

中央氣象局全球預報系統系集變分混合

資料同化系統之發展與作業化評估

周俊宇¹ 趙子瑩¹ 黃清勇² 楊明仁² 陳雯美¹ 陳登舜¹

中央氣象局氣象科技研究中心¹

國立中央大學大氣科學系²

摘 要

中央氣象局全球預報系統(Central Weather Bureau Global Forecast System；CWB/GFS)於2010年引進美國國家環境預報中心(NCEP)所開發的網格統計內差系統(Gridpoint Statistical Interpolation；GSI)後，持續對系統進行改善及更新維護，而現行作業GSI資料同化系統為三維變分資料同化系統(3-Dimensional Variational Data Assimilation；3DVAR)；但三維變分資料同化系統存在著背景場誤差(background error)無法隨時間變化的最基本弱點，因此我們使用兼具動態背景場誤差及傳統變分系統穩定性的系集變分混合資料同化系統(Hybrid)取代現階段的三維變分分析系統(3DVAR)，期望取得更佳的分析及預報結果。Hybrid資料同化系統於上年度(102)已完成建置以及初步測試，並且取得預期的成效，而接下來則需要更進一步的檢測此系統對現行資料同化系統的影響，並且評估實際上線作業的可能性；為了達到此目的，我們進行了接下來一連串的平行測試：a.將ECbogus及其他作業用觀測資料納入平行測試，確保與現行作業系統在觀測資料上的一致性。b.新增NMC-method 添加系集擾動之運用，改進先前的混合法資同化系統，增加additive inflation的機制。c.設計三組實驗，分別為現行作業用3DVAR架構下的資料同化系統(ngsiEC)、混合法資料同化系統並添加NMC-method 系集擾動(enkfNMC50)和混合法資料同化系統未添加系集擾動(enkfhyb50)，個別進行冬夏各一個月的平行測試。將藉由這三組實驗的平行測試結果，評估是否具有作業化的能力。

關鍵字：EnKF、3DVAR、Hybrid

一、前言

中央氣象局全球預報系統自2010年7月所採用的全球資料同化系統，即格點統計內差(GSI)系統，取代自2003年開始作業之波譜統計內差系統(SSl；Parrish and Derber, 1992)。原先SSl資料同化系統因屬2003年舊版本，其版本未與時更新，並不具備同化許多新的非傳統觀測資料(如衛星、雷達及剖風儀等)之能力。中央氣象局為解決此迫切的問題，經過仔細評估後，決定引進NCEP全球數值預報系統，於2007年5月更新的新一代資料同化系統(GSI)。此GSI系統不但在分析方法上對於背景場誤差具備較接近真實大氣的處理方式，同時也具備同化更多非傳統觀測資料的能力。

對於資料同化方法而言，SSl與GSI本質上均為三維變分分析，且都在模式定義的sigma面上進行分析，其中GSI則還具備在hybrid垂直座標分析的選項。SSl是在波譜空間(spectral space)上進行分析，背

景場誤差在波譜空間上定義，為均向(isotropic)及均質的(homogeneous)，緯度的變化是藉由一統計之緯度變化參數來控制，在分析質量場與風場的平衡關係時直接應用線性平衡方程(linear balance equation)來維繫。相對而言，GSI的分析是在物理網格空間(grid-point physical space)上進行分析，其主要優勢在背景場誤差直接在物理格點上定義，目前為緯度及高度的函數，可以是非均向性(anisotropic)及非均質性的(inhomogeneous)，與真實大氣特性較為接近。各分析變數間的平衡關係除了經由統計關係維繫外，還建置了正切線性正模約束(Tangent Linear Normal-Mode Constraint；TLNMC)的強約束(strong constraint)項，進行變數平衡關係的調合，以減低因分析之質量場與風場間的不平衡而影響預報的表現程度。

NCEP於2012年5月正式上線作業版本的GSI系統新增了混成(Hybrid)功能，它可以經由系集產品提供

給背景誤差另外一個重要分項。系集背景誤差協方差，透過變分資料同化求得最佳解的方式，使得所求得的最佳分析場同時具有傳統變分資料同化以及EnKF資料同化的雙重優點，以達成混成的目的。

