出水位數據品管檢核之方法，即合理性品管、連續性品管、關聯性品管及人工品管，以此資料品管程序篩選數據中之突波(spike)與異常值(outlier)，而以時序列自回歸模式結合卡門濾波法補遺經品管程序剔除之水位數據，使水位長期觀測資料有更良好的品質。

二、文獻回顧

資料品管主要目的在於提供可信度高及正確資料給使用者。王維紀(1993)及鄧中柱(1993)認為對資料進行檢測及品質管制有其必要性。一般資料品管比較顯見於海氣象觀測，如波浪統計。唐存勇、黃淑貞(1993)以國科會水文資料庫三倍標準偏差外的資料認為係壞資料。Michael,

Marshall and David (1993) 採用線性趨勢濾除法來去除潮位變化趨勢；董東環等(1997, 1998)以合理性、連續性、關聯性三項檢驗原則，來分析高雄大林埔外海觀測樁實測資料；鄔寶林等(2000)針對雨量、水位及流量之資料進行評鑑，藉以評估各站之資料品質；蔡萬宮等(2006)將地下水位資料依完整性、合理性、連續性及關聯性等品管目標程序，建置自動化之地下水位資料品質管理檢核系統。Kao *et al.* (2003)；Doong *et al.* (2007) 在海氣象觀測上亦提出相當之品管程序。

另對於資料補遺方面，Box and Jenkins (1970) 提出 ARIMA (Autoregressive Integrated Moving Average) 模式為非常有用之時間序列模式，並廣泛應用在各種實用科學上，如經濟、氣象…等上，而此一模式具有原始觀測資料的統計性質，故可以用於繁衍大量的資料與預測短期的資料。張玉田(1982)將卡門濾波應用於洪水演算。陳昶憲、林榮坤(1993)將參數視為定率之水文模式，配合機率過程無定常性限制之卡門濾波理論，進行水文量之推估。Yen *et al.* (1996)發現卡門濾波法最適合短期潮汐之即時預測。另外虞國興、魏柏村(1994)以合成資料與實測資料來討論前人提出的幾種補遺方法，研究結果發現以 Ljung 所用之時序列方法的穩定性與精確度最佳。

三、水位數據之品管程序

一般水位數據若與觀測記錄分佈差異過大

表 1 觀測資料合理性檢測標準

| | 水位 | 風速 | 氣溫 | 氣壓 |
|------|----------|------------|----------|--------------|
| 儀器限制 | 0.3~15 m | 0~60 m/sec | -20~50°C | 900~1100 hpa |
| 物理限制 | <5.5 m | - | - | - |

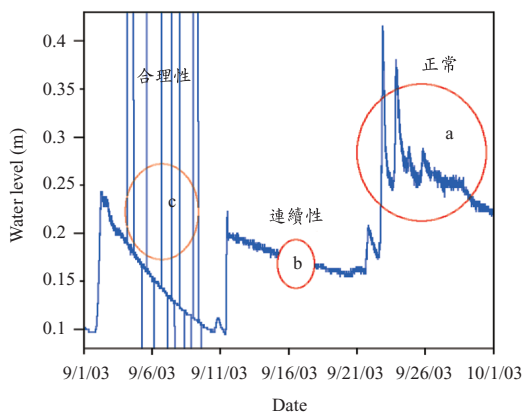
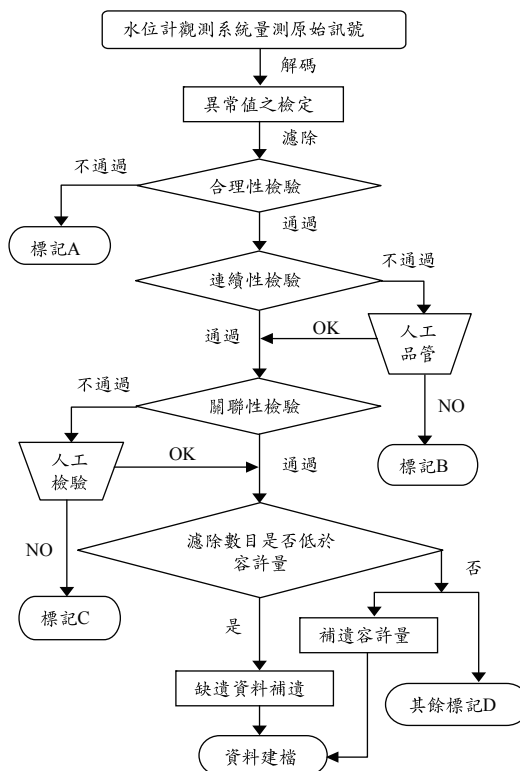


