



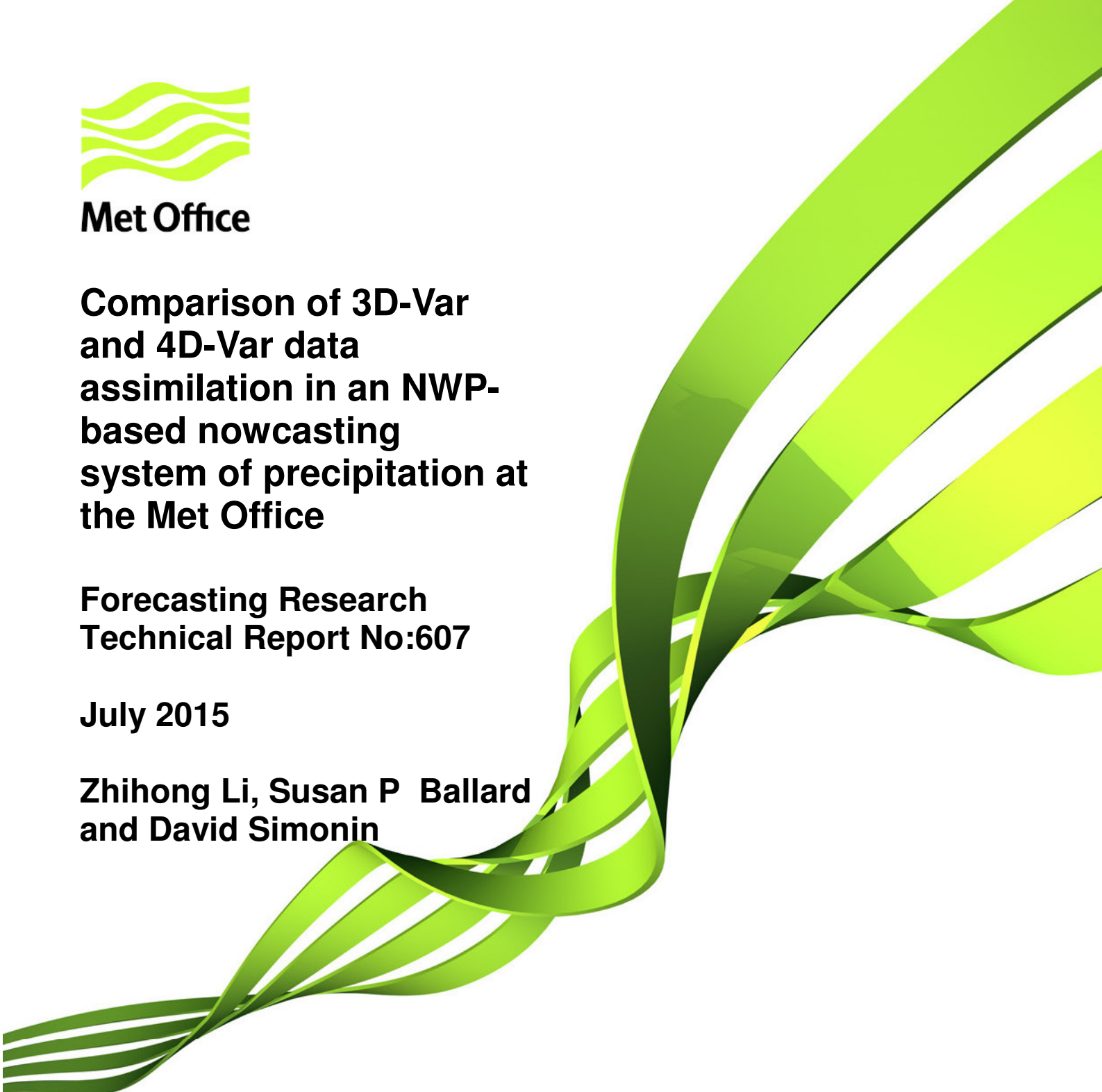
Met Office

**Comparison of 3D-Var
and 4D-Var data
assimilation in an NWP-
based nowcasting
system of precipitation at
the Met Office**

**Forecasting Research
Technical Report No:607**

July 2015

**Zhihong Li, Susan P Ballard
and David Simonin**



ABSTRACT

A high resolution Numerical Weather Prediction (NWP) nowcasting system has been developed and run in real time at the Met Office for the prediction of precipitation on a domain covering southern England and Wales. The hourly cycling nowcasting system, known as NDP (Nowcasting Demonstration Project), combined a 3km resolution 4D-Variational data assimilation (4D-Var) and a 1.5km resolution version of the Unified Model (UM) to provide hourly NWP analyses and forecasts for a period of 0 to 6 hours. Central to the NDP is the rapid updating cycles and its timely delivery of analyses and forecasts using the latest conventional and sub-hourly novel observations. In this paper the benefits of using 4D-Var assimilation compared to First Guess at Appropriate Time (FGAT) 3D-Variational assimilation (3D-Var) on forecasts of precipitation are considered by comparing model forecasts with radar-derived hourly surface accumulations for the period of June 2012, using an objective, scale-dependent verification scheme. It is shown that 4D-Var assimilation has a positive impact on precipitation forecasting skills compared to the corresponding 3D-Var assimilation for the whole nowcast period $[T+0, T+6]$. The 4D-Var assimilation system produced forecasts with greater spatial accuracy and longer lead times.

Contents

1. Introduction	3
2. NDP model configuration and data assimilation	6
2.1. Model	6
2.2. Incremental 3D-Var and 4D-Var	8
2.3. Hourly data assimilation cycle	10
3. Scale-dependent Verification Scheme of Precipitation	11
4. Comparison between NDP forecasts using FGAT 3D-Var and 4D-Var	14
4.1 A case of thunderstorms - 5th August 2012	15
4.2 Statistical Verification of FSS for the period of June 2012	18
4.2.1. Model domain averaged hourly rain accumulations	18
4.2.2. Fraction Skill Scores for one hour rainfall accumulations	20
5. Comparison between 3D-Var with sub-hourly data and 3D-Var with hourly data	24
6. A further case comparison – 28th June 2012	25
7. Concluding remarks	30
Acknowledgements	31
References	32

1. Introduction

An accurate and timely prediction of severe weather associated with convection is a subject of great importance. There is an increasing demand for improved very short range forecasts of hazardous high impact weather events, such as flash floods, from both the emergency planning authorities and the general public. More accurate local-scale rainfall forecasts delivered earlier would allow better and more informed risk management by a range of organisations such as the Flood Forecasting Centre, a partnership between the Environment Agency and the Met Office.

Very short range forecast of precipitation for periods of 0 to 6 hours ahead is commonly referred to as precipitation nowcasting. One early and widely-used nowcasting technique is the so-called observation/extrapolation-based method which involves the spatial and temporal extrapolation of observational fields such as precipitation and cloud from radar and satellite imagery in the first few hours of the forecast (see Browning and Collier, 1989; Witt and Johnson, 1993; Li *et al.*, 1995; Germann and Zawadzki, 2002). Considerable progress have been made in the extrapolation-based method for nowcasting of precipitation since the concept was first introduced by Ligda (1953). However extrapolation alone is unable to forecast any developing weather systems, which often cause hazardous weather including subsequent severe flood events.

From the 1980s onwards limited area Numerical Weather Prediction (NWP) systems were introduced into operations in National Meteorological Services and also run in government and academic research centres. With increasing availability of computer resources their resolution reduced from 10s of kilometres to kilometre scale (Saito *et al.*, 2006; Stephan *et al.*, 2008; Seity *et al.*, 2011). To extend the time period of the predictive capability of precipitation nowcasts, blended NWP/extrapolation nowcast techniques were subsequently developed in which the NWP model forecasts are merged with radar extrapolation (e.g. Golding, 1998, 2000; Mueller *et al.*, 2003; Bowler *et al.*, 2006, and reviews by Pierce *et al.*, 2012; Sun *et al.*, 2014).

The accuracy of these blended nowcasts will depend on the skill of the underlying NWP systems which will in turn depend on the model formulation, resolution and quality of the initial conditions. An important question with high resolution models is how to provide the

initial conditions. The simplest approach is to start each high resolution forecast from a lower resolution analysis without incorporating any data assimilation. This method has been attempted, among others, by Lean *et al.* (2008) who used a 12 km model to provide initial conditions for nested high resolution 1 km and 4 km models. They concluded that a high resolution NWP (1 km or 4 km) when starting from much lower resolution analysis (12 km in their case) was of little use for very short range rainfall forecasts to T+6, as it suffered from serious spin up problems while the high resolution structure spun up. An assimilation system to initiate high resolution models is required if very short range forecasts are of the prime concern.

The skill will also depend on the frequency at which the NWP forecasts are updated. Traditionally operational limited area short-range NWP systems produce 24 to 36 hour forecasts every 6 hours and often with 3-hourly data assimilation cycles. There is also a delay before they are run and the output produced, e.g. about 2-3 hours for the Met Office UK domain forecasts, as they need to wait for observations to arrive for the data assimilation (so called cut-off time) and for availability of runtime slots on the computer facility as well as taking up actual runtime for the observation processing, data assimilation and forecasts. If the extrapolation nowcasts are produced every 15 minutes or every hour within 15 minutes of observation time, the merged NWP forecasts will be 2/3 to 8/9 hours old in terms of data time. Ballard *et al.* (2015) studied the relative performance and skill of an hourly cycling 1.5 km resolution forecast model for southern England and Wales using 4D-Variational data assimilation (4D-Var) at 3 km resolution (Nowcasting Demonstration Project known as NDP) compared to the deterministic Met Office operational extrapolation/merged nowcast system STEPS (Short Term Ensemble Prediction System, Bowler *et al.*, 2006; Seed *et al.*, 2013) and a real-time, now operational, 3-hourly cycling 1.5 km resolution model for the UK that uses 3D-Variational data assimilation (3D-Var) and produces forecasts every 6 hours (UKV). The STEPS system used the UKV forecasts for its NWP merged component. They showed that, for the period June to August 2012, the hourly cycling NDP had better skill for precipitation forecasts out to 6 hours than the latest available UKV forecasts and better skill than the STEPS forecast from T+2 onwards. STEPS, if merged with the NDP, should increase the skill of its nowcasts. Alternatively with further development, the NDP could potentially also be used for nowcasting alone to replace STEPS in the whole UK domain.