Hybrid資料同化系統結合了變分資料同化與系集卡爾曼濾波同化方法兩者。主要以變分資料同化系統為架構，將背景誤差協方差由原本僅使用靜態背景誤差協方差，置換為同時包含靜態以及利用EnKF所計算出的系集背景場誤差協方差。因此，一個完整的Hybrid資料同化系統必須結合EnKF系統與已擴充Hybrid相關項於價值函數的變分資料同化系統。而中央氣象局現階段已建置完Hybrid資料同化系統，在101年度完成了單點測試並得到預期的表現，其後，在上一年度(102)開始進行部分的平行測試，都取得了預期的成效，而接下來的發展目標則是更進一步的測試且改善Hybrid資料同化系統。

二、研究方法

(一) EC bogus以及其餘作業用觀測資料之運用

Hybrid資料同化系統相關實驗中，必須要確保觀測資料的一致性，避免與中央氣象局實際作業所用的觀測資料產生落差。其中最重要的部份為CWB/GFS資料同化過程中一直當作觀測資料使用的EC bogus資料，在CWB/GFS的資料同化系統中，加入EC bogus能夠給予預報結果明顯改善。而EC bogus的作法是擷取ECMWF的分析場或預報場，用於CWB/GFS流程中post或major的資料同化步驟。而我們所測試的Hybrid資料同化系統實驗，是建基於CWB/GFS的post預報流程，並且納入EC bogus的資料。

此外，由於衛星資料同化的運用在中央氣象局的分析流程也有持續進展，因此本次的相關實驗也調整了設定，包含了CWB/GFS作業流程中所使用的衛星觀測資料、部份未同化的AMSUA衛星資料，以及今年中央氣象局加入的IASI衛星資料；如此才能盡量與作業環境相當，也較能進一步評估Hybrid資料同化系統對於CWB/GFS的影響。

(二) NMC-method 增加系集擾動之運用

NCEP的Hybrid資料同化流程中是有包含additive inflation的機制，但現階段CWB/GFS測試下的Hybrid實驗中並沒有此項機制；而延續先前的研究分析，發現現階段的Hybrid資料同化系統中所使用的系集組員，沒有辦法產生足夠程度的系集擴散(Spread)；因此我們增加了NMC-method (National Meteorological Center)添加系集擾動(perturbation)的步驟，希望能夠改善系集擴散不足的現況。

NMC-method最主要的概念，是使用48小時的預報場與24小時的預報場相減，得到該時間背景場誤差的特徵；而實際運用，首先需要準備一整年的每一天所有預報時間之48小時預報減24小時預報，建立成資

料庫，接著從資料庫擷取該預報時間前後各40天的資料，隨機排序後再將背景場誤差的特徵加入每一組系集組員中，達到增加擾動的效果。以本計劃為例，是選用2012年的資料作為資料庫，每六小時進行一次預報，一天有四個預報時間，擷取前後各40天的資料，也就是共有320筆資料；最後將擷取出的320筆資料重新隨機排序，再依序選出36筆資料加入至36組系集組員當中。

三、實驗設計

我們進行三組平行測試實驗，第一組是使用現行CWB/GFS為3DVAR架構下的資料同化系統，並包含與現行作業相同的觀測資料設定，包含EC bogus資料，以及其他的作業用衛星資料；實驗代稱為ngsiEC。第二組則是Hybrid資料同化系統，使用36組EnKF系集成員，並加入NMC-method擾動，調整靜態與系集背景誤差之權重參數(beta1_inv)為0.5，觀測資料皆與作業相同；實驗代稱為enkfNMC50。第三組實驗與enkfNMC50類似，但並沒有增加NMC-method擾動，主要是用以檢視增加NMC-method擾動步驟對Hybrid資料同化系統的影響；實驗代稱為enkfhyb50。三組實驗的各項設定比較參照列表(表3.1)。