圖 1 異常值時序列型態

資料稱之為異常值。根據以往水位觀測經驗，約可彙整出異常值或缺漏數據之時序列型態如圖 1 所示，圖中之分析資料係取自於曾文水庫集水區上游之達德安攔砂壩水位觀測站，圖中橫軸為時間軸，縱軸為壩頂到水面之距離即水位。其中 a 為一正常無異常值之水位時序列；b 為持續一段時間無水位記錄(缺漏)或產生一段與原水位變化固定偏差之水位訊號；c 出現了數筆不固定之水位值且其訊號趨勢偏差較大。為了有效解決上述異常值情況，本研究以品管工作的準則必須符合資料合理性、連續性及關聯性三項檢驗為原則。其品管流程圖如圖 2 所示。

3.1 合理性品管

合理性檢定為品管中最基本的檢驗工作，一般而言，合理性品管分為兩種限制，第一是指水位量測數據不應超過觀測儀器規格所設定之範圍；第二則是觀測數據不應超過量測地點之水位臨界條件，亦即物理上的限制，如攔砂壩水位觀測站所量測到的水位不應超過攔砂壩壩頂高度或是水位計架設高度，否則即視為不合理值。表 1 分別列出各觀測物理量受儀器量測範圍的限



*標記A、B、C、D表資料可保留一年後再刪除

圖 2 水位觀測資料品管流程

制及物理現象限制合理性品管標準。

3.2 連續性品管

自然界中大部份物理量的變化是循序漸進且連續的，連續性品管即基於此假設，認為各種觀測物理量在時間及空間上都有連續的變化存在，也就是說其時空的變化量應落在合理的範圍內，如果在時間上的改變量超過某一限度或是在空間上相鄰測點的資料無法以物理觀點解釋，都視為可疑資料，以水位變動為例，即水位在某段時間內的變化量應小於某一容許值。而對於連續性品管，本研究提出兩種方法來檢驗，即採用簡單又快速濾除異常值的方法~『信賴區間推求法』，來推求容許變動值，及具有表現前後相鄰時段狀態間變化特性之『馬可夫鏈機率模式 (Markov Chain Probability Model)』來檢驗其連續性。

3.2.1 信賴區間推求法

為使有效地濾除異常值，本研究採用統計學上之信賴區間概念來作為水位數據異常值的過濾，利用此法簡便易作業之特性，可以合理的檢測出時序列資料在時域上的異常值。以上述達德安攔砂壩水位觀測站為例，其分析程序首先計算觀測資料前後時刻間隔設定為 5 分鐘之差值 Δ_5 ，若有 N 個分析數據，則將會有 N-1 個 Δ_5 ，本研究假設以上所計算出來的稽延 5 分鐘差值符合一統計分佈。計算上述差值分佈的 90% 信賴區間 Δ_5^{90} ， Δ_5^{90} 代表在 90% 的信賴區間範圍水位數據在下 5 分鐘後所能接受的最大變動範圍，亦即所需過濾的異常值檢定標準，故可建立連續性品管之鑑別公式如下：

$$El_5(n-1) - \Delta_5^{90} \leq El_5(n) \leq El_5(n-1) + \Delta_5^{90} \dots (1)$$

式中， $El_5(n-1)$ 、 $El_5(n)$ 為每 5 分鐘第 n-1 筆、第 n 筆之水位數據，

$$\Delta_5^{90} = \frac{0.1}{n-1} \sum_2^n [El_5(n) - El_5(n-1)] \dots \dots \dots (2)$$

同理亦可求出 10 分鐘、15 分鐘...的檢定標準 (Δ_{10}^{90} 、 Δ_{15}^{90} ...)，將此品管標準與 $t_5 = \Delta t = 5$ 、 $t_{10} = \Delta t = 10$ 、 $t_{15} = \Delta t = 15$... 迴歸得一公式：

$$Y = 0.000114 + 0.000278X + 4.68 \times 10^{-6} X^2$$

式中，Y 為最大容許變動範圍 (Δ_5^{90} 、 Δ_{10}^{90} 、 Δ_{15}^{90} ...)，X 為時間量測間距 Δt ($\Delta t = 5 = t_5$ 、 $\Delta t = 10 = t_{10}$ 、 $\Delta t = 15 = t_{15}$...)。

上式係以達德安攔砂壩歷史水位 57334 筆資料為例分析而得，此公式僅供參考，並不適合所有水位觀測站應用，其相關係數達 0.997，今以量測時間間距 $X = \Delta t = 10$ 分鐘說明之，其異常值濾除之最大容許變動範圍 Y 為 0.003362 m，則下一時刻容許變動範圍是以第 n 筆水位值再加減 0.003362 作為第 n + 1 筆水位資料的最大濾除範圍之上下限，依此類推，其成果如圖 3 所示。