The accuracy of very short range NWP depends critically upon the accuracy of the initial conditions and the ability of the model to retain information derived from frequent high

resolution observations. For a NWP nowcasting system to be successful, fast and efficient methods are required to frequently assimilate the latest high resolution observations. These are also required to provide initial conditions sufficiently accurate for the model to be able to predict convective weather development within the short nowcasting time frame. Fields have to be well balanced right from the start of the forecast to minimise effects of spin-up or spin-down of cloud and precipitation.

The NDP described in Ballard *et al.* (2015) can use either 3D or 4D-Var data assimilation in its hourly update cycle and includes sub-hourly radar, geo-stationary satellite and wind profiler data. The 4D-Var version of the NDP system was run in real time at the Met Office from March 2012, ahead of London 2012 Olympics, until April 2013 to showcase its capability. Ballard *et al.* (2015) reported on the performance of the 4D-Var system. Simonin *et al.* (2013) reported on experiments with a 3D-Var version of the system using just hourly observations to show the impact of radar Doppler radial winds.

Owing to the tight time constraints on forecasts, it is important to demonstrate that a nowcasting system incorporating a more expensive 4D-Var data assimilation produces improved skills when compared with a system using a simpler and cheaper assimilation method to produce the initial conditions.

In this paper, we concentrate on the impact of the model initial conditions on precipitation nowcasting arising from the use of 3D-Var compared to 4D-Var assimilation in the NDP. A key consideration in designing a NWP-based nowcasting system is its ability to deliver the short range forecasts promptly after assimilating frequent high resolution data into the model. In such a system, the data assimilation method may be the simpler, faster 3D-Var, or a computationally more intensive 4D-Var, depending on availability of computing resources and their relative performance. The performance of the NDP system in precipitation forecasts is considered on nowcasting timescales of 0-6 hours using a statistical analysis. The June 2012 period was rerun using 3D-Var with both the high time frequency observations used in the 4D-Var system described in Ballard *et al.* (2015) and also just hourly observations. Impacts of the 3D-Var and 4D-Var assimilation on model forecast skill are discussed using the scale-dependent scheme of Roberts and Lean (2008), applied to hourly accumulations of precipitation.

In Section 2, we describe the data assimilation system that is based on a combination of a 3D-Var or 4D-Var scheme and latent heat nudging (Ballard *et al.*, 2015; Dixon *et al.*,

2009; Jones and Macpherson, 1997) used for the assimilation of conventional and sub-hourly novel observations. The implementation of the rapid hourly cycling data assimilation system within a 1.5 km grid length version of the UM is presented. Section 3 summarises key aspects of a verification method where forecast skill is measured by a scale-dependent, statistical analysis procedure. Impacts of 3D-Var and 4D-Var assimilation on nowcast skills are presented in Section 4, first for an example of 28 August 2012 and then for the 30 day period of June 2012. In Section 5 3D-Var with hourly and subhourly data is compared and in Section 6 a comparison of forecasts from the NDP system for the case of 28 June 2012 is shown for 3D-Var with hourly and subhourly data and 4D-Var. Concluding remarks are presented in Section 7.

2. NDP model configuration and data assimilation

2.1. Model

The NDP model configuration is described in Ballard *et al.* (2015). The NDP used a 1.5 km fixed grid length version of the Unified Model, UM (Davies *et al.*, 2005), with a time step of 50 seconds. The UM is a non-hydrostatic, deep-atmosphere forecast model used for applications of atmospheric modelling at the Met Office, ranging from regional and global NWP to climate research. The UM is configured to run in each application using a common computer code and operating system.

Figure 1 shows the NDP model domain which covered the southern half of the United Kingdom. Also shown in Figure 1 is the larger domain of the operational UKV model which provides lateral boundary conditions (LBCs) for the NDP. The UKV model, which also has a horizontal resolution of 1.5 km, was run using a 3-hourly 3D-Var assimilation cycle and, at the time of the experiments in 2012, was producing a 36 hour forecast every 6 hours. The LBCs were time interpolated from UKV model forecast output provided at 30 minute intervals. The LBCs were updated to use the latest UKV run in the next NDP run after they would have become available in real-time.

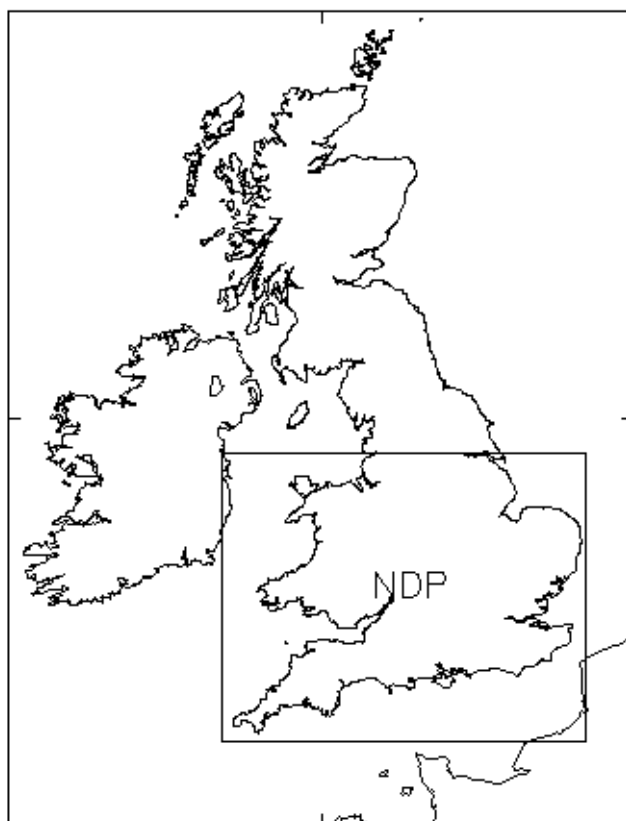


Figure 1. The NDP (inner rectangle) and UKV (outer rectangle) 1.5 km grid-length model domains

The NDP model domain consists of 366 x 288 horizontal grid points and 76 vertical levels. The size of the model domain was chosen during the NDP development to allow for a prompt delivery of hourly analyses and 6 hour forecast (which is within 15 mins from the start of the model run including the data processing, assimilation and forecast) when using an hourly 4D-Var assimilation cycle on the computer system available at the time

The NDP model configuration is essentially the same as the UKV. Detail is given in Ballard *et al.* (2015) and is not repeated here. Neither the NDP nor the UKV use a convection parameterization scheme on the assumption that the 1.5 km grid length is sufficiently small to allow for a reasonable representation of convective rainfall. The implication of using this assumption will be discussed later in the context of modelled rainfall accumulations.

2.2. Incremental 3D-Var and 4D-Var

The Met Office variational data assimilation system uses either an incremental three-dimensional scheme, 3D-Var (Lorenc *et al.*, 2000), or an incremental four-dimensional scheme, 4D-Var (Rawlins *et al.*, 2007). For the experiments described here both versions have been configured in limited area mode for the NDP system which has a nonlinear Unified Model resolution of 1.5 km whilst an analysis resolution of 3 km has been used for both 3D-Var and 4D-Var. In the basic runs, both use the same set of observations \mathbf{y}^o distributed within an assimilation time window $[T - m, T + m]$ where m is 30 mins and T is the nominal analysis time in hours UTC i.e. 0000, 0100, 0200, ..., 2300.

Given a full resolution, non-linear forecast model background state \mathbf{x}^b , the incremental variational assimilation seeks a simplified, (and usually lower resolution) incremental perturbation model state (or increments in short) $\delta\mathbf{w}^a$ to a full resolution guess field \mathbf{x}^g such that the analysis at full resolution \mathbf{x}^a is given by

$$\mathbf{x}^a = \mathbf{x}^g + S^{-1} \delta\mathbf{w}^a \quad (1)$$

where S^{-1} is the incrementing operator; it is the generalised non-linear inverse of a simplification operator, S which reduces the full model's complexity and resolution to that of the perturbation. In general, S^{-1} transforms from a lower resolution of $\delta\mathbf{w}^a$ to a full resolution of \mathbf{x}^a when S is just a decrease in resolution (Ide *et al.*, 1997). In the Met Office variational schemes where the full resolution non-linear model is the UM, S is also used to simplify multiple moisture and cloud variables to a single variable (Rawlins *et al.*, 2007).

