

海氣象觀測資料之自動品管流程

莊士賢¹

台南水工試驗所
副研究員

董東璟²

成功大學水利及海洋工程所
博士班研究生

高家俊³

成功大學水利及海洋工程所
教授

摘要

高雄市環保局於大林埔外海約1公里處已設置一海氣象觀測樁，量測之資料計有氣溫、水溫、氣壓、風速、波浪水位等。為提供可信度高之資料給使用者，有必要建立一套自動品管流程，將可疑的資料檢查出來，以提供進一步的判別。

基本上品管作業乃經由合理性、連續性及關聯性三原則進行檢驗，本文以統計學中的馬可夫鏈模式推求連續性的容許變動值，分析結果顯示，馬可夫鏈可根據不同之前期狀態提供一較合理的品管標準。針對不同物理量間的關聯性，探討不同風速計、3秒陣風與平均風速、有義波高與有義週期、氣溫與水溫、氣壓與水溫間的關係，作回歸分析，再以迴歸曲線的95%信賴區間作為後續關聯性檢驗的上、下限。除此亦探討當地風速與波浪頻譜中高頻能量的相關性。最後訂定出品管系統的檢測流程以及關聯性品管的檢測順序以供觀測系統的品管作業之進行。

一、前言

海氣象觀測資料理應「量」與「質」並重，在規劃環島海氣象觀測網站以增加觀測「量」的同時，對於「質」的要求亦刻不容緩。現有海氣象觀測浮標及觀測樁之主要目的在於提供長期的觀測資料供各界運用，欲達成此目的除其他各子系統的正常運作外，最後仍需資料品管系統的檢測與處理。通常資料在經由觀測、編碼、通訊傳遞、接收解碼、儲存等複雜的流程中，可能因為儀器設備的故障、人為疏忽或是一些不明原因的干擾，造成資料的錯誤或漏失，這些失真的資料若不經過適當的品質檢測，其謬誤極易誤導學術研究的結果，或造成工程應用上的設計錯誤等重大影響，故對資料進行檢測及品質管制有其必要性(王維紀，1993；鄧中柱，1993)。總而言之，若觀測系統所提供的資料不正確或不可靠，那可能比沒有量測到資料還要糟。

資料品管的消極意義在找出並除去可疑的資料，而在積極面上則是進一步對資料作驗證以及發現儀器的故障等問題。一般來說，資料品管分為兩種，一為自動品管、另一種則為人工品管，本文主要是研究自動品管的流程。以台灣實測海氣象資料為分析對象，參考 NDBC 品管藍圖，以合理性、連續性以及關聯性為自動品管的準則。

二、品管標準制訂

(一) 分析資料

本文的分析資料有波浪資料及氣象資料兩類，觀測站是高雄環保局設置於高雄大林埔外海約 1 公里的南星海氣象觀測樁，當地水深約 15m，採用的波浪資料包含有義波高 H_y 與有義週期 T_y ，觀測系統每兩小時觀測一次，每次 10 分鐘，取樣頻率為 5Hz；氣象資料則有氣溫、

分鐘，取樣頻率為 5Hz；氣象資料則有氣溫、水溫、氣壓、風速等資料，觀測資料經無線電自動傳回岸上接收站後，再傳回監控中心。本文選用自民國 85 年 7 月至民國 86 年 6 月共計 1 年的資料作為分析的對象，其中觀測樁上共設置有兩支風速計，可為本文關聯性品管時分析應用。

(二) 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，第一是指海氣象觀測數據不應超過量測儀器的可測範圍，如大林埔觀測樁的風速計可量測最大值為 60 m/sec，若觀測資料中出現負值或超過 60m/sec 的數據，則為錯誤的資料需刪除；第二則是觀測數據不應超過量測地點的海氣象臨界條件，亦即物理上的限制，如量測波高不應超過當地水深的碎波波高，否則即視為不合理值。

(三) 連續性品管

連續性品管是基於自然界中物理量的變化是循序漸進的假設，各種觀測物理量在時間及空間上的變化都必須有連續的，也就是說其時空的變化量應落在某一容許範圍內，否則視為可疑資料。