3.2.2 馬可夫鏈模式

一般水文歷程之遺傳效應可用馬可夫過程 (Markovian Process) 來說明，今假設有一水文量

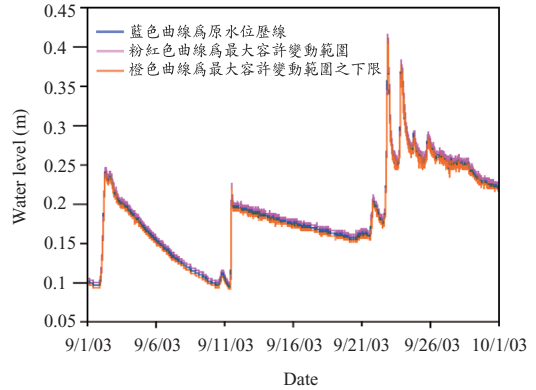


圖 3 異常值濾除之最大容許變動範圍

時間系列 $X(t_1)$ 、 $X(t_2)$ 、...、 $X(t_n)$ ，且 $t_1 < t_2 < \dots < t_n$ ，在時間 t_n 之水文量 $X(t_n)$ 所發生之機率係受臨前水文量 $X(t_1)$ 、 $X(t_2)$ 、...、 $X(t_{n-1})$ 之綜合影響，可用條件機率 $P[X(t_n) | X(t_{n-1})、X(t_{n-2})、\dots、X(t_1)]$ 表示之，此即為 n-1 階馬可夫鏈 (n-1 order Markov chains)。由於水文歷程之遺傳效應受前一時刻之影響最大，於實際之應用上僅需推求一階馬可夫鏈之條件機率 $P[X(t_n) | X(t_{n-1})]$ 即可。

換言之，表示在下一階段可能出現某種狀態之機率，只依現在階段的狀態而定，與過去的歷史狀態無關，此即為馬可夫過程之無記憶 (no memory) 特性。前後相鄰兩階段之狀態存在有移轉機率矩陣之關係，透過移轉機率矩陣，可根據目前之狀態推求未來任一階段發生不同狀態之機率。

進一步而言，由階段 t_m 的狀態 i 轉變到階段 t_n 的狀態 j 其移轉機率可以用下式表示：

$$P_{i,j}(m, n) = P[X(t_n) = j | X(t_m) = i] \quad n > m \dots \dots \dots (3)$$

假使由狀態 i 經 k 階段後轉變為狀態 j 之移轉機率適用於任一時段 (s)，與時間無關，則稱此過程為齊次 (homogeneous) 馬可夫過程。定義如下式， $s \geq 0$

$$P_{i,j}(k) = P_{i,j}(s, s+k) = P[X(t_{k+s}) = j | X(t_s) = i] \dots \dots \dots (4)$$

表 2 移轉機率矩陣

| | 0~0.05 | 0.05~0.1 | 0.1~0.15 | 0.15~0.2 | 0.2~0.25 | 0.25~0.3 | 0.3~0.35 | 0.35~0.4 | 0.4~0.45 |
|----------|--------|----------|----------|----------|----------|----------|----------|----------|----------|
| 0~0.05 | 99.97 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.05~0.1 | 0.06 | 99.7 | 0.23 | 0.01 | 0 | 0 | 0 | 0 | 0 |
| 0.1~0.15 | 0 | 1.17 | 98.42 | 0.41 | 0 | 0 | 0 | 0 | 0 |
| 0.15~0.2 | 0 | 0 | 0.54 | 97.9 | 1.56 | 0 | 0 | 0 | 0 |
| 0.2~0.25 | 0 | 0 | 0 | 4.67 | 90.16 | 5.17 | 0 | 0 | 0 |
| 0.25~0.3 | 0 | 0 | 0 | 0 | 9.43 | 90.19 | 0.38 | 0 | 0 |
| 0.3~0.35 | 0 | 0 | 0 | 0 | 0 | 5.4 | 86.95 | 7.65 | 0 |
| 0.35~0.4 | 0 | 0 | 0 | 0 | 0 | 0 | 8.1 | 86.5 | 5.4 |
| 0.4~0.45 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 50 | 50 |

此為 k 階段移轉機率(k-step transition probability)，代表經過 k 階段後由狀態 i 變為狀態 j 之移轉機率。

根據歷史觀測資料，可適當得將觀測值分為 c 個不相互重疊之狀態區間，並使用下式估算一階段移轉機率 $P_{i,j}(1)$ 值：

$$P_{i,j}(1) = \frac{f_{i,j}}{\sum_{j=1}^c f_{i,j}} \quad i, j = 1, \dots, c \quad \dots\dots\dots(5)$$