The perturbation model state $\delta\mathbf{w}^a$ is found by minimising a penalty function J which in the basic form measures the misfit of the model state to the background state and observations:

$$J(\delta\mathbf{w}) = 1/2(\delta\mathbf{w} - \delta\mathbf{w}^b)^T \mathbf{B}^{-1} (\delta\mathbf{w} - \delta\mathbf{w}^b) + 1/2(\mathbf{y} - \mathbf{y}^o)^T (\mathbf{E} + \mathbf{F})^{-1} (\mathbf{y} - \mathbf{y}^o) \quad (2)$$

where $\delta\mathbf{w}^a$ minimizes J , $\delta\mathbf{w} = S(\mathbf{x}) - S(\mathbf{x}^g)$ and $\delta\mathbf{w}^b = S(\mathbf{x}^b) - S(\mathbf{x}^g)$, \mathbf{B} is the background error covariance matrix, \mathbf{E} and \mathbf{F} are instrumental and representativeness

error covariance matrixes, respectively, \mathbf{y}^o and \mathbf{y} are the observations and model simulations of the observations throughout the time window.

The notation of Rawlins *et al.* (2007) is also used in (2) to make the 4D-Var penalty function look more like 3D-Var. We use underlines to represent four-dimensional variables. Omission of the underline in variables denotes fields valid at the analysis time. A significant difference between the 4D-Var and 3D-Var scheme is that $\delta\mathbf{w}^a$, and hence \mathbf{x}^a , is found for $t = T - m$ for 4D-Var and $t=T$ for 3D-Var.

In the NDP implementation of 4D-Var only one outerloop is performed, i.e. $\mathbf{x}^b = \mathbf{x}^g$. The same \mathbf{B} is used for the NDP in both 3D-Var and 4D-Var and is described in Ballard *et al.* (2015). Also the observation errors, $\mathbf{R} = \mathbf{E} + \mathbf{F}$, are kept the same in both 3D-Var and 4D-Var versions.

In the 3D-Var formulation (as used in Simonin *et al.*, 2013 and the UKV), we use the First Guess at Appropriate Time (FGAT) method (Lorenc and Rawlins, 2006). In the simulation of the observations $\mathbf{y} = H(\mathbf{x}^g + S^{-1}\delta\mathbf{w})$, where H is the observation operator and \mathbf{x}^g is 4D variables and calculated using full 1.5 km resolution UM forecast output at 10 min intervals from the previous NDP cycle (background forecast) by linear interpolation in time to the actual time of the observations. However persistence is used in the process of minimizing J , where the incremental fields $\delta\mathbf{w}$ are held time-invariant over the assimilation window $(T-m, T+m)$ and calculated for time T . The addition of 4D and 3D variables \mathbf{x}^g , $\delta\mathbf{w}$ in the observation operator H is achieved by copying $\delta\mathbf{w}$ over the time-dimension.

Instead of the assumption of persistence, the 4D-Var formulation uses a linear Perturbation Forecast (PF) model with a lower (3km) resolution to generate the time-varying model trajectory over the assimilation window (Ballard *et al.*, 2015). At each iteration of the minimisation (which we call the inner loop), the PF model is first stepped forward in time from the initial time $T-m$ to provide increments $\delta\mathbf{w}$ at the actual time of each observation (to the nearest timestep of the PF model). The model prediction of the observations is given by $\mathbf{y} = H(\mathbf{x}^g + S^{-1}\delta\mathbf{w})$. An adjoint (backward) integration is then performed from the observation time using the adjoint of the PF model so that the penalty gradients at the initial time, $T-m$, needed by the actual minimization, are calculated. The PF model needs a linearization state (LS) from the previous nonlinear NDP forecast reduced to 3km from 1.5km resolution, $S(\mathbf{x}^g)$, and this is also provided by

output at 10 min intervals. The PF model has a 100 second timestep. The LS states are snapshots, not time interpolated, and the LS state nearest to the appropriate PF timestep is used in the PF and adjoint model time integrations (see Ballard *et al.*, 2015).

As in Lorenc *et al.* (2000), minimization of J is carried out with respect to pre-conditioned control variables after a transformation of model variables to independent, uncorrected modes of stream function, velocity potential, unbalanced pressure, relative humidity and log(aerosol concentration) for visibility analysis.

2.3. Hourly data assimilation cycle

The observations used in the hourly updating 4D-Var NDP system are shown in Table 2 of Ballard *et al.* (2015) compared to those used in the 3-hourly, 3D-Var operational UKV model. The latter uses mainly hourly observations but the NDP extended the observation usage to include also sub-hourly observations of wind profiler, radar radial Doppler winds and Meteosat Second Generation (MSG) SEVIRI radiances.

Figure 2 shows a schematic diagram of the hourly 3D-Var data assimilation (DA) cycle used in the NDP experiments reported in this paper. This should be compared to Figure 2 of Ballard *et al.* (2015) which shows the 4D-Var system. They both consist of a 3D- or 4D-Var analysis combined with latent heating nudging (LHN) for radar-derived surface precipitation rate. The UM model background/first guess state is obtained from the previous cycle's one hour UM forecast. In both systems, in each cycle, the one hour long assimilation window is centred on the nominal analysis time T ($= 0000, 0100, 0200 \dots 2300\text{UTC}$) and ranges from $T-30\text{mins}$ to $T+30\text{mins}$. At each cycle, the UM is run from $T-30\text{mins}$, using the model state valid at $T-30\text{mins}$ from the previous cycle as initial conditions, (which is a one hour forecast from the start of the previous assimilation window at $T-90\text{mins}$) to provide forecasts up to $T+6$ hours. The 1.5 km UM forecasts also provide, at 10 minute intervals from $T+30\text{mins}$ to $T+90\text{mins}$, the fields for the simulation of the observations and the linearization states for the next cycle (in 3D-Var, only one linearization state valid at $T+0$ is required) and the first guess/background field for the next cycle (at $T+30\text{mins}$ in 4D-Var and $T+60\text{mins}$ in 3D-Var). During the DA period, the UM forecast is corrected by the addition of increments produced by both the variational analysis and the latent heat nudging procedure. In the 4D-Var assimilation, the 4D-Var increments are added into the model (UM) at the start of the forecast at $T-30\text{mins}$. The 3D-Var increments, on the other hand, are added in gradually using an

Incremental Analysis Update (IAU) initialisation scheme (Bloom *et al.*, 1996) during the first hour of the forecast from T-30mins to T+30mins. The IAU adds in the 3D-var increments $1/N$ th at a time over the N model time steps in the first hour of the forecast from T-30mins. The use of the IAU in 3D-Var and not 4D-Var follows the implementation of these schemes in the operational systems at the Met Office.

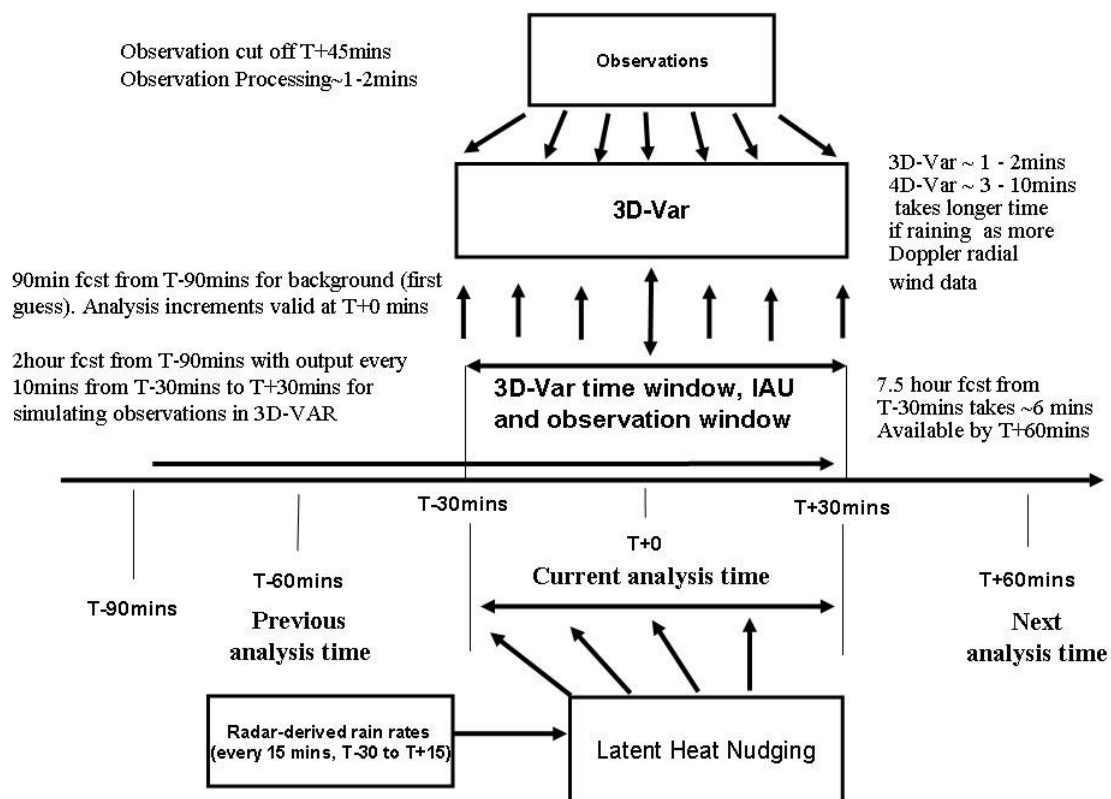


Figure 2. Schematic diagram of 3D-Var DA cycle together with LHN used in NPD 1.5 km hourly cycling nowcasting system.