將此三組實驗個別進行冬季(2012/12)和夏季(2013/07)各一個月的平行測試，互相比較實驗結果以及接下來的分析。

實驗代稱	ngsiEC	enkfNMC50	enkfhyb50
DA系統	3DVAR	Hybird	Hybrid
系級組員	none	36組	36組
beta1_inv	none	0.5	0.5
EC bogus	有	有	有
衛星資料	同作業	同作業	同作業
添加擾動	none	NMC	none

表3.1 三組實驗各項設定比較。

四、實驗結果分析

(一) 冬季結果分析

首先分析三組實驗在冬季的結果與預報表現；我們選擇較具有代表性的分層和資訊，分析在500mb上的高度場距平相關、850mb上的溫度場均方根誤差和300mb上的風場均方根誤差的五天預報時序變化表現(圖4.1)；結果顯示北半球三組實驗的差異甚小，預報表現在各層是各有優劣，整體來看北半球兩組Hybrid的實驗比起ngsiEC的預報表現為持平。南半球的部份則是以enkfNMC50的預報表現最好，在500mb上的高度場距平相關，相較於ngsiEC有0.005左右的提升，溫度場以及風場的均方根誤差也較低；enkfhyb50的預報表現介於兩個實驗之間，比起ngsiEC的結果略有提升，但是不及enkfNMC50的表

現。分析在熱帶區域的表現(圖4.2)，由於高度場距平相關在熱帶區域的較無參考性，故分析溫度場和風場；而溫度場均方根誤差的情形三組實驗差異不大，風場則是以enkfNMC50的均方根誤差最小。

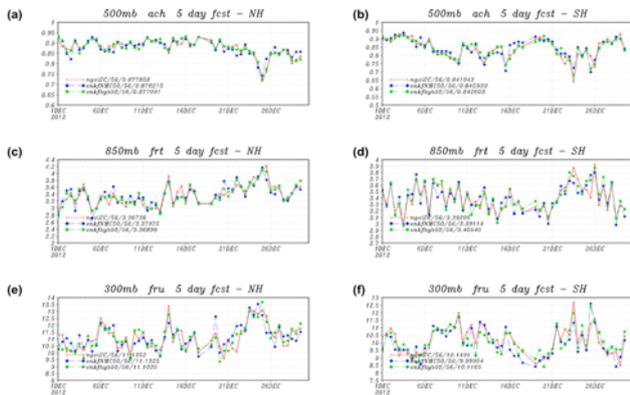


圖4.1 三組實驗(ngseEC、enkfNMC50和enkfhyb50)一個月期間(2012年12月)五天預報的500mb高度場距平、850mb溫度場的均方根誤差和300mb風場時間序列圖(由上而下)。a、c、e為北半球的結果；b、d、f為南半球的結果。

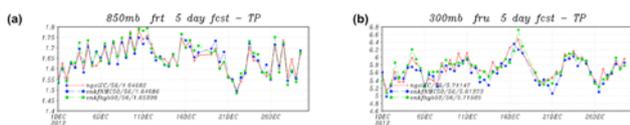


圖4.2 為三組實驗在熱帶地區的五天預報850mb溫度場及300mb風場均方根誤差之時序變化情形。

接著比較在南北半球冬季的每一日預報的月平均表現(圖4.3)，分析在500mb上的高度場距平相關；北半球部分的結果是 enkfhyb50相較於 ngseEC表現略差，而 enkfNMC50則是與 ngseEC持平，其中二、三日的預報略好；南半球則明顯看出兩組Hybrid的實驗相較於 ngseEC都有所改善，前四日的95%信心區間都顯示高度場距平相關較大，不過第五日的信心區間則沒有辦法完全通過，而enkfNMC50改善的程度較多。

