

壹、緒論

一、研究背景

資料品質對資訊化的社會而言，扮演著非常重要的角色（Chen, Chi, & Wolfe, 2005; Srinivasan, 1999; Strong, 1997; Svanks, 1998; Wang, Stroey, & Firth, 1995）；這對於以非常依賴數據資料的氣象領域而言，更是如此。氣象觀測數據不僅是氣象預報作業的關鍵依據，也是氣候資料建立的重要基礎。除此之外，長期累積下來的觀測數據也常被引用於綠色科學研究與各式經營管理。Jeffrey、Carter、Moodie 與 Beswick（2001）指出，多數使用者所引用的氣象數據大多來自於政府單位的長期記錄，且完全相信數據是準確的或從未察覺其中的不合理性，導致研究結論或作業決策可能有所失當。由此可知，氣象觀測數據品質不單是影響氣象預報之準確性，舉凡任何據以為策之研究或管理皆在其影響範圍之內。

在氣象領域中，資料品質是一項重要的研究議題；其探討如何利用現有技術產生可靠的數據集，以提供正確的氣象相關資料給氣象預報中心或外界組織來應用（Feng, Hu, & Qian, 2004; González-Rouco, Jiménez, Quesada, & Valero, 2001; Klein Tank et al., 2002; Manton et al., 2001; Reek, Doty, & Owen, 1992; Vincent, Zhang, Bonsal, & Hogg, 2002）。有鑑於氣象數據品質的重要性及其影響範圍之大，許多國家或地區廣泛設立氣象觀測站以取得更完整的氣象資料，並將這些大量資料透過資訊技術輾轉傳輸及彙集至中心站，同時也發展各式的檢驗準則，以確保這些氣象數據集是可靠且準確的。

二、研究動機與目的

本研究在觀察中發現，目前有許多檢核方式來確保氣象資料之品質，例如範圍測試、改變率測試、連續無變化測試等；這些方法主要是以物理特性為主，並且著重在最後的資料來進行檢核。然而，若從全面品質管理的觀點來看，氣象資料的產生與呈現是有過程性的（process-oriented），不應僅著重於結果面（results-based）（Chen, Chen, & Chen, 2012; Vassiliadis, Bouzeghoub, & Quix, 2000），例如，從感測器蒐集資料、將資料儲存至資料庫，以及人工校正等過程中，是否也受其他因素干擾，進而影響到資料最後呈現出來的品質？若從資訊科技（Information Technology, IT）角度來看，氣象資料的產生與應用之過程實也是一種資訊系統的展現；具體而言，其為處理大量資料的資料倉儲系統（Data Warehousing, DW）

類型，因此自有一套相關的資料品質保證要求。而在這些跨領域的考量下，當我們在檢核氣象資料時，是否也能以較全面性之角度來進行之？

為解決上述問題，本研究將以氣象領域現有的品質保證學理方法為基礎，嘗試加入資訊科技與流程專注的角度，進一步提出「氣象資料品質檢核綜效架構」，以協助關鍵人員較全面地掌握氣象資料品質。此架構是以資料倉儲系統的資料品質理論（Ballou & Pazer, 1985; Chen et al., 2005）為學理依據，建立氣象資料過程品質保證的四個品質構面——即準確性（accuracy）、完整性（completeness）、一致性（consistency）及時效性（timeliness），並以本研究案例的實際需求而定義出相關的度量操作內容。此外，本研究運用資訊科技中的統一塑模語言（Unified Modeling Language, UML）來設計並開發一個應用系統，除了用以驗證所提的架構的可行性，也為巨量的氣象資料檢核過程中，加以自動化並提升效率。在系統展示方面，本研究以實際的案例，即中央氣象局淡水觀測站，說明系統功能與使用情形。最後，本論文針對目前的研究成果與限制加以探討，並提出未來的研究方向。

在接下來的文章內容，其結構如下：第貳節為文獻探討，包括資料倉儲及資料品質領域的學理文獻的介紹與整理，以及現有在氣象領域之資料檢核方式的探索與評論；第參節為研究方法與設計，即氣象資料品質檢核綜效架構的說明；第肆節為系統實作與展示；第伍節為分析與討論；第陸節則針對本研究做一總結。

貳、文獻探討

一、資料倉儲

資料倉儲的概念起源於 1960 年代，隨著科技的進步以及企業資訊系統的資料量大幅增加，近年來資料倉儲的相關應用與發展愈來愈受到重視。大多數企業使用資料倉儲為資料庫核心，再建置成各種不同用途的分析式的應用系統，如財務金融分析系統、顧客關係管理系統、決策支援系統等（Sen, Sinha, & Ramamurthy, 2006）；而在氣象領域中，各國的氣象中心利用資料倉儲來儲存氣象資料也漸為普遍（Huang, Yuan, Liu, & Shang, 2010; Musa et al., 2003）。資料倉儲是一常見的商務資料分析之重點業務。根據先驅學者 Inmon（1996）的定義，資料倉儲系統是以主題導向，經過整合，並考慮時間關聯的資料集合，其用以協助於管理決策過程。從此一定義可看出，資料倉儲的特性即是累積巨