In the experiments reported here, in each of the hourly time windows [T-30min, T+30mins], the same observations were assimilated in the 3D-Var and 4D-Var NDP. These are hourly 3D moisture derived from cloud observations (satellite + surface reports), surface temperature, relative humidity, wind, pressure and visibility as well as MSG cloud and humidity tracked winds (AMV) and aircraft temperature and winds (AMDAR) once per cycle, Doppler radial wind from 5 radars every 10mins and wind every 15mins from 4 wind profilers as well as satellite radiances from MSG SEVIRI channel 5 (clear and over low cloud) and channel 6 (clear) plus, over sea only, clear window channels (Table 2 in Ballard *et al.*, 2015). In both systems radar-derived surface rain rates, available every 15 mins from the UK radar network, were also assimilated in

the observation window, using Latent Heat Nudging (LHN) to make adjustments to the temperature and moisture fields aloft in response to the precipitation (Dixon *et al.*, 2009).

Both the 3D-Var and 4D-Var systems also used the same background error covariances derived from (T+6-T+3) forecast differences using lagged NMC method (see Ballard *et al.*, 2015).

In the NDP system, for each hourly cycle, 3D-Var took about 1-2 minutes to run whereas 4D-Var took about 3-10 minutes. The 3D-Var or 4D-Var runtime is very dependent on the number of Doppler radar radial winds assimilated which in turn depends on the areal coverage of rain within the model domain. At each hourly cycle, the total runtime in the real-time 4D-Var NDP including observation processing, 4D-Var and UM forecasts to T+6hrs was about 15 minutes so the availability of the T+6hrs forecasts would be 5 minutes earlier on average if the 3D-Var NDP system were used.

3. Scale-dependent Verification Scheme of Precipitation

In this study, the hourly surface precipitation accumulations from the NDP forecasts are verified against the quality controlled radar-derived surface rainrate composite used in the STEPS nowcasting system and also for latent heat nudging in the NDP system. The verification uses the Fraction Skill Scores (FSS) scheme of Roberts and Lean (2008). Hourly precipitation accumulations from T+0 to T+6 are first calculated from 1.5km model predictions at every time-step and also from the 5-minute, 1km observational surface rainrate composites from UK radar networks. Both model and radar accumulation fields are then interpolated onto the same 5 km grid-length mesh. The verification is subsequently carried out on this 5 km grid length mesh.

FSS is a method of objective verification which measures how the skill of a precipitation forecast varies with changing spatial horizontal scale and accumulation threshold. This is done by computing model forecast rainfall fractions of grid squares *exceeding* a particular accumulation threshold within different sized sampling squares (neighbourhood) in 5 km verification grid-length mesh. The fractions from radar observations using the same neighbourhood size and accumulation threshold are also computed. A skill score, FSS, is then calculated in terms of normalised mean square

error (MSE) between the model forecasts and the observations in the fractions (probabilities).

For a given accumulation threshold, FSS at a sampling square of width l is given by

$$FSS^{(l)} = 1 - \frac{MSE^{(l)}}{MSE_{worst}} = 1 - \frac{\frac{1}{N} \sum_{j=1}^N (O_j^{(l)} - M_j^{(l)})^2}{\frac{1}{N} \sum_{j=1}^N [(O_j^{(l)})^2 + (M_j^{(l)})^2]} \quad (3)$$

where N is the total number of 5km grid points in the verification area, and $O_j^{(l)}$, $M_j^{(l)}$ are observed and model forecast fractions at each 5 km verification grid point j exceeding the given accumulation threshold within the sampling square of width l . MSE_{worst} is the largest possible MSE which represents the worst forecast when no overlap occurs between the observed and model forecast fractions.

For a given accumulation threshold, we therefore compute FSS for different spatial scales by changing the neighbourhood size at each 5km grid point, j . Starting from the smallest 5km verification grid-length mesh, the neighbourhood size increases progressively with the width of the neighbourhood size being $l = 5(2i - 1), i = 1, 2, 3, \dots$ or $l = 5km, 15km, 25km, \dots$ and the sampling squares corresponding to $l \times l$.

In a model forecast, FSS has a range of values between 0 (zero skill) to 1 (perfect skill). Typically FSS increases with increasing spatial scale l as progressively large spatial errors in the model prediction are allowed for.

As the spatial scale l increases, the sharpness of the observed and model forecast fractions and hence the useful information content is gradually diminishing. Following Roberts and Lean (2008), we can choose a target level of skill for the model forecast to reach:

$$FSS_{target} = 0.5 + 0.5 f_o \quad (4)$$

where f_o is the fraction of observed rainfall pixel points exceeding the given threshold. FSS_{target} is exactly halfway between a random forecast and the perfect forecast and would be the skill by assigning the model forecast fraction at every 5 km grid point equal to f_o over the whole verification domain. It may be regarded as a minimum skill we choose to achieve for a forecast to become skilful (see Roberts and Lean, 2008).

In what follows, we define the relative skill between FSS and FSS_{target} as

$$\Delta FSS = FSS - FSS_{target} \quad (5)$$

with the minimum spatial scale L_{min} corresponding to the scale length when FSS_{target} is first reached by FSS (i.e. $\Delta FSS = 0$). L_{min} is regarded as a measure of the overall model forecast scale over which the model forecast may contain sufficiently useful information (Roberts and Lean, 2008).

The accumulation threshold can be expressed either as an absolute precipitation (e.g. 1 mm in one hour period) or as a percentile of grid points. If the model has a substantial bias in the rain intensities, the FSS based on the absolute accumulation threshold will be dominated by this bias, and the model skill on spatial accuracy (i.e. location of precipitations) will be masked in FSS. We can use the percentile (relative) thresholds to remove the model bias in rain intensities. This is achieved by comparing the top fractions of the rainfall so that spatial quality of the forecasts can be assessed. The 90th percentile threshold for example selects the 10% grid points with the highest accumulations in the both the model and radar fields respectively.

4. Comparison between NDP forecasts using FGAT 3D-Var and 4D-Var

In this section, we present forecasts from the hourly cycling FGAT 3D-Var assimilation run and the 4D-Var run to compare their relative performance. 4D-Var is the version used in the real-time NDP.

In order to compare 3D-Var results as closely as possible to the 4D-Var, we used the exact same observations in the 3D-Var as in the 4D-Var. The 3km resolution was also

used in obtaining the 3D-Var analysis increments – the same as in the 4D-Var PF and its adjoint model.

As mentioned in Section 2, 3D-Var and 4D-Var NDP differ in the initialization of the forecast model. At each hourly cycle, the IAU initialization method is used in 3D-Var version where increments calculated for $T+0$ are added from $T-30\text{min}$ to $T+30\text{min}$, while analysis increments in 4D-Var version are added instantaneously into UM at the (correct) initial time for that system at $T-30\text{min}$.

4.1 A case of thunderstorms - 5th August 2012

Before presenting the statistical analysis for the NDP rainfall forecasts, we describe a case to illustrate how the subjective assessment of precipitation forecast could be related to the objective method such as the Fraction Skill Scores (FSS) verification scheme.

There was extensive convective cloud and storms over England and Wales on the 5th August 2012. Flash flooding was reported in Pembrokeshire and in North Somerset. Lightning and heavy rain also disrupted the Olympic events in and around London, although the sailing in Weymouth was unaffected as the cloud and rain was confined to land areas and Weymouth Bay basked in sunshine apart from a non-precipitating cloud street triggered from Portland Bill.

Figure 3 shows, in the left hand column, the precipitation accumulation for the hour 1000-1100 UTC from 3D-Var and 4D-Var NDP 1000 UTC forecast cycle. Radar-derived accumulation is also shown for comparison. Both 3D-Var and 4D-Var NDP had an overall good indication of the locations of heavy rain to the north of London, in Pembrokeshire and Bristol, though rain intensity from the 3D-Var NDP forecast was much heavier. Subjectively, the 4D-Var NDP seems to have a better forecast than the 3D-Var as the 4D-Var NDP appeared more accurate spatially in the location of the main rain areas compared with the 3D-Var NDP.

Plots of the corresponding fraction fields $O_j^{(l)}$, $M_j^{(l)}$ from the radar hourly accumulations and model forecasts are also shown in Figure 3, right hand column, using a neighbourhood length scale of ($l=$) 35km and the 90th percentile threshold. Consistently with the accumulation fields themselves, the better 4D-Var forecast is

confirmed by the fraction fields and the resulting FSS, being 0.782 as compared with 0.656 in the 3D-Var forecast. The corresponding minimum length scale, another measure of the model forecast skill, is 10.6 km and 20.9 km respectively.

Table 1 lists the skill scores for the rainfall accumulations over the period [1000-1100UTC] using absolute thresholds of 0.2, 0.5, 1.0, 2.0 mm, which again shows the scores from the 4D-Var run are better than those from the 3D-Var run.

Table 1 Fraction Skill Score for one hour rainfall accumulations from forecasts from the 3D-Var and 4D-var systems for 1000 to 1100 UTC from data time 1000UTC 5 August 2012 cycle for different accumulation thresholds and 35km scale as well as the minimum lengthscales for acceptable skill.