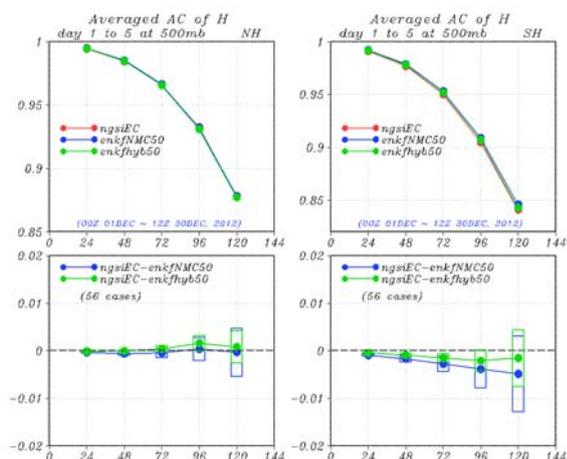


圖4.3 為冬季(2012/12)在北半球(左)及南半球(右)在

500hPa 高度場距平相關各預報時間的平均值與差異圖，橫軸為預報時間。上圖紅色實線為ngseEC，藍色實線為enkfNMC50，綠色實線為enkfhyb50；下圖藍色及綠色實線為ngseEC減去enkfNMC50、enkfhyb50的差異值，方框表示95%信心區間。

同時分析三組實驗在冬季垂直各層南北半球的預報表現(圖4.4)；北半球的部分，五天預報高度場距平相關的結果，顯示 enkfNMC50實驗與 ngseEC的差異甚小，大部分垂直分層都為持平的情形；而溫度場及風場的均方根誤差也是類似情形，各實驗間的差異不大。不過注意到200mb至100mb的分層間，高度場距平相關的預報結果是enkfNMC50表現略差，但是同一區間的風場及溫度場均方根誤差，則是 enkfNMC50較小。分析南半球的結果就明確顯示，enkfNMC50無論是在高度場距平相關或是溫度場及風場的均方根誤差的表現，都是優於 ngseEC實驗。因此使用Hybrid資料同化系統的 enkfNMC50實驗在冬季的預報表現，顯示北半球持平，而南半球能提供正面效益的結果。

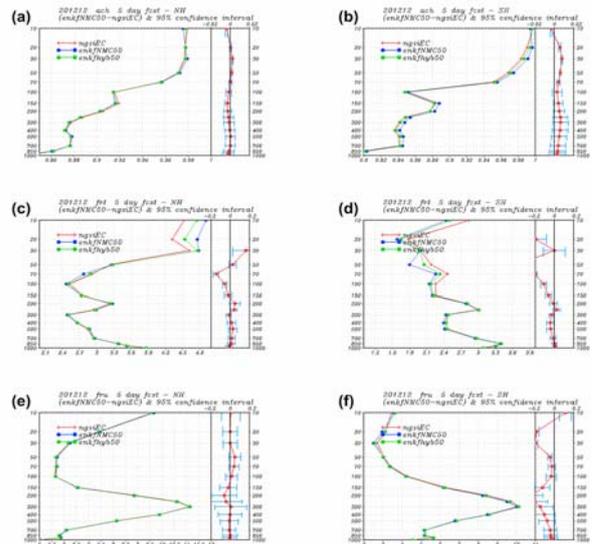


圖4.4 為三組實驗垂直各層的5天預報高度場距平相關、溫度場及風場的均方根誤差。圖a、c、e依序為北半球的高度場距平相關、溫度場、風場的均方根誤差垂直分佈。圖b、d、f依序為南半球的高度場距平相關、溫度場、風場的均方根誤差垂直分佈圖。右側則是代表enkfNMC50與ngseEC實驗差值之95%信賴區間。

熱帶地區的預報表現(圖4.5)，enkfNMC50在風場的均方根誤差較小，而溫度場的部分比起 ngseEC則是在各層互有優劣，而高度超過30mb因為接近模式層頂，會受到模式調整影響。

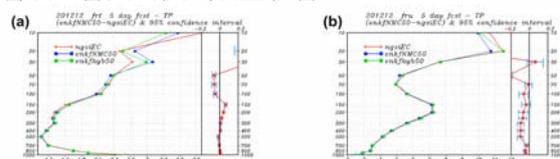


圖4.5 為三組實驗在熱帶地區垂直各層的5天預報溫度場及風場的均方根誤差。圖a為溫度場的均方根誤差，圖b為風場的均方根誤差；右側則是代表enkfNMC50與ngsiEC實驗差值之95%信賴區間。