連續性品管分為時間連續性與空間連續性兩種品管，本文以下採用統計學中的馬可夫鏈模式(Markov chain model)來建立時間連續性的品管標準。該模式具有表現前後相鄰時段狀態間變化之特性，可用來描述前後時刻具有高度相關性的海氣象資料特性。

1. 馬可夫鏈模式

今假設有一水文量時間系列 $X(t_1)$ 、 $X(t_2)$ 、...、 $X(t_n)$ ，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，此即為 $n-1$ 階馬可夫鏈。由於水文歷程之遺傳效應以先前一時刻之影響最大，實際應用上僅採用一階馬可夫鏈之條件機率 $P[X(t_n)|X(t_{n-1})]$ 即可。該理論指出相鄰兩階段之狀態存在有移轉機率矩陣 (transition probability matrix) 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：

$$P_{i,j}(m,n) = P[X(t_n)=j | X(t_m)=i] \quad n > m \quad (1)$$

實際應用時、根據歷史觀測資料，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：

$$p_{i,j}(1) = \frac{f_{i,j}}{\sum_{j=1}^c f_{i,j}} \quad i, j = 1, \dots, c \quad (2)$$

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)、將其納入一個 $c \times c$ 之矩陣中，即為一階段移轉機率矩陣，如下。

$$P(1) = \begin{bmatrix} p_{11}(1) & p_{12}(1) & \dots & p_{1c}(1) \\ p_{21}(1) & p_{22}(1) & \dots & p_{2c}(1) \\ \vdots & \vdots & \vdots & \vdots \\ p_{c1}(1) & p_{c2}(1) & \dots & p_{cc}(1) \end{bmatrix} \quad (3)$$

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

2. 分析結果

以有義波高為例，將波高區分為十個狀態，由歷史資料中統計在前期狀態情形下發生後期狀態的移轉機率矩陣，如表 1 所示。由計算結果中發現前後兩階段發生相同狀態的事件機率最大，此亦證實小波高接著發生小波高，大波浪後發生大波浪是最常發生的。

然而上述的移轉機率矩陣並不適合本文進行連續性品管標準的制訂，本文將上法略做修正，同樣地以十個波高區間作為前期狀態，但改以前後觀測量(2小時)的變動差值作為下一階段的狀態變數，分為十個區間 (0~5、5~10、10~15、15~20、20~25、25~30、30~35、35~40、40~50、>50公分)，由歷史資料中統計移轉機率矩陣如圖 1 所示，內差求出累加 90% 機率 (填滿部份) 對應的後期狀態作為連續性品管檢驗容許範圍。各種前期波浪狀態下，稽延 2 至 12 小時的容許變動值列於表 2 中，舉例說明如現階段為狀態一，也就是發生 0~20 公分的小波浪(狀態

