應用 CorrDiff 模型於 MLWP 系集預報

Application of the CorrDiff Model to MLWP Ensemble Forecasting

蘇胤瑞 [1] (Su Y.-R.)　王志嘉 [1] (Wang C.-C.)　張惠玲 [1] (Chang H.-L.)

[1] 中央氣象署

[1]Central Weather Administration

摘　要

傳統基於物理方程之數值天氣預報（NWP）動力系集模型，受限於龐大的計算資源需求，長期難以同時兼顧高解析度與足夠的系集成員數。近年資料驅動之機器學習天氣預報（MLWP）快速發展，其中 FourCastNet 3（FCN3）在運算效率上較傳統模式提升數十倍，並能有效掌握大尺度環流特徵；然而，如何將其粗網格預報精準轉化為在地化細尺度天氣資訊，並有效定量其不確定性，仍是當前氣象科技亟待突破的課題。本研究建立全自動化作業流程，將生成式擴散模型（CorrDiff）降尺度技術應用於 FCN3 多成員系集預報，將粗解析度預報場細緻化為公里級、對流尺度之精細網格產品。並以 2025 年 5 月 15 日至 6 月 15 日期間 32 個起始時間之個案進行階段性評估，結果顯示：在溫度預報誤差方面，不論 CorrDiff 降尺度前後，MLWP（FCN3）之均方根誤差與平均絕對誤差均明顯低於傳統 NWP，尤其在陸地區域差異更為顯著，且誤差並未隨預報時間明顯增長；在系集離散度方面，CorrDiff 降尺度後之系集預報可大致維持降尺度前的離散度特徵，於第 1、2 週略呈離散度不足現象，第 6 週則能較佳反映預報不確定性；在降水預報技術方面，逐週累積降水 FSS 評估顯示，中低門檻（10 至 30 毫米/週）以 NPM 表現最佳，高門檻（80 至 150 毫米/週）則以 PM 較具競爭力。上述結果顯示，此技術架構能以極低計算資源，於極短時間內產出過去仰賴超級電腦密集運算方能達成的高解析度系集預報，可為建置新一代細尺度機率預報作業，以及防救災、水利與農業等跨領域決策支援，奠定堅實的科學與技術基礎。

關鍵字：生成式 AI 降尺度、CorrDiff、機器學習天氣預報、系集預報、機率預報

Abstract

Traditional dynamical ensemble NWP systems face severe computational bottlenecks, limiting simultaneous high resolution and sufficient ensemble size. Machine learning weather prediction (MLWP), represented by FourCastNet 3 (FCN3), achieves dozens-fold speedups over conventional models while capturing large-scale circulation, but converting its coarse output into localized, fine-scale information with quantified uncertainty remains challenging. This study establishes a fully automated workflow applying CorrDiff, a generative diffusion downscaling model, to the FCN3 ensemble, producing kilometer-scale, convection-permitting products, with staged evaluation conducted for 32 initializations from 15 May to 15 June 2025.

Results show that, before and after downscaling, MLWP (FCN3) temperature RMSE and MAE are substantially lower than NWP, especially over land, with errors not growing markedly with lead time. Ensemble spread after downscaling largely preserves pre-downscaling characteristics, showing slight underdispersion in weeks 1-2 and better-calibrated spread by week 6. For precipitation, weekly FSS shows NPM performs best at low-to-moderate thresholds (10-30 mm/week) while PM is more competitive at high thresholds (80-150 mm/week). These findings show the framework can generate high-resolution ensemble forecasts, once dependent on intensive supercomputing, at minimal cost, supporting fine-scale probabilistic forecasting and disaster-prevention decisions.

Key words : generative AI downscaling, CorrDiff, machine learning weather prediction, ensemble forecast, probabilistic forecasting