Threshold	Var	FSS, $l = 35km$	$l_{min}(km)$
90th	3D/4D	0.656/0.782	20.9/10.6
0.2 mm/h	3D/4D	0.689/0.699	20.5/17.5
0.5 mm/h	3D/4D	0.658/0.662	23.4/20.3
1.0 mm/h	3D/4D	0.618/0.619	25.5/23.5
2.0 mm/h	3D/4D	0.557/0.572	30.8/24.8

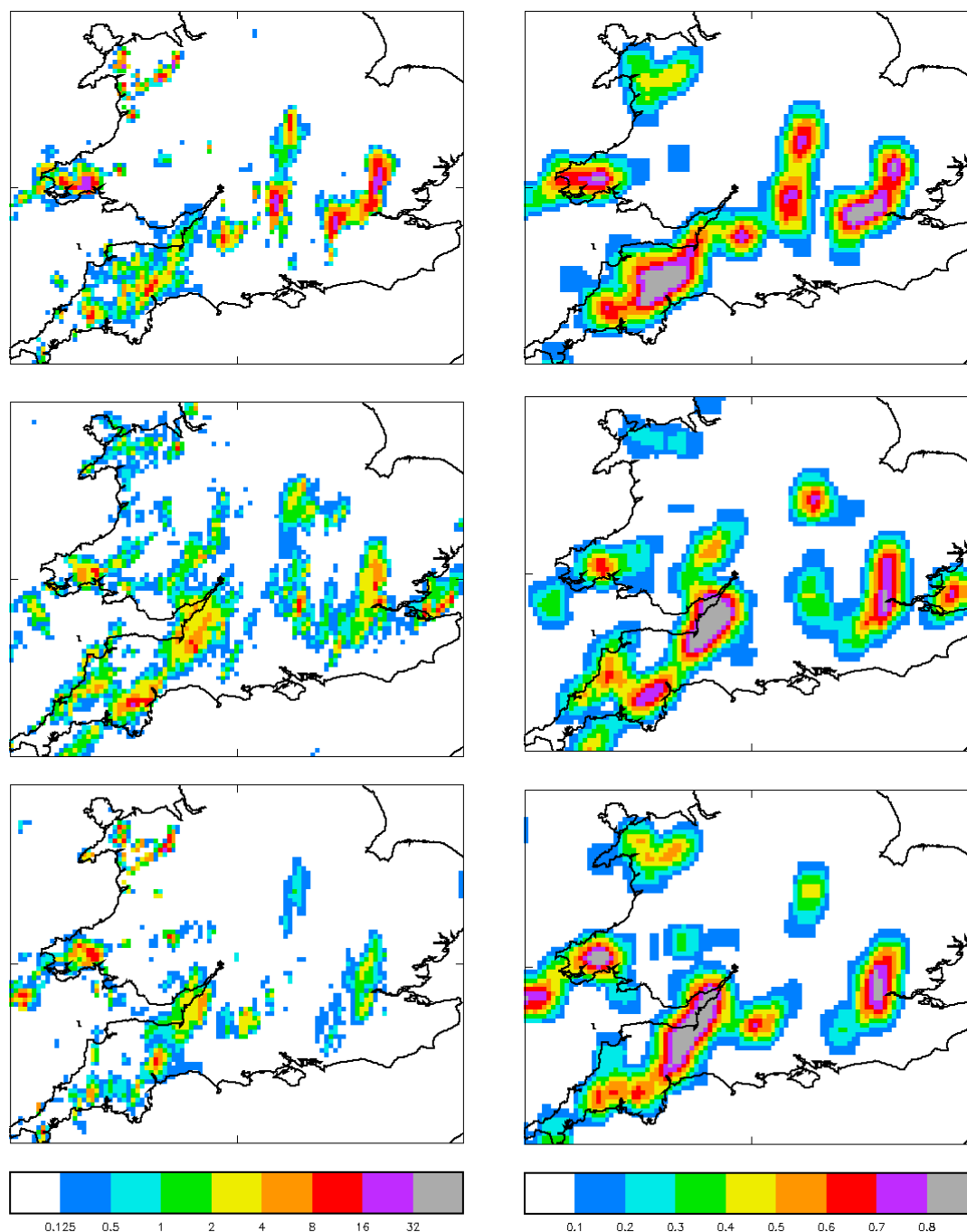


Figure 3. NDP forecasts for the hour 1000-1100UTC 5th August 2012 from the 1000UTC cycle. In the left column are hourly precipitation accumulations with the radar derived surface rain accumulation (middle) , the 3D-Var run (top) and the 4D-Var run (bottom) . In the right column are the fraction fields using ($l =$) 35km and the 90th percentile threshold from radar derived rain accumulation (middle) , the 3D-Var run (top) and the 4D-Var run (bottom) .

4.2 Statistical Verification of FSS for the period of June 2012

Performance of the NDP nowcasting system for very short range forecasts is assessed statistically here for the period of June 2012 using the two DA configurations, namely 3D-Var and 4D-Var. June 2012 was chosen as it was the UK's wettest June since 1910 and the weather was dominated persistently by low pressures with complete absence of settled spells (it in fact rained every day in the NDP model domain for the whole month). There were numerous widespread flooding events during the month, with most serious events occurring on the 8th in Aberystwyth area of Wales, and the 28th in Midlands as a result of supercell thunderstorms and torrential rain.

4.2.1. Model domain averaged hourly rain accumulations

When a high resolution model is initialised from a lower resolution analysis, a finite time is required to spin up high resolution structures during the subsequent forecast. Obviously, the length of the spin up period varies depending on the model resolution and also the lower resolution driving model that also provides the initial conditions. For example, when a 4 km resolution model is initialised and driven by a 12 km model that uses a parameterized convection scheme, it can take up to 6 hours for explicit convection to spin up (Lean *et al.*, 2008). The 4 km model produces low rain rates in the first few hours of the forecast but this is followed by a large overshoot (peaking at about T+6) as early unrealistic build up of CAPE is released later on in the forecast.

When models are run continuously as in the experiments reported here with the hourly cycling NDP system and in the operational 3-hourly cycling UK domain forecasts and 6-hourly cycling global forecasts that type of spin up should not occur as the background previous forecast will already be consistent with the model resolution and formulation. However spin up or down can occur due to imbalanced initial fields as a result of the insertion of the analysis increments.

The time-dependent model domain averaged hourly rainfall accumulations for the month of June 2012 from the 3D-Var, 4D-Var NDP runs and radar are plotted in Figure 4. Radar data are accumulations on a 5 km grid aggregated from 5 minute, 1km quality controlled, derived surface rainrate composites. Model rain accumulations are those

from the 1.5km model fields, interpolated to the 5 km verification grid before averaged over the area of the 1.5 km NDP model domain.

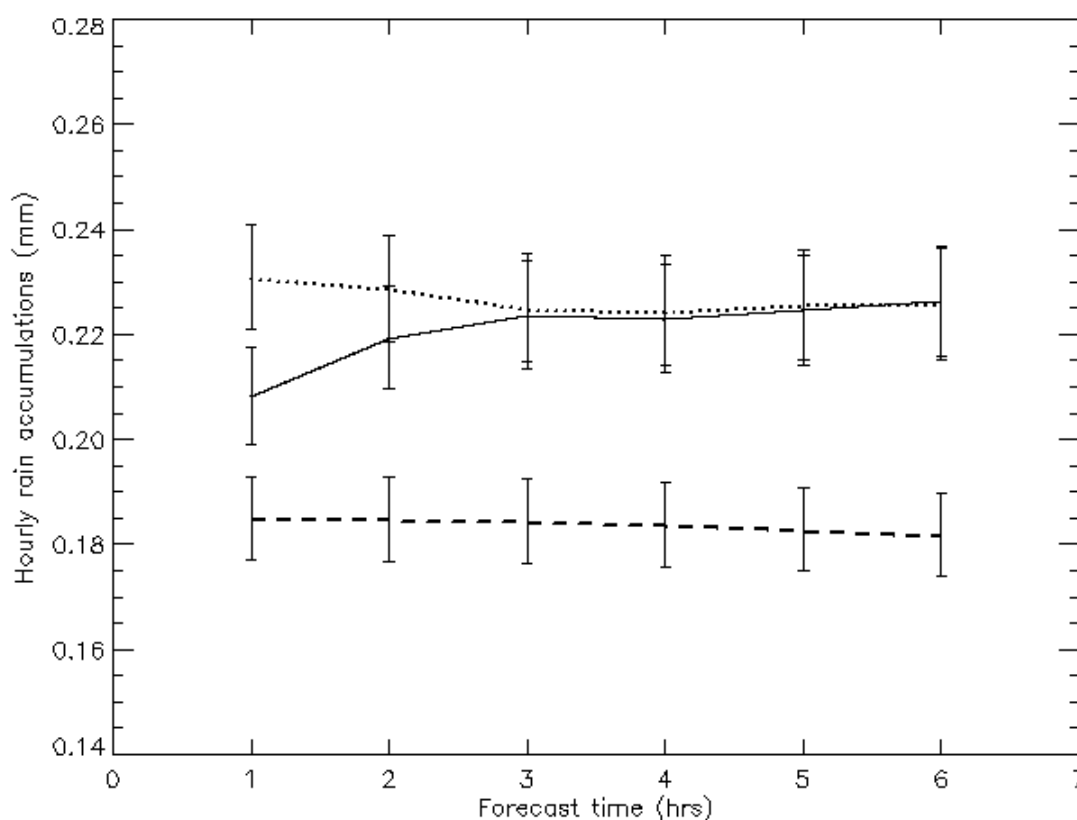


Figure 4. Mean model and radar hourly precipitation accumulations over the NDP model area. Black line is the 4D-Var NDP, dotted is the 3D-Var NDP and dashed line is the radar. Uncertainties of the means are given by error bars of standard error of the mean. 1 in the horizontal axis refers to accumulations [T+0, T+1] and 6 refers to accumulations [T+5, T+6]

It can be seen from Figure 4 that the domain averaged rain accumulations from 3D-Var and 4D-Var NDP are about 20% higher than the radar-derived observations. This seems to be a characteristic of high resolution NWP models like the Unified Model at 1.5km resolution where the convection is treated explicitly but is only in fact partially resolved. The radar returns can also suffer from attenuation. Both the 3D-Var and 4D-Var NDP systems over-predict the accumulations.. However, compared with the 3D-Var NDP, use of 4D-Var in NDP leads to a reduction in rainfall accumulations, with the largest improvement occurring at the first hour of the forecast [T+0, T+1] but it is almost identical to 3D-Var at the end of the forecasting hour [T+5, T+6]. There is a slight reduction in the accumulation with increasing forecast time in the forecasts from 3D-Var but no major

spin down of precipitation indicating that the initial conditions due to continuous DA cycling and without using a convection scheme in both the NDP and driving UKV model LBCs are well balanced.