整體看來三組實驗在冬季的表現，於北半球的部分差異不大，Hybrid資料同化系統並未有顯著的改善，enkfNMC50比起ngsiEC的預報表現算是持平；但是在南半球的部分則能看到enkfNMC50以及enkfhyb50相較ngsiEC都能有改善的結果，顯示Hybrid資料同化系統可以給予明顯的正面效益，而同時enkfNMC50實驗所提升的改善程度較enkfhyb50更大，意味著加入NMC-method擾動系集組員，可以使整體的預報有更好的結果。

(二) 夏季結果分析

接下來是分析三組實驗在夏季(2013/07)的表現；同樣選擇較具有代表性的分層和資訊，分析在500mb上的高度場距平相關、850mb上的溫度場均方根誤差和300mb上的風場均方根誤差的五天預報時序變化表現(圖4.6)；在北半球的部分則能看到enkfNMC50的表現是較ngsiEC的預報結果有明顯的提升，而從時序變化的情形除了數天較差之外，大部分天數的五天預報高度場距平相關都是高於ngsiEC，整體平均提升0.009；而enkfhyb50的在北半球區域表現也是優於ngsiEC，但是不及enkfNMC50。而溫度場及風場的表現也是enkfNMC50在北半球的表現最佳。南半球區域的五天預報高度場距平相關則是enkfhyb50的表現是最好的，不過enkfNMC50相較ngsiEC亦是有正面的提升，只是不如北半球區域的顯著，溫度場及風場的表現類同。

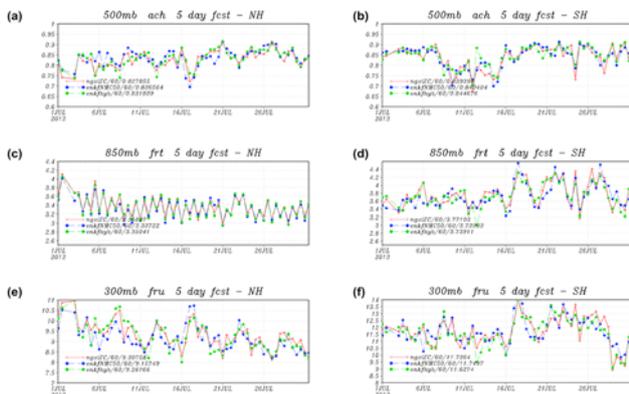


圖4.6 三組實驗(ngsiEC、enkfNMC50和enkfhyb50)一個月期間(2013年07月)五天預報的500mb高度場距平、850mb溫度場的均方根誤差和300mb風場時間序列圖(由上而下)。a、c、e為北半球的結果；b、d、f為南半球的結果。

熱帶地區的表現與北半球的預報結果相似(圖4.7)，enkfNMC50在風場及溫度場的均方根誤差都是

三組實驗中最低。

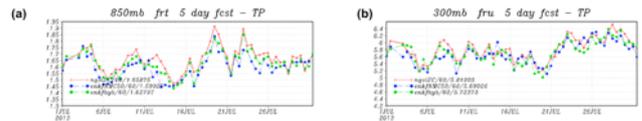


圖4.7 為三組實驗在熱帶地區的五天預報850mb溫度場及300mb風場均方根誤差之時序變化情形。

同樣比較在南北半球夏季的每一日預報的月平均表現(圖4.8)，分析在500mb上的高度場距平相關；在北半球區域可以明顯看到enkfNMC50的一至五天的預報，都比ngsiEC有更好的預報表現，而且皆能通過95%的信心區間，代表有足夠的信心提供正面效益，而enkfhyb50在北半球雖然不及enkfNMC50，但一至五天預報所呈現的趨勢依然是較優於ngsiEC。而在南半球的區域，兩組Hybird實驗的預報也是較ngsiEC表現更佳，不過一至四天的預報結果，顯示enkfNMC50的表現最好，也都能通過信心區間，但是在第五日的預報表現則是enkfhyb50的結果超越enkfNMC50的結果，所以對應先前水平時序變化圖的結果，也是enkfhyb50在南半球有較好的表現。