一)時, 2個小時後的波浪有90%的機率其變動量不超過4.5公分。

經由馬可夫鏈模式的分析可合理的考慮前期的波浪狀態, 授與不同的後期計算結果, 本文後續研究中以此為連續性品管的檢定標準。

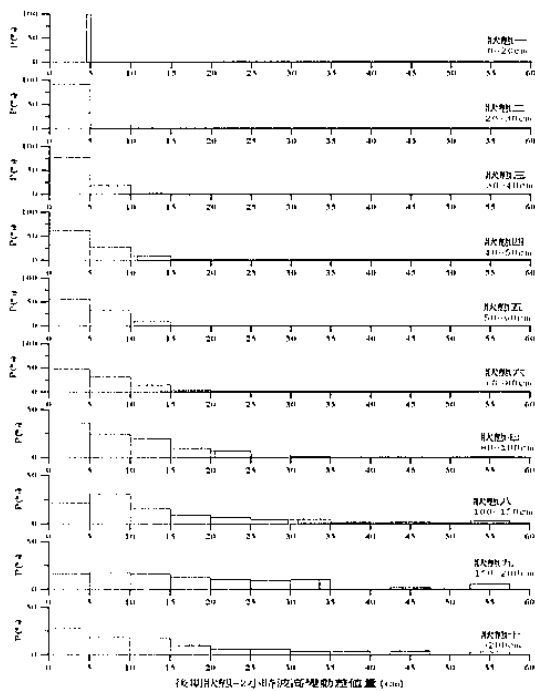


圖 1：修正型馬可夫鏈移轉機率矩陣圖

表 1：有義波高之移轉機率矩陣

	狀態一 0~20	狀態二 20~30	狀態三 30~40	狀態四 40~50	狀態五 50~60	狀態六 60~80	狀態七 80~100	狀態八 100~150	狀態九 150~200	狀態十 >200
狀態一 0~20	40.0	60.0	.0	.0	.0	.0	.0	.0	.0	.0
狀態二 20~30	5.0	75.0	19.5	.5	.0	.0	.0	.0	.0	.0
狀態三 30~40	.0	11.1	65.2	21.7	2.0	.0	.0	.0	.0	.0
狀態四 40~50	.0	.0	12.6	55.0	28.0	3.8	.2	.2	.0	.2
狀態五 50~60	.0	.0	1.2	19.6	54.8	23.8	.4	.1	.0	.1
狀態六 60~80	.0	.0	.0	1.4	23.1	67.0	7.6	.7	.1	.1
狀態七 80~100	.0	.0	.0	.0	1.4	31.5	86.3	9.9	.9	.0
狀態八 100~150	.0	.0	.5	.0	.0	1.5	12.3	74.5	10.8	.5
狀態九 150~200	.0	.0	.0	.0	.0	1.8	.0	35.7	58.9	3.6
狀態十 >200	.0	.0	.0	.0	.0	7.1	.0	10.7	.0	82.1

表 2：有義波高容許變動範圍

	稽延時間					
	2小時	4小時	6小時	8小時	10小時	12小時
狀態一: 0~20	4.5	6.2	7.6	9.2	9.4	9.8
狀態二: 20~30	4.8	8.6	9.6	12.5	15.1	19.6
狀態三: 30~40	8.4	12.1	16.5	18.9	21.4	23.4
狀態四: 40~50	10.8	14.7	18.3	19.4	19.7	19.7
狀態五: 50~60	10.5	14.0	16.2	17.9	18.2	18.4
狀態六: 60~80	13.8	18.0	19.9	20.6	21.2	21.8
狀態七: 80~100	20.5	25.0	29.0	30.1	30.1	32.2
狀態八: 100~150	31.2	30.5	38.0	36.4	37.1	38.0
狀態九: 150~200	33.6	45.0	47.5	49.4	48.9	59.2
狀態十: >200	36.4	64.2	66.8	72.9	76.5	80.3

(四) 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, 除了可判別資料正確性外, 對於儀器老化或毀損, 更可顯示出其偏差量。一般在波浪資料浮標或樁上安裝二套風速計, 降低儀器故障的影響, 並可相互檢測。在確定風速的正確性後, 利用風速與波浪能量之關係, 檢測波高記錄。

1. 一般觀測量間的關聯

本文選擇①兩支風速計同時所測得的平均風速、②3秒陣風與平均風速、③有義波高與有義週期、④氣溫與水溫以及⑤氣壓與氣溫來探討其關連性, 選擇各物理量同時有觀測資料的樣本, 套配不同的迴歸模式, 經比較相關係數與無因次的均方根誤差(root mean square error, RMSE)[(4)式]後發現以線性迴歸最能擬合分析樣本, 式中 x_i 為原始樣本, \hat{x}_i 為由迴歸曲線推估值, N 為樣本數。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \hat{x}_i}{x_i} \right)^2} \quad (4)$$

分析結果繪於圖 2~圖 6, 圖中亦繪製迴歸曲線的 95%信賴區間作為後續關聯性檢驗的上下限, 各回歸線的相關係數最高達 0.99, 最低為氣溫與氣壓之關係, 僅為 0.53, 至於均方根誤差均低於 20%, 在可接受之範圍內。