The 4D-Var initial conditions seem to have led to the forecast precipitation closer to the observed values as was seen in Figure 3. The closer agreement in the rainfall accumulations between 3D-Var and 4D-Var NDP in the latter part of the forecast period may be related to the limited size of the NDP domain and the increasing dominance of the LBCs from the driving UKV model or a relaxation of the 4D-Var forecast fields back towards the characteristics of the Unified Model.

The model bias towards too high intensities by both the 3D-Var and 4D-Var NDP is indicative of an inherent predisposition in high resolution convection-permitting models. Such models, like the present 1.5 km NDP model, run resolving convection explicitly (i.e. without a convection scheme) but still don't have high enough resolution. Likewise, the operational 1.5km UKV also tends to have very high peak rain rates in isolated convection cells compared to the radar where more light precipitation is observed over a slightly larger area.

4.2.2. Fraction Skill Scores for one hour rainfall accumulations

The model domain averaged rainfall accumulations, like the ones shown in Figure 4, are useful in assessing the bias characteristics of the 1.5 km model in forecasting convection events. The total amount of rainfall in the model domain however does not address a key and more important question about the spatial accuracy and scales over which the model can produce reliable information. For this purpose, we here use scale-dependent Fraction Skill Scores (FSS) verification method. The procedure described earlier has been applied to hourly precipitation accumulations for the entire month of June 2012. Comparison is then made between the hourly cycling 3D-Var and 4D-Var runs to demonstrate their relative skills, using both absolute and percentile (relative) rainfall thresholds.

The results presented below are for the month June 2012 where a total of 720 (30 days x 24 cycles) samples, applied to hourly accumulations, are used to generate the mean value of the skill \overline{FSS} at each forecast time and varying neighbourhood length for

different rainfall thresholds. Comparisons are then made between the 3D-Var and 4D-Var NDP to demonstrate their relative forecasting skills.

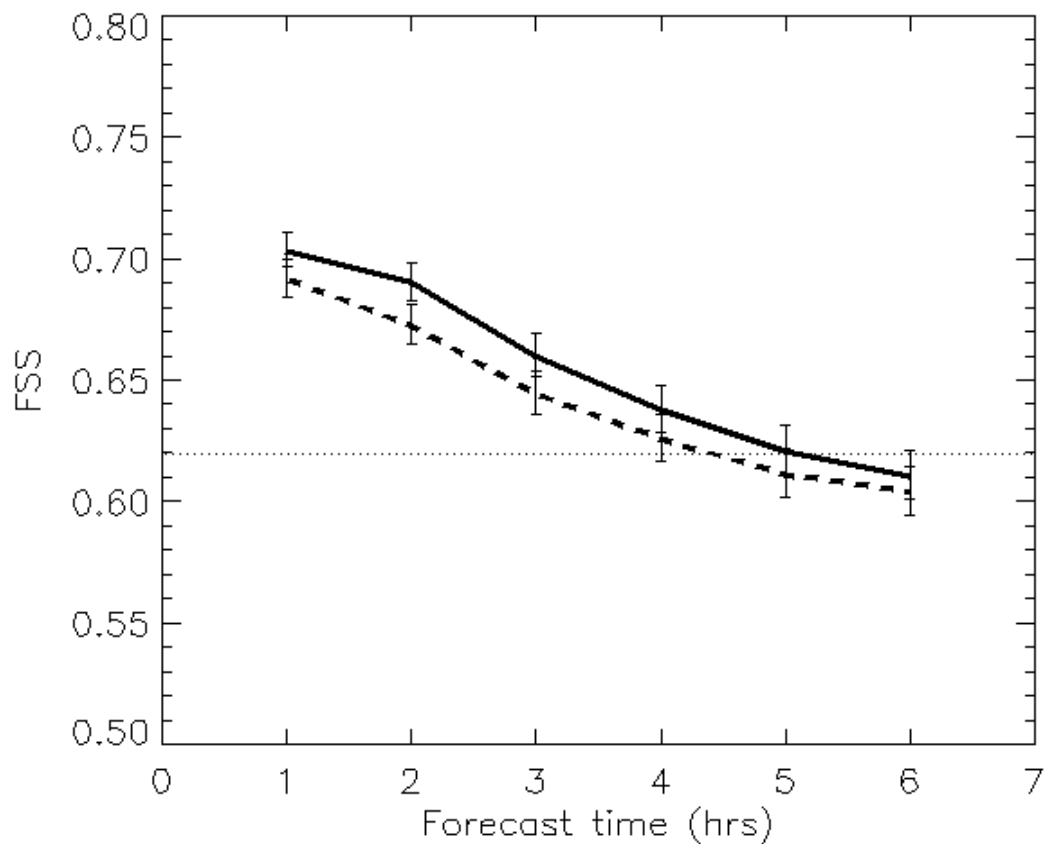


Figure 5 Fraction skill score for 0.2mm/h threshold and 25km scale for forecast time ranges T+1 to T+6 from NDP 3D-Var (dashed) and 4D-Var (black) forecasts in June 2012 where the threshold was achieved for at least 10% of the model domain. Uncertainties of the mean FSS are given by error bars of standard error of the mean. FSS_{target} is represented by the horizontal dotted line.

Forecast time series of \overline{FSS} using the 0.2mm/hr threshold with $l=25\text{km}$ is plotted in Figure 5 for both 3D-Var and 4D-Var forecasts in June 2012. A number of features are apparent here. First, the skill from both runs decreases with forecast time as a result of model error growth. Second, 4D-Var forecasts has a better skill in predicting spatial rainfall locations compared with the 3D-Var forecast. $\overline{L_{min}}$ is evidently smaller than 25 km in the 3D-Var and 4D-Var for the first 4 and 5 hours forecast, respectively wherein the \overline{FSS} is above with the target line $FSS_{target} = 0.5 + 0.5 f_o$. The runs using 4D-Var have almost an hour's gain in skill over the 3D-Var runs. The threshold of 0.2mm/hr is chosen

as it effectively verifies locations of precipitation and is not, or is less, affected by any biases in precipitation forecasts from the model.

Figure 6 shows contours of the mean value of skill, \overline{FSS} as a function of forecast time and spatial neighbourhood length scale using absolute thresholds of 0.2 and 1.0 mm/h. Inspection of Figure 6 reveals that the 4D-Var run is more skilful spatially than the 3D-Var run throughout the forecast period at each of the absolute thresholds. 4D-Var run has smaller neighbourhood sizes for the same forecast skill. A reduction of 5-10 km in the scale, for example, is evident (i.e. by drawing a line vertically) for the same fraction score of 0.7 for 0.2mm/h threshold in the 4D-Var compared to the 3D-Var forecasts for the whole forecasting period. \overline{FSS} in both model runs displays a decrease in skill as forecasts progress. The loss of accuracy in the model with time is due to error growth.

It is also clear in Figure 6 that for a given neighbourhood length scale the 4D-Var retains skill for longer than the 3D-Var forecasts. During the first few hours, there is a gain of an hour in forecast lead time in the 4D-Var run when compared with the 3D-Var run (e.g. comparing contours of 0.7 for 1mm/h by drawing a line horizontally).

\overline{FSS} is, as expected, higher for the lower accumulation threshold 0.2mm/h than the higher threshold 1mm/h (e.g. by comparing curves in Figure 6a and Figure 6b). This occurs partly because it is more difficult to forecast the higher accumulations, which is usually associated with more severe, localised events, and partly because of the model biases in rainfall intensity, which become more apparent for the higher accumulation thresholds.

As discussed earlier, the model forecast skill can be assessed quantitatively in terms of mean minimum spatial length scale, $\overline{L_{min}}$. For a given model forecast, the smaller $\overline{L_{min}}$ is, the better the prediction will be.