其中 $f_{i,j}$ 為由狀態 i 經一階段後轉變為狀態 j 之發生頻率(次數)。如此可將這些移轉機率歸納入一個 $c \times c$ 之矩陣中，此即為一階段移轉機率矩陣(1-step transition probability matrix)。

$$P(1) = \begin{bmatrix} p_{11}(1) & p_{12}(1) & \dots & p_{1c}(1) \\ p_{21}(1) & p_{22}(1) & \dots & p_{2c}(1) \\ \vdots & \vdots & \vdots & \vdots \\ p_{c1}(1) & p_{c2}(1) & \dots & p_{cc}(1) \end{bmatrix} \quad \dots\dots\dots(6)$$

由一階段移轉機率矩陣 $P(1)$ 即可利用 $P(k) = P(1)^k$ 進一步決定 k 階段移轉機率矩陣 $P(k)$ 。

由建立從狀態 i 經一階段後轉變為狀態 j 之移轉機率矩陣 $P(1)$ ，只需給定初始階段($t = 0$)之不同狀態的發生機率，透過移轉機率矩陣，即可獲得未來任一前置階段($t > 0$)發生不同狀態之機率。

以達德安攔砂壩水位站 2003 年 7 月至 2004 年 1 月之水位資料為例，將水位區分為九個狀態，由歷史資料中統計計算在前期狀態情形下發

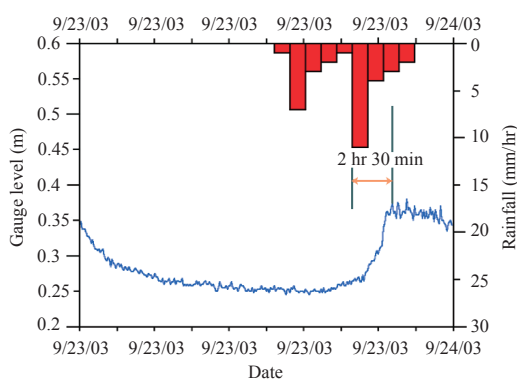


圖 4 達德安攔砂壩水位-雨量歷線

生後期狀態的移轉機率矩陣，其計算結果如表 2 所示，由表中可以發現前後兩階段發生相同狀態的事件機率最大。亦即當水位在任何狀態下，在其前後兩狀態外之水位，皆可視為不連續(異常值)。

3.3 關聯性品管

關聯性檢驗是檢驗所量測之資料與其他相關物理量之間原有之關係是否有異常現象，用不同物理量間之關聯性，可互相檢驗資料的品質。例如上游雨量異常的大，則下游水位不可能沒有升高現象，然而到底要以多長的時間作為關聯性的濾除範圍，本研究將其分為兩類來討論：

3.3.1 上游雨量站~下游水位站之關聯性

以達德安水位站及其上游水山雨量站為例，從水位-雨量歷線圖之延時(lag time)關係可作為比較趨勢的參考(圖 4)。

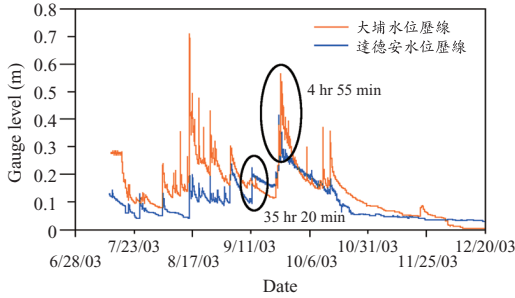


圖 5 達德安攔砂壩及大埔壩水位歷線比對

3.3.2 上游水位站~下游水位站之關聯性：

以 2003/07/12~2004/01/27 之達德安(上游)及大埔(下游)兩水位站為例，因在水文應用上所關心的是大流量，因此具體作法上可利用水位歷線疊圖求最小延時作為濾除範圍依據(圖 5)或以求斜率方式比較趨勢正確與否。此外同一物理量使用不同測量儀器或是相同儀器同時量測，除了可明顯的比較資料正確性外，對於儀器老化或毀損，更可顯示出其偏差量。

3.4 人工品管

當品管系統流程架構建立之後，其整個檢測流程皆可以用計算機自動控制，但人工檢測的工作仍然不可避免，如颱風等異常氣候來臨時，此時觀測資料可能就無法通過相關的檢測項目，故必須要用人工檢測的方式，並且需要調閱附近的測站或相關資訊來判斷欲保留或刪除該筆可疑資料，因極端資料得之不易，但若保留錯誤過大或過小紀錄也會造成日後分析上的困擾，因此人工檢測為不可或缺的一環。

四、水位數據補遺

當失真的水位資料在經過上述異常值的檢定濾除後，原始時序列將產生某些缺口，因而造成對後續資料分析或計算上的困擾，故必須進行補遺工作。而補遺的原則為盡可能保留原有觀測資料所隱含之訊息，以避免後續計算、統計分析時所產生的錯誤資訊，而誤導學術研究的結果，或造成工程應用上的設計錯誤或相關決策之重大疏失。因此，本研究以時間序列模式 AR(2)搭