由圖 2 至圖 6 分析結果指出, 兩支風速計間的關係良好, 這一方面說明了觀測系統的穩定性, 另一方面也提昇了風速資料的可信度, 由

於 3 秒風速是由計算平均風速的同一筆風速時序求得，因此 3 秒風速與平均風速亦有良好的關係存在。一般而言，除湧浪外，大波高均伴隨較長的波浪週期產生，而在高雄大林埔外海大部份波浪週期均介於 3~5 秒之間，湧浪的成分較少，這也造成了圖 4 的線性迴歸關係，由圖中也可看出少數為由外地傳入的湧浪（波高小，週期大於 95%上限者），至於圖中週期較長且波高較大的資料則為去年賀伯颱風前期所帶來的湧浪。氣溫與水溫如圖 5 呈現線性關係，受到太陽照射及風的吹拂等因素的影響，分析結果顯示出部份時刻水溫均高於氣溫。

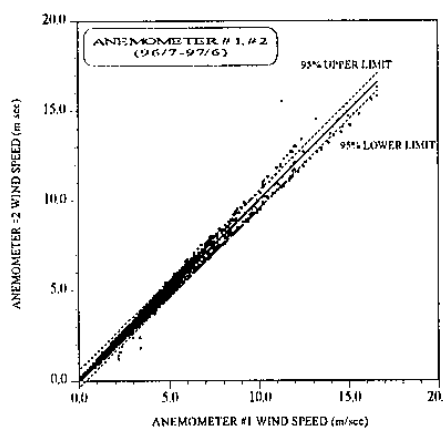


圖 2：兩支風速計平均風速關係圖

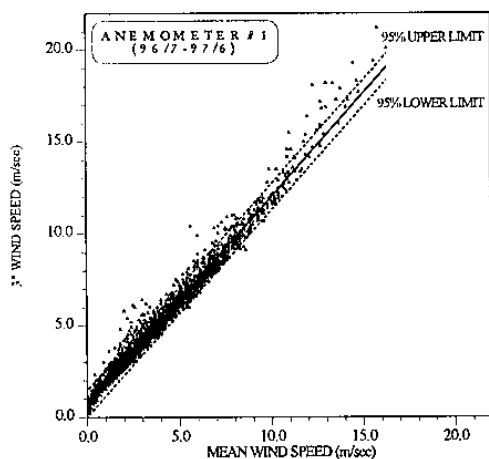


圖 3：平均風速與 3 秒陣風關係圖

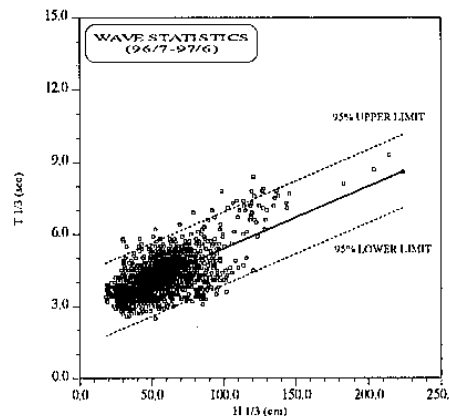


圖 4：有義波高與有義週期關係圖

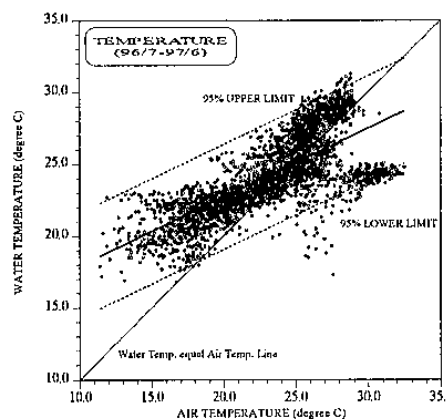


圖 5：氣溫與水溫關係圖

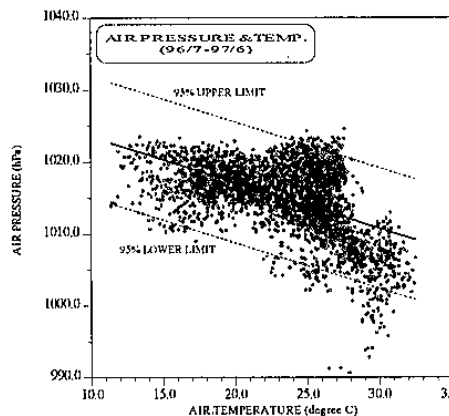


圖 6：氣溫與大氣壓力關係圖

2. 風浪間的關聯

Steele & Marks(1979)提出頻譜與風速間的線性模式，如(5)式所示。

$$S(f) = A(f) + B(f)U \quad (5)$$

式中 $S(f)$ 為頻譜能量， $A(f)$ 與 $B(f)$ 為係數， U 則

為風速，Steele & Marks 研究結果顯示當頻率介於 0.2~0.27Hz 時有最佳之風浪關係。Draper et al.(1981)提出頻譜能量應與 U^4 有較佳的關係，但分析結果仍證實線性模式較佳。Lang(1987)延續 Steele & Marks 的研究指出，0.2~0.27Hz 頻率之間的總能量與前四小時平均風速的平方有良好的關係存在，此結果目前為 NDBC 所引用。