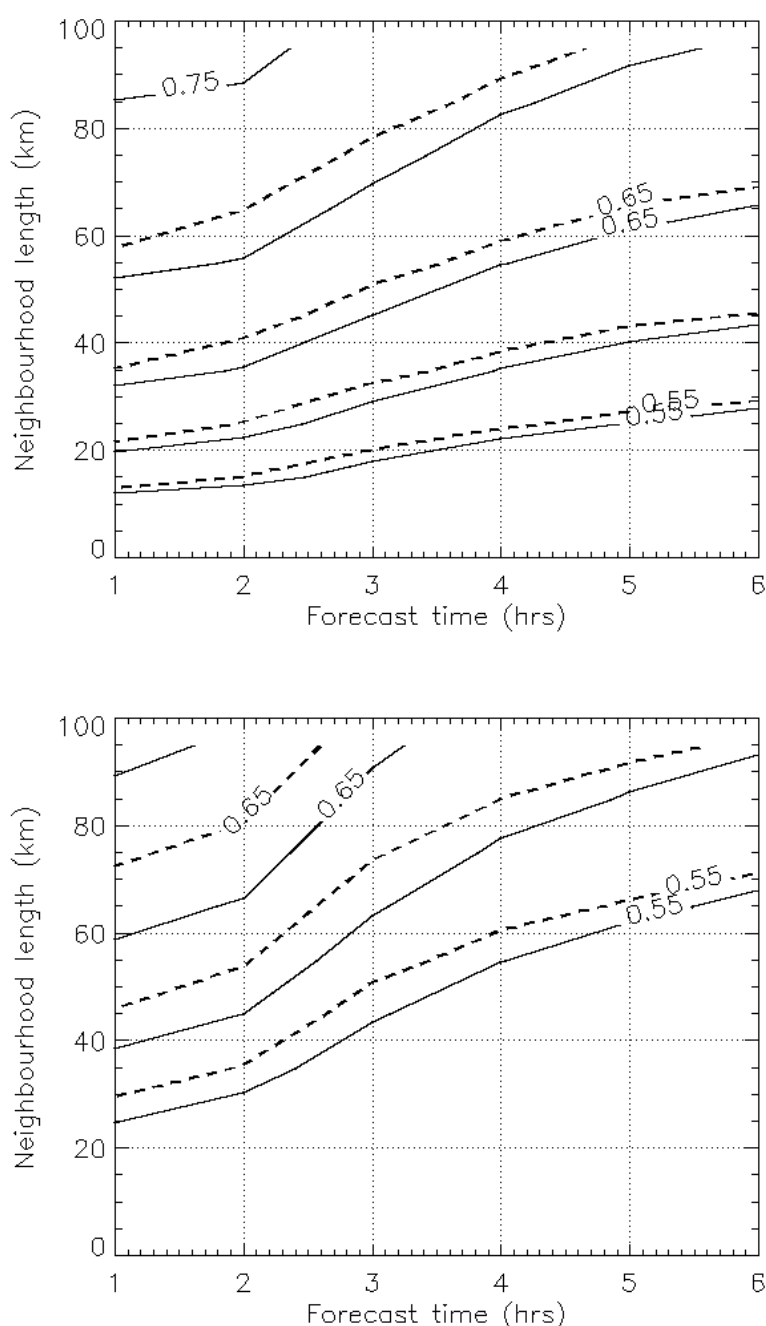


Figure 6. Contours of mean \overline{FSS} (in intervals of 0.05) from the 3D-Var (dashed) and 4D-Var (black) forecasts using rainfall threshold of (a) 0.2mm/h (top) and (b) 1.0 mm/h (bottom)

The dependence of $\overline{L_{min}}$ on forecast time from both the 3D-Var and 4D-Var NDP is shown in Figure 7 for the 90th percentile threshold. Figure 7 shows that $\overline{L_{min}}$ increases with forecasting time range, reflecting the gradual decrease in the model skills. But it

remains smaller than 40 km at the end of the forecast period. Acceptable skill is reached at scales of ~15 km at [T+0, T+1], and dropping to ~37 km at [T+5, T+6] in the 4D-Var runs. By comparison, the 3D-Var runs has a minimum scales of ~18 km and ~39 km at [T+0, T+1] and [T+5, T+6], respectively.

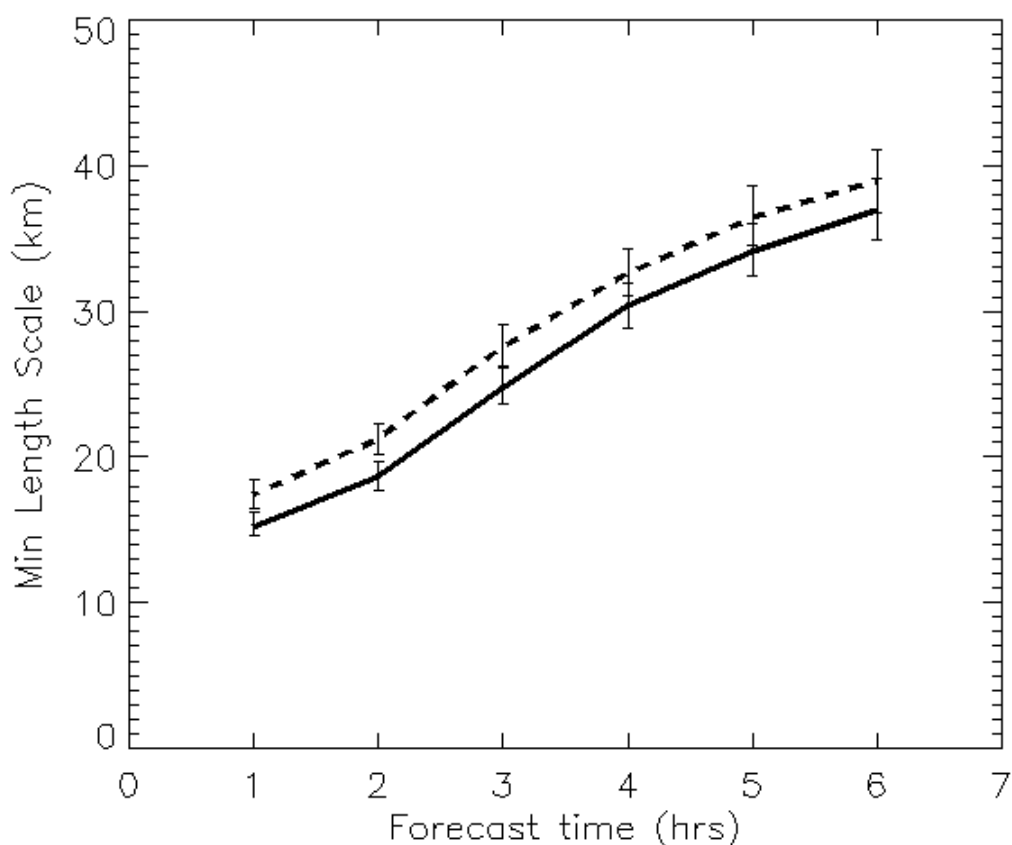


Figure 7. Mean length scale at which the level of desired skill is first reached, $\overline{L_{min}}$, as a function of forecast time from 3D-Var (dashed) and 4D-Var (black) forecasts for the 90th percentile threshold. Uncertainties of the mean L_{min} are given by error bars of standard error of the mean.

5. Comparison between 3D-Var with sub-hourly data and 3D-Var with hourly data

Results from the FGAT 3D-Var run shown so far are obtained using observations including sub-hourly data. It is not immediately obvious that, at convective scale resolution and with hourly assimilation, the FGAT will still be beneficial compared to the

traditional 3D-Var where only observations closest to the nominal analysis time are assimilated. This is because the assumption that the increments at the observation time are valid at the analysis time may be less beneficial when we are trying to capture fast moving, or changing, fine scale details of the weather. When the observation locations are not changing, as with wind profilers, we will be effectively assimilating an increment averaged over the analysis window but with higher weight.

We investigated this by re-running 3D-Var experiment for the month June 2012 using only the observations closest to the nominal analysis time (referred to here as Hourly-Data 3D-Var). Figure 8 shows the FSS for 0.2 mm/h and 1.0 mm/h from this experiment in comparison with the FGAT 3D-Var run. We upgraded the UM version from 7.9 to 8.2 in both runs as a result of the upgrade in our super computer software.

For the small rainfall threshold 0.2 mm/h, Figure 8 shows the skill in the Hourly-Data 3D-Var is almost the same as the FGAT 3D-Var. The FGAT 3D-Var however has a greater skill than the Hourly-Data 3D-Var for the larger 1 mm/h threshold. There is a reduction of 5km in the spatial scale in the FGAT 3D-Var forecast when compared with the Hourly-Data 3D-Var. Additionally, a longer lead time of 1 hour is apparent for the FGAT 3D-Var than the Hourly-Data 3D-Var in the first few hour forecast.

Therefore the FGAT is seen to still be beneficial in the hourly cycling convective scale system. The forecast has clearly benefitted from the use of extra, higher time frequency observations. However we have not investigated whether a similar result would be obtained by giving more weight to the observations in the Hourly-Data 3D-Var runs by reducing the observation errors.

6. A further case comparison – 28th June 2012

To illustrate the greater skill in the 4D-Var forecasts than those using 3D-Var initial conditions we look at an example forecast for the 28th June 2012 case shown in Ballard *et al.* (2015). It can be seen from Figure 9 that the forecast from the 4D-Var NDP system has a better prediction of the southerly extent of the storms than that from the 3D-Var system forecasts. Also the location of the precipitation in the north of the domain is better in the forecast from the 4D-Var than that from the 3D-Var system. The figure shows the 3D-Var subhourly data forecasts with UM v7.9 and v8.2 as well as the Hourly-Data 3D-Var and they are all more similar to each other than to the 4D-Var forecast.