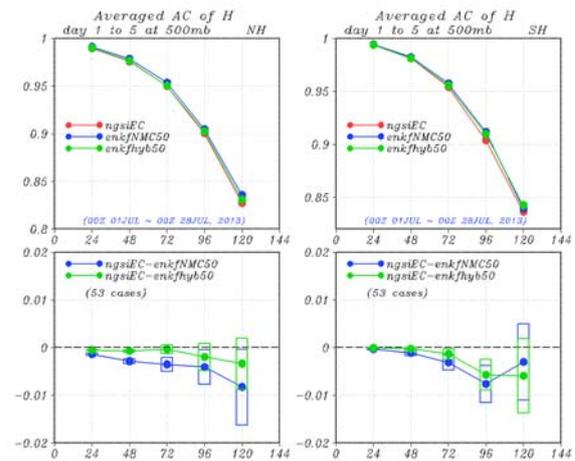


圖4.8 為夏季(2013/07)在北半球(左)及南半球(右)在500hPa高度場距平相關各預報時間的平均值與差異圖，橫軸為預報時間。上圖紅色實線為ngsiEC，藍色實線為enkfNMC50，綠色實線為enkfhyb50；下圖藍色及綠色實線為ngsiEC減去enkfNMC50、enkfhyb50的差異值，方框表示95%信心區間。

接下來分析三組實驗在夏季垂直各層南北半球的預報表現(圖4.9)；在北半球區域的部分，enkfNMC50在150mb至最底層的五天預報高度場距平相關都是相較ngsiEC有更好的結果，也都能通過信賴區間，但是在100mb附近則是有突然下降的偏差；對照溫度場及風場在此高度的變化情形，enkfNMC50的風場均方根誤差在垂直各層都是小於ngsiEC，而溫度場均方根誤差則是在200mb至100mb間較大，但是變化沒有高度場距平相關明顯，因此原因還尚需要更進一步的釐清。觀察南半球區域的垂直變化，實驗間的差異則較小，enkfNMC50在五天預報高度場距平相關及溫度

場均方根誤差的表現與ngsiEC持平，風場均方根誤差平均雖然較小，但並未能完全通過信賴區間。

熱帶區域的預報表現(圖4.10)，在溫度場的部分在100mb以下的高度差異不大，enkfNMC50的均方根誤差略小；風場的垂直分佈則是150mb以下的高度是enkfNMC50的均方根誤差較小，但100mb有較大的偏差。

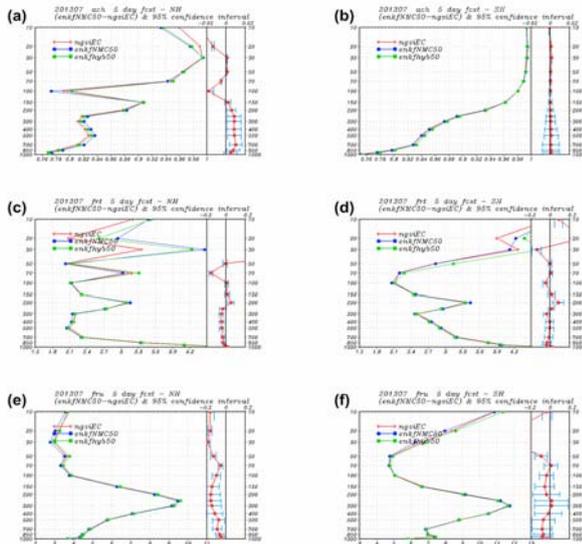


圖4.9 為三組實驗垂直各層的5天預報高度場距平相關、溫度場及風場的均方根誤差。圖a、c、e依序為北半球的高度場距平相關、溫度場、風場的均方根誤差垂直分佈。圖b、d、f依序為南半球的高度場距平相關、溫度場、風場的均方根誤差垂直分佈圖。右側則是代表enkfNMC50與ngsiEC實驗差值之95%信賴區間。