本文以(5)式的線性模式套配風浪資料，分析結果顯示在高雄大林埔外海的風浪關係，介於 0.6~0.67Hz 頻率間的總能量與當時的平均風速有最佳的線性模式，如圖 7 所示，其相關係數為 0.70，均方根誤差為 17.8%。此結果可說明高雄大林埔地居台灣海峽內，在受限制的風域中，風的吹拂反映在較小週期的波浪能量上，還不足以形成較低頻的長浪，這可由當地有義週期僅為 3~5 秒看出，這也就說明了本文尋得圖 7 中某時刻平均風速與同時的波浪高頻能量間明顯的線性關係。

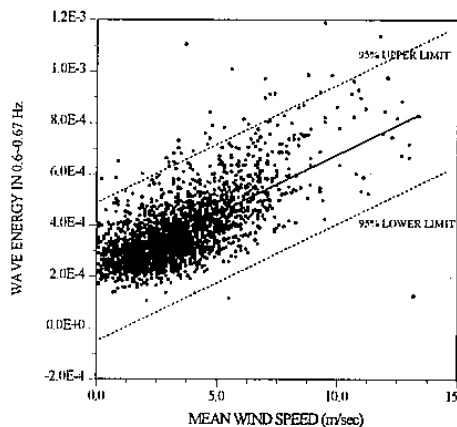


圖 7：平均風速與 0.6~0.67Hz 間頻譜能量關係

三、品管流程

品管流程如圖 8 所示，由合理性的檢驗後接續連續性的檢驗，再執行關聯性檢驗。

連續性檢測方法是判斷新的一筆觀測資料與上一筆觀測記錄差值是否在對應狀態的容許變量範圍內，當資料有遺缺或上一筆記錄未通過檢測時則繼以稽延 4 小時或更長稽延時刻之連續性檢驗為之。

關聯性的檢驗方法是取在該時刻對應關聯

的物理量代入迴歸關係式，求得欲品管量的 95%信賴區間，若欲品管量居於 95%信賴區間內則通過檢定。然關聯性品管的順序亦為重要之一環，由於裝設的兩支風速計可互相檢驗以確保風速資料的正確性，此為關聯性檢驗的第一項目；風速資料檢測完畢後視為正確，繼而再利用風速與 0.6~0.67Hz 間頻譜能量關係檢測波浪譜，繼而以波浪譜來檢測波浪統計值，本文分析發現大林埔地區海域頻譜能量與波高之關係為 $H_y = 3.55\sqrt{m_0}$ ，進行品管時若 $2.84\sqrt{m_0} \leq H_y \leq 4.26\sqrt{m_0}$ ，則視 H_y 通過品管。依序再由 $H_y \sim T_y$ 關係檢測有義週期，由氣溫資料檢測水溫與氣壓資料。本文將關聯性品管順序整理條列如下，

- (1)兩支風速計間互相進行檢測
- (2)以平均風速檢測 3 秒陣風
- (3)以平均風速檢測波浪頻譜(0.6~0.67Hz)
- (4)以頻譜能量檢測有義波高 H_y
- (5)以 H_y 檢測 T_y
- (6)以氣溫檢測水溫觀測值
- (7)以氣溫檢測氣壓觀測值