配卡門濾波法求 AR 模式中之參數來作水位數據之補遺。

4.1 補遺應用之理論

對於水位數據的補遺，本研究係利用時間序列之自迴歸模式(Autoregressive model, AR model)進行補遺，其中 AR(2)中之參數是以卡門濾波法來求解，敘述如下：

4.1.1 自回歸模式[AR(p)model]

時間序列是以時間為基軸，對應過去觀測資料之特性做序列統計分析，並建立一套可供參考模擬推估的序率模式。利用此一模式分析之各種資料，經由模式參數的檢定(calibration)和驗證(verification)求得一可以信賴的時間序列，而此一模式具有原始觀測資料的統計性質，故可以用於繁衍大量的資料與預測短期的資料。

本研究係利用 ARIMA 模式中之自迴歸模式部分，其限制條件為只適用於定常性(stationary)之時間序列，亦即水文時間序列尚在平均值有效範圍內變動，其機率分配不因時間而改變，所以可用於繁衍及預測資料。AR(p)模式可以用下式表示：

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \dots + \phi_p Z_{t-p} + \varepsilon_t \quad \dots\dots\dots(7)$$

$$Z_t = \frac{y_t - \mu}{\sigma} \quad \dots\dots\dots(8)$$

$$\varepsilon_t = \sigma_s \times \zeta_t \quad \varepsilon_t \sim NID(0, \sigma_s^2) \quad \dots\dots\dots(9)$$

式中：

ϕ_i : 為 AR(p)模式之參數($i = 1, 2, 3 \dots p$)

P : AR 模式之階數

ε_t : 白噪音(White Noise)，亦稱擾動值

σ_s^2 : 白噪音之變異數

ζ_t : 常態分佈之隨機變數， $\zeta_t \sim NID(0, 1)$

Z_t : 標準化後之時間數列

T : 時間

y_t : 呈常態分佈之時間數列

μ : y_t 數列之平均值

σ : y_t 數列之標準偏差

上述(7)式表示由過去之水文觀測量可以推求未來之水文量。

4.1.2 卡門濾波法

對於水位數據短期之補遺，本研究採用卡門濾波法，因其具有：(1)不假定機率過程為定常性，(2)系統的狀態量係經觀測系統而測得，觀測值輸入系統後，能修正系統輸出變數及參數值，(3)演算時無須保存以往的資料，如施以逐項計算，即可求得最適宜之解答，(4)其理論完備，可用以預測觀測輸出值，且可定出模式參數之特性。

卡門濾波理論係由 Kalmann (1960)所發表，其假設物理系統為一動態系統(dynamic system)，而此系統中具有若干個參數，此等參數可經由一觀測系統測得之，然後由此等參數之觀測值，進而推求系統之輸出。其過程為一『預測→校正→更新』之反覆過程，演算時無須保留太多以往之資料為其一大特色。

卡門濾波法為一時間數列，有以下兩個方程式：

1. 狀態方程式(State Equation)

$$X(t+1) = \Phi(t+1, t) \cdot X(t) + V(t) \dots\dots\dots (10)$$

2. 觀測方程式(Measurement Equation)

$$Y(t) = M(t) \cdot X(t) + W(t) \dots\dots\dots (11)$$

式中：

$X(t+1)$ ：為 $t+1$ 時刻之狀態向量， n 維

$\Phi(t+1, t)$ ：為 t 時刻至 $t+1$ 時刻狀態變遷矩陣， $n \times n$ 階

$X(t)$ ：為 t 時刻之狀態向量， n 維

$V(t)$ ：系統干擾向量， n 維

$Y(t)$ ：自觀測系統所得變量， m 維

$M(t)$ ：系統觀測矩陣， $m \times n$ 階

$W(t)$ ：觀測誤差向量， m 維

其演算方法乃是利用(11)式之觀測方程式求出觀測系統之輸出變量，然後利用(10)式狀態方程式求出下一時刻之狀態向量，狀態向量求出後在將其代入觀測方程式求出下一時刻觀測系統之輸出變量，如此反覆之運算，即可得到每一時刻之系統輸出值。

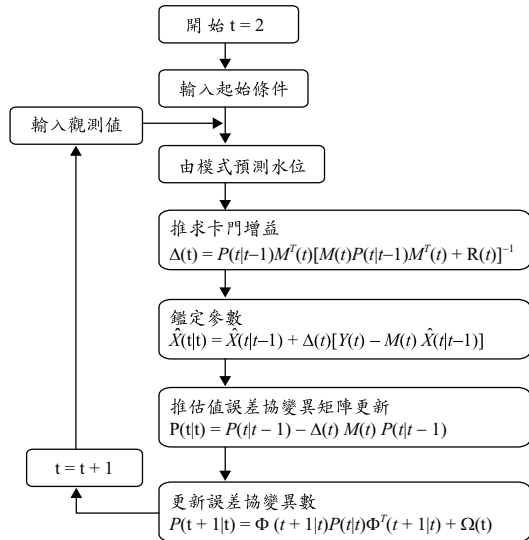


圖 6 卡門濾波法流程圖(修改於李友平, 1997)

