應用隨機森林於對流尺度資料同化系統反演水物及其對分析與預報之影響

Impacts of Random Forest-Based Hydrometeor Retrieval on Analysis and Forecasting in a Convective-Scale Data Assimilation System

蔡金成[1,2] (Tsai C.-C.) 蔡子衿[1] (Tsai T.-C.) 劉正欽[1] (Liu C.-C.)

[1]中央氣象署科技發展組 [2]國立臺灣大學

[1]Technology Development Division, Central Weather Administration

[2]National Taiwan University

摘　要

隨著人工智慧技術的蓬勃發展，本研究利用隨機森林(Random Forest)演算法，透過逐時更新訓練模型以反演水物(Hydrometeors)，並與現行WRFDA系統所採用之雷達回波反演水物方法進行比較，進而評估其對三維變分(3DVAR)與混合式系集變分同化系統(Hybrid 3DEnVar)在分析及預報效果之影響。本研究針對鋒面及午後降水個案，分別進行49次預報之統計分析。

兩個案之實驗結果顯示，引入隨機森林演算法反演水物並同化到模式，能有效修正中高層模式回波分析場顯著低估，並改善初始一小時回波預報場的低估、提升整體準確度。針對6小時定量降水預報的校驗結果顯示，隨機森林方法能顯著提升第一小時降水在20毫米門檻以上的預報能力且具備統計顯著性。

關鍵字：資料同化、雷達回波反演、隨機森林

Abstract

With the rapid development of artificial intelligence technology, this study utilizes an hourly updated Random Forest algorithm to retrieve hydrometeors. The performance of this approach is compared against the radar reflectivity-based hydrometeor retrieval method currently employed in the WRFDA system, aiming to evaluate its impacts on the analysis and forecasting efficacy of the 3DVAR and Hybrid 3DEnVar data assimilation systems. Statistical analyses were conducted based on 49 forecasting cycles for both a frontal system and an afternoon thunderstorm case.

Results from both cases show that incorporating the Random Forest algorithm effectively mitigates the significant underestimation of radar reflectivity in the mid-to-high-level model analysis fields. Furthermore, it alleviates the underestimation in the initial 1-hour reflectivity forecasts and enhances overall accuracy. Regarding the validation of the 6-hour Quantitative Precipitation Forecast (QPF), the Random Forest method yields a statistically significant improvement in forecasting capability for the first-hour rainfall exceeding the 20 mm threshold.

Key words : Data Assimilation, Radar Reflectivity Retrieval, Random Forest