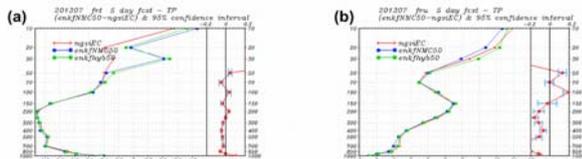


圖4.10 為三組實驗在熱帶地區垂直各層的5天預報溫度場及風場的均方根誤差。圖a為溫度場的均方根誤差，圖b為風場的均方根誤差；右側則是代表enkfNMC50與ngsiEC實驗差值之95%信賴區間。

而圖4.9中我們能發現，enkfNMC50的高度場距平相關在北半球100mb處附近有突然下降的偏差，但其他層並未有這樣的現象；針對此現象，我們比較100mb處的高度場距平相關、溫度場及風場均方根誤差的五天預報表現(圖4.11)，結果顯示enkfNMC50的高度場距平相關較另外兩組實驗低，但是溫度場及風場的均方根誤差並未顯示enkfNMC50有特別偏高的現象，因此原因可能還需要再進一步的分析及探討。

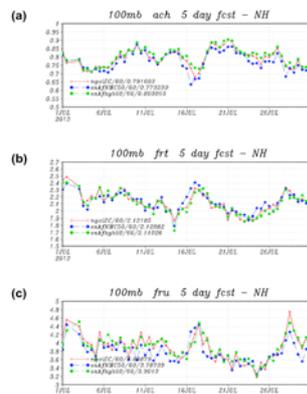


圖4.11 為三組實驗(ngsiEC、enkfNMC50和enkfhyb50)一個月期間(2013年07月)五天預報的500mb高度場距平、850mb溫度場的均方根誤差和300mb風場時間序列圖(由上而下)在北半球的結果。

整理夏季的結果，綜合來說依然是enkfNMC50的預報表現

是最佳的，在北半球區域的預報表現相較ngsiEC有顯著的提升，南半球區域雖然第五天的預報是enkfhyb50有較好的表現，但一至四日的預報仍然是enkfNMC50有較好的表現，兩組Hybrid實驗在夏季的表現都優於ngsiEC。

綜合冬夏兩個月的結果，其中發現兩組Hybrid實驗有一個共通的現象，就是溫度場的均方根誤差在兩個月的北半球高層(50mb以上)，相較ngsiEC都有明顯過高的偏差，因此我們針對enkfNMC50及ngsiEC兩組實驗，比較在30mb分層上的溫度場減去NCEP結果的水平分佈表現(圖4.11、圖4.12)。而冬夏兩個時間的溫度分析場，都是enkfNMC50的分析場較接近於NCEP的結果，但是溫度場的五天預報結果反而是enkfNMC50實驗偏離NCEP較多，尤其夏季的部分更為明顯，可以藉由這個現象，更進一步評估使用Hybrid資料同化系統對於模式在高層的影響，推測可能發生的原因以及改進方法，皆是接下來能夠努力的方向。

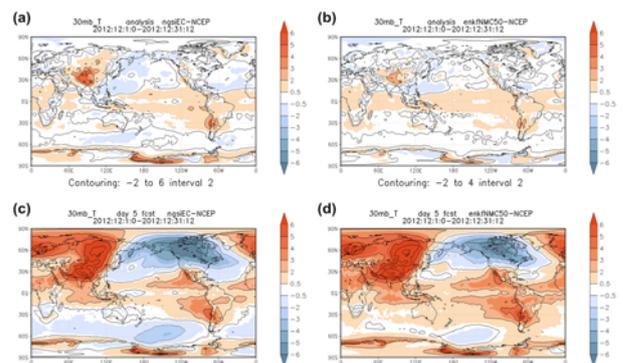


圖4.11 兩組實驗(ngsiEC和enkfNMC50)在30mb上，分析場及五天預報場分別減去NCEP的結果，時間為冬季(2012年12月)的平均。a、c為ngsiEC實驗，b、d為enkfNMC50實驗。