4.2 模式建立與驗證

為了建立卡門濾波法之水位預測模式，以便於可以將其應用在水位資料缺漏或異常值濾除時之補遺用，首先須將(7)式變為 AR(2)，如下(12)式所示，此即為卡門濾波之觀測方程式。

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \varepsilon_t \dots\dots\dots (12)$$

利用上式之觀測方程式求出觀測系統之輸出變量，然後利用(10)式狀態方程式求出下一時刻之狀態向量，狀態向量求出後在將其代入觀測方程式求出下一時刻觀測系統之輸出變量，如此反覆之運算，即可得到每一時刻之系統輸出值，其執行流程如圖 6 所示。

在執行卡門濾波演算時，必須輸入系統狀態向量起始值 X 、推估值誤差協變異數矩陣起始值 P 、系統誤差協變異數矩陣起始值 Ω 、觀測誤差協變異數矩陣起始值 R 等，為了選擇適合模式之最佳起始值，本研究卡門濾波法演算時之初始值假設如下：

1. 推估值誤差協變異數矩陣 $P(2/1)$ ：

因系統參數狀態起始值係假設值，所推估的水位起始值誤差必定很大，而其假設值執行後如圖 7 所示， $P \times I$ 值越大時誤差越小，而圖中以 $P(2/1) = 10^4 \times I$ 與原始水位最為吻合，故以此值為推估值誤差協變異數矩陣起始值，其中 I 為三階

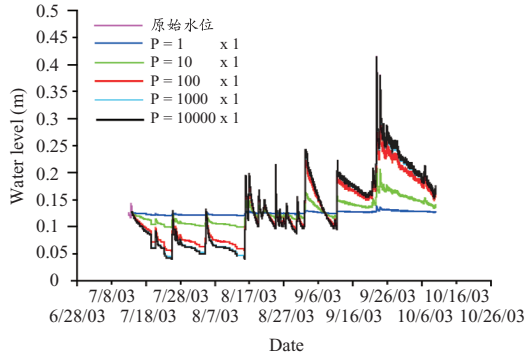


圖 7 卡門濾波法推估值誤差協變異數矩陣之起使條件對水位鑑定之檢定圖

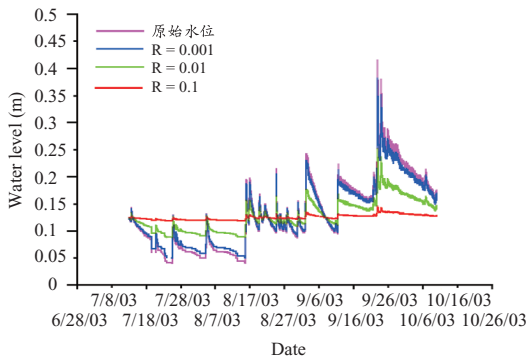


圖 8 卡門濾波法觀測誤差協變異數矩陣之起使條件對水位鑑定之檢定圖

單位矩陣，如此足夠大的推估誤差協變異數起始值才足以包容卡門濾波執行過程中之推估誤差而不致發散。

2. 系統干擾之協變異數矩陣 $\Omega(t)$:

因模式系統噪音 $V(t)$ 係由環境狀況反應而導致之改變值，而在現場水位觀測中系統誤差比觀測誤差相較穩定，故系統干擾誤差之協變異數矩陣在此過程皆假設 $\Omega(t) = 0 \times I$ ，其中 I 為三階單位矩陣。亦即本研究假設模式中之系統噪音誤差包含在觀測誤差中。

3. 觀測誤差之協變異數矩陣 $R(t)$:

因觀測噪音 $W(t)$ 係由現場量測誤差所導致，而水位觀測水位計其量測解析度(resolution)可達公厘(mm)級，因此在本模式中量測誤差協變異數矩陣在整個試驗過程皆假設 $R(t) = 0.001$ 。由圖 8 也可以看出 $R(t)$ 值越小越好。