品管系統流程架構建立後，整個檢測流程皆可以計算機自動控制，但不可避免地仍必須進行人工檢測的工作，譬如颱風等異常氣候發生時，觀測資料可能無法通過相關的檢測項目，必須人工檢測並調閱附近測站或相關資訊來判斷欲保留或刪除該筆可疑資料，因為極端資料得之不易，但若保留錯誤過大或過小記錄也會造成日後分析的困擾，因此人工檢測的工作為不可或缺的一環。另一方面，也可由人工檢測看是否發生長時間觀測記錄不正確的現象，依此作為監測儀器是否正常運作的資訊，及早發現故障問題並予以修復以確保觀測工作的順利進行。

標記(mark)是對有問題資料的一種基本處理方法，品管前必須先進行資料處理的工作，若發現有缺失資料則標記 A，基本檢驗的檢驗過程中，若沒有通過合理性檢測，刪除該筆資料，並在該處標記 B 數字，同理若不通過連續性檢測則標記 C 數值，最後以標記 D 來代表未通過關連性檢測者。事後分析資料時即可根據

各種標記數值來了解資料缺乏的原因，波高與週期末通過檢測的標記列於表 3 中。

表 3：波高與週期末通過品管的標記

項目	標記 A	標記 B	標記 C	標記 D
波高	999.9	999.8	999.7	999.6
週期	99.9	99.8	99.7	99.6

四、結論與建議

建立完善的資料品管系統，可減少資料的錯誤或是發現珍貴資料的存在。本文探討自動品管的問題，以合理性、連續性再到關聯性的檢測作為品管流程，採用高雄大林埔海域資料為分析對象，獲得以下的結論：

- (1)連續性品管方面，修正馬可夫鏈模式，考慮前期狀態來訂定下階段變動量得較合理之結果。
- (2)關聯性品管方面，利用線性迴歸建立數種物理量之間的關聯性；至於風浪之間，本文發現介於 0.6~0.67Hz 間的波譜能量與該時平均風速有良好的關係，反映了當地的風浪特性。另外，求出頻譜與波浪統計值的關係為 $H_{\lambda} = 3.55\sqrt{m_0}$ ，作為由波譜檢測統計值的依據。
- (4)制訂出整體自動品管工作的流程，及細部的關聯性品管順序。

人工品管流程的制定為下一階段所欲探討的主題。

謝誌

本文承蒙中央氣象局及國科會補助研究經費；高雄環保局提供南星觀測樁觀測資料，特此一併致謝。

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編彙，交通部中央氣象局，189 頁-221 頁。

2. 鄧中柱(1993)"作業化資料浮標之評介"，海象分析與預報研討會論文編彙，交通部中央氣象局，33 頁-45 頁。
3. Draper, M. K., H. A. Gamber and T. A. Watkins (1981) "A Statistical Approach to Modelling Ocean Waves Using Buoy Data," in *Proc. Oceans 81*. New York: Marine Tech. Soc. And IEEE, pp. 1189-1191.
4. Lang, N. C. (1987) "An Algorithm for the Quality Checking of Wind Speeds Measured at Sea Against Measured Wave Spectral Energy", *IEEE Journal of Oceanic Engineering*, Vol. OE-12, No.4, pp. 560-567.
5. Steel, K. E. and G. E. Marks (1979) "Detection of NDBO Wave Measurement Systems Malfunctions," in *Proc. Oceans 79*. New York: Marine Tech. Soc and IEEE, pp.226-236.

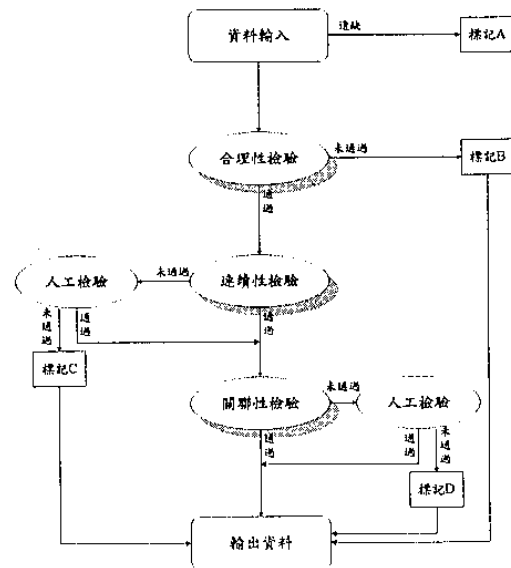


圖 8：品管系統流程圖