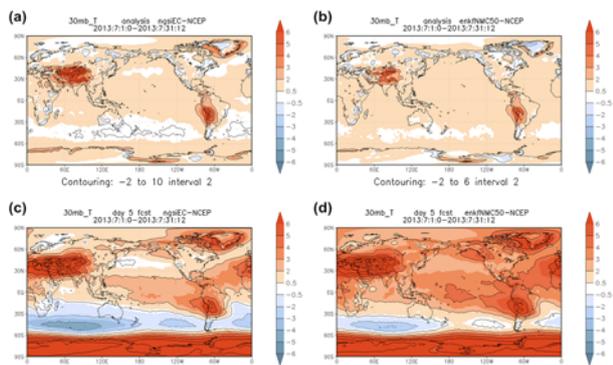


圖4.12 兩組實驗(ngsiEC和enkfNMC50)在30mb上，分析場及五天預報場分別減去NCEP的結果，時間為夏季(2013年07月)的平均。a、c為ngsiEC實驗，b、d為enkfNMC50實驗。

五、結論

綜合一系列的結果與分析，我們進一步確認Hybrid資料同化系統相較於現行3DVAR資料同化系統能提供的改善空間；整體看來Hybrid資料同化系統在冬半球區域的改善情形雖然較不明顯，但至少也能達到持平的表現，並不會劣於現行的資料同化系統；而在夏半球的提升就較為顯著，兩組Hybrid實驗都能給予正面效益。分析兩組Hybrid實驗之間的差異，平均都是以enkfNMC50的表現更為優異，這也顯示了添加NMC-method的系集擾動，能夠讓整體的預報趨向更好的結果。

Hybrid資料同化系統未來尚能調整的部分，還有很大的進步空間，除了各項參數的調整之外，也可以更深入的針對各系集組員間的品質及來源有更深人的探討，因此接下來的目標，則是分析並調整出更佳的參數設定，也希望能夠藉由參考這些調整資訊，對現階段的Hybrid資料同化系統，進行更深入的測試，並且改善需要擴充的部分，更加完善此資料同化系統。

六、參考文獻

楊明仁，2012：先進資料同化技術之建置與發展之完

整報告。交通部中央氣象局委託研究計畫，66頁。

Bishop, C. H., B. J. Etherton, S. J. Majumdar, 2001: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420-436.

Buehner, M, P. L. Houtekamer, C. Charette, H. L. Mitchell, B. He, 2010: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single Observation Experiments. *Mon. Wea. Rev.*, **138**, 1550-1566.

Charron, M., G. Pellerin, L. Spacek, P. L. Houtekamer, N. Gagnon, H. L. Mitchell, and L. Michelin, 2010: Towards random sampling of model error in the Canadian ensemble prediction system. *Mon. Wea. Rev.*, **138**, 1877-1901.

Etherton, B. J., and C. H. Bishop, 2004: Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error. *Mon. Wea. Rev.*, **132**, 1065-1080.

Hamill, T. M., C. Snyder, 2000: A Hybrid Ensemble Kalman Filter - 3D Variational Analysis Scheme. *Mon. Wea. Rev.*, **128**, 2905-2919.

——, J. S. Whitaker, 2011: What Constrains Spread Growth in Forecasts Initialized from Ensemble

Kalman Filters? *Mon. Wea. Rev.*, **139**, 117-131.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796-811.

Parrish, D. F. and J. Derber, 1992: The National Meteorological Center's spectral and statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Wang, X., C. Snyder, and T. M. Hamill, 2007a: On the theoretical equivalence of differently proposed ensemble/3D-Var hybrid analysis schemes. *Mon. Wea. Rev.*, **135**, 222-227.

——, D. M. Barker, C. Snyder, and T. M. Hamill, 2008a: A hybrid ETKF-3DVAR data assimilation scheme for the WRF Model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116-5131.

——, D. Barker, C. Snyder, T. M. Hamill, 2008b: A hybrid ETKF-3DVar data assimilation scheme for the WRF model. Part II: real observation experiments. *Mon. Wea. Rev.*, **136**, 5132-5147.

——, 2010: Incorporating ensemble covariance in the Gridpoint Statistical Interpolation (GSI) variational minimization: a mathematical framework. *Mon. Wea. Rev.*, **138**, 2990-2995.