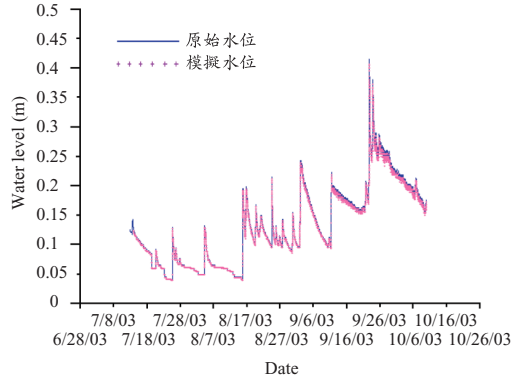


圖 9 利用卡門濾波法模擬水位成果圖

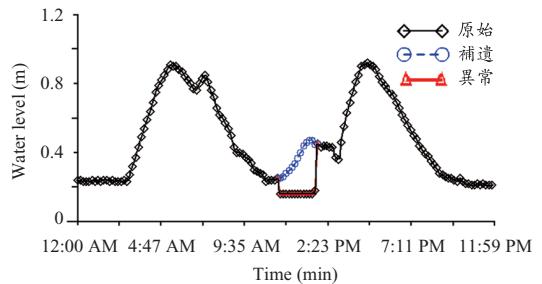


圖 10 濾除及補遺成果圖

4. 狀態矩陣之初始值 $X(2/1)$:

狀態矩陣之初始值在模式進行前並無法得知，故須事先作假設，因此本文假設狀態矩陣之初始值

$$X(2/1) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

根據上述初始值假設，本研究以達德安水位觀測站 2003/7/13~2003/10/9 之水位記錄來做驗證，其結果如圖 9 所示。

本研究另以朴子河流域上之灣內橋水位觀測站 2004/4/27 之水位資料為例，發現所設置之超音波水位計在中午時段出現約二小時之異常值，此時即可將此二小時異常情形以上述各方法進行濾除及補遺，其成果如圖 10 所示。

4.3 最大缺遺資料數的決定

倘若一水文時序列中某些資料發生遺失，導致資料的連續性被破壞，將會造成研究人員僅能

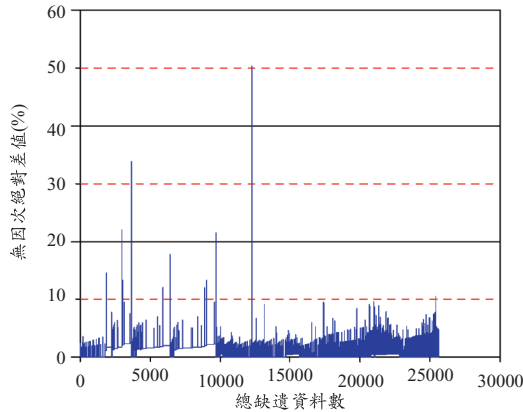


圖 11 最大缺遺資料數的決定

以具連續性部分資料分析之遺憾。如乾旱頻率分析中之聯程分析(Run Length)，就需要使用連續的時序列資料來分析。因此遺缺資料的補遺有其必要性，然而遺缺的樣本數必須有一定的限制，亦即某筆觀測數據中最多不得超過一定的遺缺量，否則補遺的模式將無法有效的補遺，該筆記錄將會失去其原有之物理特性。因此最多容許遺缺多少測點數據，仍可藉由補遺的技術來尋回原有資料特性，亦為本研究探討之重點。

本研究選用無異常值的實測記錄(以達德安攔砂壩水位站之水位數據為例)，採用無因次絕對差值(absolute error, AE)作為評鑑指標，分析結果如圖 11 所示。

$$AE = \left| \frac{P - \hat{P}}{P} \right| \dots\dots\dots (13)$$

式中：

P 為水位 h 實際值，

\hat{P} 則為隨機捨棄不同數目樣本數後，經補遺模式差補後計算所得之水位值。

由圖中發現即時缺遺達 25000 多筆水位資料，使用卡門濾波法補遺其誤差幾乎都在 90% 信賴區間內，然而約在 2000~10000 筆缺遺數間，誤差明顯偏大，經補遺後計算之水位值與實際值有可能超過 10% 的誤差，因此為了使補遺效果可靠性增加，本研究認為總缺遺筆數不應超過總檢核筆數之 7.32% (1878 筆)，亦即水位資料的現場

或遠端遙控接收盡量以不超過 8% 為原則，以避免經補遺後之數據有超過 10% 之誤差。

五、結 論

1. 本研究對所蒐集之水位數據提出觀測資料必須符合資料合理性、連續性及關聯性三項檢驗原則之品管程序，以信賴區間推求法及馬可夫鏈機率模式從事合理性及連續性品管，以不同物理量間的關係作為關聯性檢測，並以卡門濾波法解算自迴歸 AR(2) 模式之參數，從事水位數據之補遺，提供一套水位數據之品管程序與檢核方法，供有關單位參考應用。
2. 為維持資料之連續性，對於缺漏或因異常值遭檢核剔除的資料，本研究以達德安站為例分析顯示，最大容許資料遺缺數為總檢核筆數之 7.32% (達德安防砂壩補遺 1878 筆，1 筆/5 分鐘，補遺誤差約 3% 內)，故為避免資料缺漏或異常值太多而造成無法補遺之憾，總缺遺筆數不應超過總檢核筆數之 8%。
3. 研究係選用無異常值的實測記錄，再假設此段時程中皆缺漏或有異常值而遭濾除的狀態下，重新補遺的結果，得知在 8% 之總缺遺筆數內的補遺效果可靠性較佳，故應可適用在其他案例中。
4. 雖然資料品管作業大部分可藉由電腦做自動化處理，然而面對變化萬千的自然界氣象狀況，總會掛一漏萬，當資料發生異常現象時，電腦或將無法做出正確或是合理的判斷。再者程式語言的撰寫只能對既定的規則做常態性的處理，超出規則範圍以外的特例，就需要做額外的人為判斷，而其判斷的準則往往來自於豐富的經驗，及時對於資料的品管做決策性的判斷，因此本研究建議人工檢校作業不可或缺。
5. 資料品管的消極意義在找出並除去可疑資料，而在積極面則是進一步對資料作驗證以及發現儀器可能已出現異狀等問題。由於資料品管與整個資料處理流程息息相關，故本研究建議，對測量儀器的特性、設計及校正、測量方法的理論及限制、通訊儀器的維修、資料處理

的方法、環境背景資料的相關性、關聯性檢驗及資料保管方法都必須深入研究，如此方能對任何時空之資料特性作長期而完整的分析。

參考文獻

1. 王維紀，1993，「浮標測波資料處理之評介」，海象分析與預報研討會論文集，第 189-221 頁。
2. 李友平，1997，「抽水試驗改良與即時參數鑑定」，國立成功大學水利暨海洋工程學系研究所博士論文。
3. 唐存勇、黃淑真，1993，「國內海氣象資料庫概況及評估」，海象分析與預報研討會論文集，第 113-124 頁。
4. 張玉田，1982，「Kalman Filter 應用於水文學之研究」，台灣水利，第 30 卷，第 1 期，第 1-6 頁。
5. 陳昶憲、林榮坤，1993，「卡門濾波法在水文實用性之探討」，逢甲學報第 26 期，第 269-278 頁。
6. 董東璟、莊士賢、高家俊，1997，「海氣象觀測資料品質管系統之建立」，第十九屆海洋工程研討會論文集，第 477-484 頁。
7. 董東璟、周恆豪、翁俊鴻、高家俊，1997，「海氣象觀測資料品質管制系統之建立」，海下技術季刊，第 7 卷，第 2 期，第 2-9 頁。
8. 董東璟、莊士賢、高家俊，1998，「海氣象時序資料品質管」，第二十屆海洋工程研討會論文集，第 351-358 頁。
9. 虞國興、魏柏村，1994，「水文時間序列補遺之研究」，第七屆水利工程研討會論文集，第 D137-D147 頁。
10. 鄔寶林、黃俊宏、莊友欽，2000，「雨量及流量資料品質之評鑑與品質管」，第十一屆水利工程研討會，第 B47-B52 頁。
11. 鄧中柱，1993，「作業化資料浮標評介」，海象分析與預報研討會論文集，第 33-45 頁。
12. 蔡萬宮、韓光恩、李友平、賴鴻成、林家棋、孫培基、林璟翔、張順年，2006，「自動化地下水位資料品質管理程序建立」，第十五屆水利工程研討會，第 E117-E125 頁。
13. Box, G.E.P. and Jenkins, G.M., (1970) Time Series analysis : Forecasting and control. San Francisco: Holden Day.
14. Doong, D.J., Chen, S.H., Kao, C.C., Lee, B.C., (2007) Data Quality Check Procedures of an Operational Coastal Ocean Monitoring Network, *Ocean Engineering*, Vol. 34, 234-246. (SCI)
15. Kalman, R.E., (1960) A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering*, Mar.
16. Kao, C.C., Doong, D.J., Chuang, Laurence Z.H., Lee, B.C., (2003) Quality Check Procedures on In-Situ Data from the Coastal Ocean Monitoring Net Around Taiwan, *Proceedings of the 13th International Offshore and Polar Engineering Conference (ISOPE 2003)*, Honolulu, Hawaii, USA, May 25-30, 196-201. (EI)
17. Michael. T., Marshall, E. and David, M. (1993) The Development of a Wave Data Analysis Standard for a National Wave Measurement Program, *Proceedings, Wave93, The Second Symposium on Ocean Wave Measurement and Analysis*, 547-553.
18. Yen P.H, Jan C.D, Lee Y.P, and Lee H.F., (1996) Application of Kalman filter to short-term tide level prediction. *Journal of Waterway Port Coastal And Ocean Engineering-ASCE*, 122(5), 226-231.

收稿日期：民國 98 年 3 月 30 日

修正日期：民國 98 年 7 月 6 日

接受日期：民國 98 年 7 月 7 日