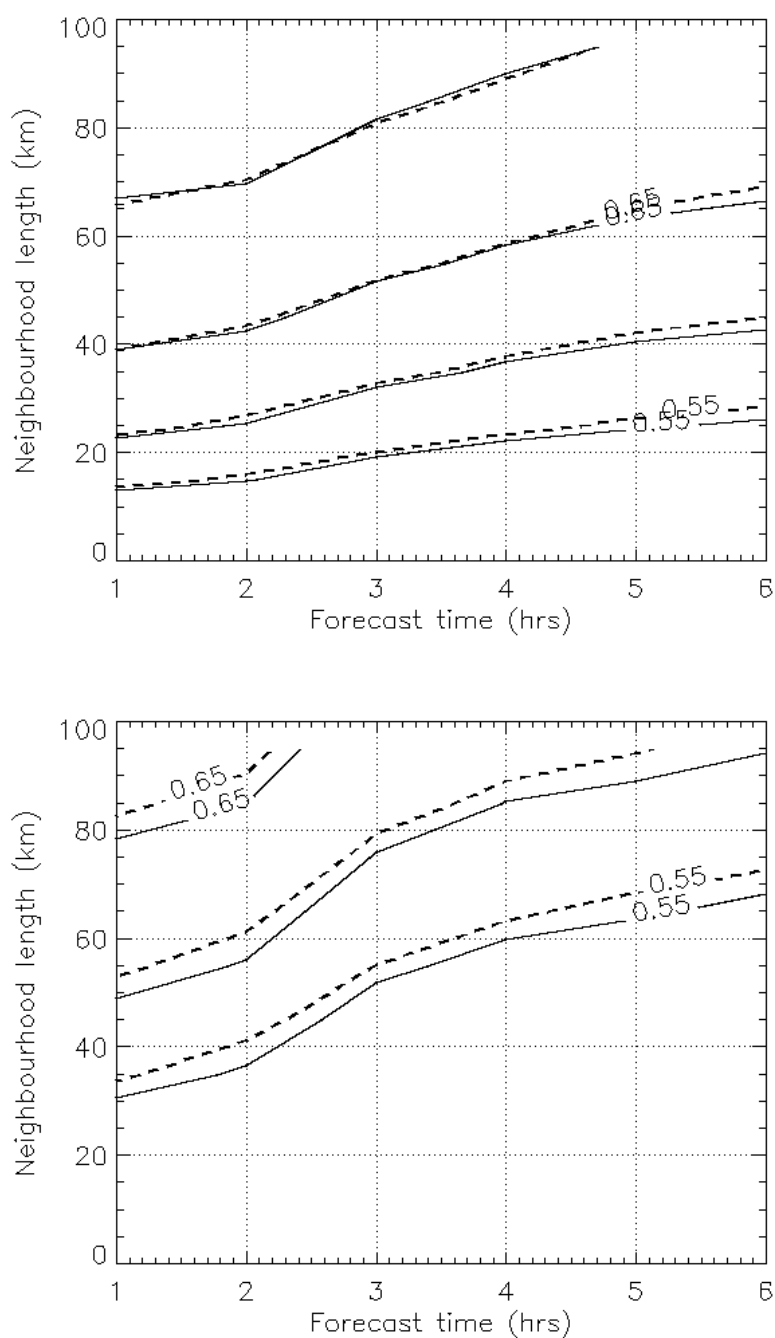


Figure 8. Contours of mean \overline{FSS} (in intervals of 0.05) from the FGAT 3D-Var (black) and Hourly-Data 3D-Var (dashed) forecasts using rainfall threshold of (a) 0.2mm/h (top) and (b) 1.0 mm/h (bottom)

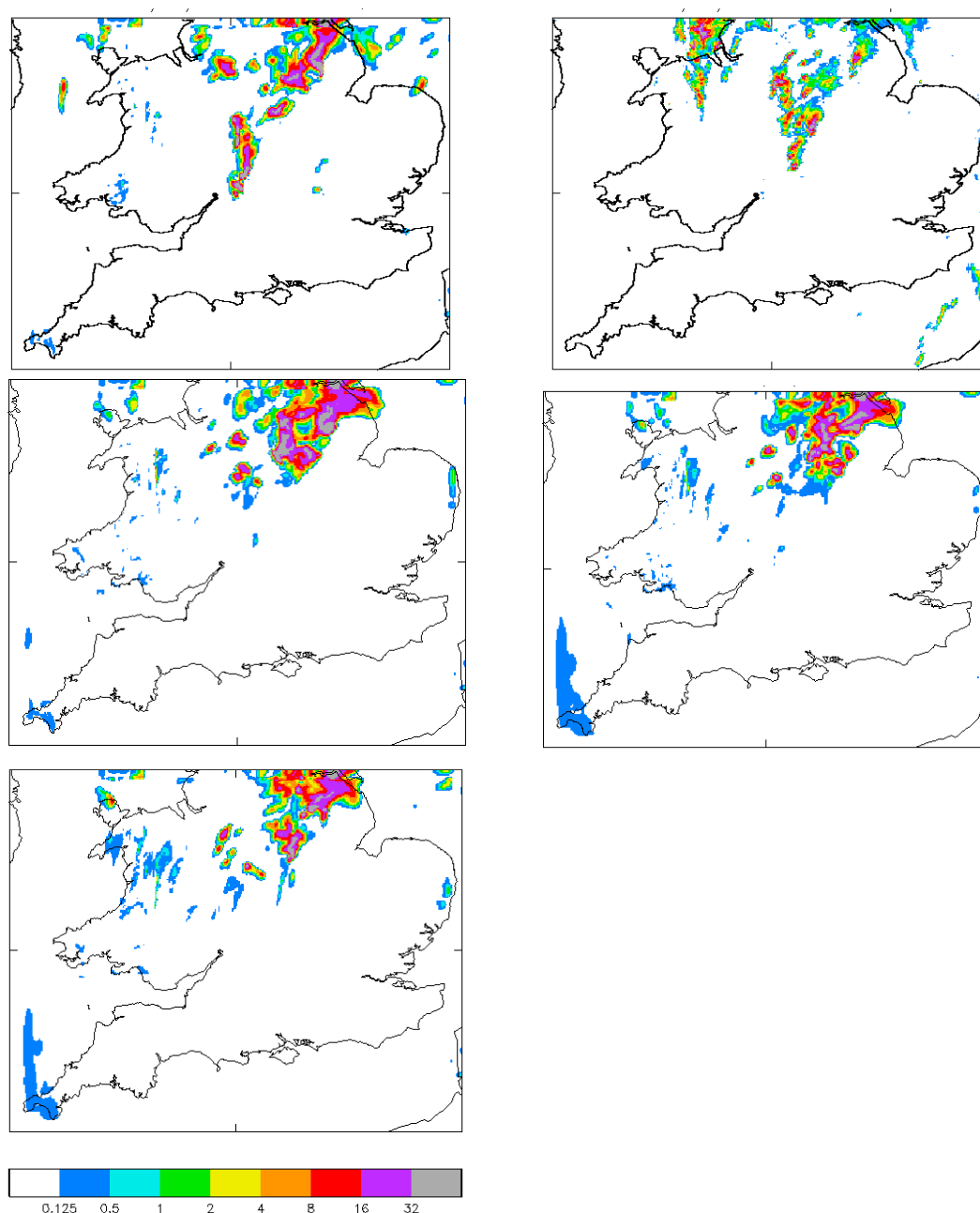


Figure 9 Comparison of forecast precipitation rate at 12UTC 28 June 2012 from 10UTC 4D-Var cycle (top left) compared to radar derived rain rates (top right), 3D-Var cycle (middle left) v 7.9, 3D-Var cycle v 8.2 (middle right) and Hourly-Data 3D-Var (bottom left)

On the 28th June 2012 there were 24 forecasts, ie one from each hourly cycle. Table 2 lists the differences in skill between the 3D-Var and 4D-Var forecasts for the threshold of 0.2 mm/h and 35km scale for each hourly period of rainfall accumulation ending at T+1

to T+6. Out of 24 cycles, 22 cycles were used in the comparison in Table 2 as another 2 cycles had very light or no rain at all when the 0.2 mm/h threshold was not reached.

Table 2 Relative skill for one hour rainfall accumulations from the 3D-Var and 4D-Var forecasts for the 28th June 2012 using Fraction Skill Score (FSS) threshold of 0.2mm/h and 35km scale

	T+1	T+2	T+3	T+4	T+5	T+6
No. and % of cycles 4D-Var worse than 3D-Var in FSS	11 (50%)	8 (36%)	9 (41%)	11 (50%)	7 (32%)	9 (41%)
No. and % of cycles 4D-Var better than 3D-Var in FSS	11 (50%)	14 (64%)	13 (59%)	11 (50%)	15 (69%)	13 (59%)
Mean Lmin (km) in 3D/4D-Var and improvement by 4D-Var $\frac{[(Lmin)_{3D-Var} - (Lmin)_{4D-Var}]}{(Lmin)_{3D-Var}}$	27/18 (33%)	40/29 (38%)	80/53 (34%)	107/73 (32%)	124/84 (32%)	133/110 (17%)

Table 2 shows that on average the 4D-Var system was performing better than the 3D-Var system with small mean Lmin and the same or more forecasts better than the 3D-Var system than vice versa. The forecasts from the 4D-Var system were significantly better than those from the 3D-Var system on some cycles and forecast time ranges. For instance at [T+0,T+1] or [T+3,T+4] ranges, the mean Lmin from the 4D-Var system was smaller (i.e. greater skill) than the 3D-Var system even though, on the individual cycle basis, the 3D-Var system had the equal number of cycles (i.e. 50%) performing better than the 4D-Var system.

To see how typical the 28th June 2012 case was, we also list the relative skill from the 3D-Var and 4D-Var system for the whole month of June 2012 in Table 3.

Table 3 Relative skill for one hour rainfall accumulations from the 3D-Var and 4D-Var forecasts for whole month June 2012 using threshold of 0.2mm/h and 35km scale

	T+1	T+2	T+3	T+4	T+5	T+6
4D-Var worse than 3D-Var in No. of cycles in FSS	288 (48%)	251 (41%)	266 (44%)	291 (48%)	280 (46%)	302 (50%)
4D-Var better than 4D-Var in No. of cycles in FSS	312 (52%)	358 (59%)	343 (56%)	318 (52%)	328 (54%)	306 (50%)
Mean Lmin (km) in 3D/4D-Var and improvement by 4D-Var $\frac{[(Lmin)_{3D-Var} - (Lmin)_{4D-Var}]}{(Lmin)_{3D-Var}}$	21/18 (14%)	25/21 (16%)	32/28 (12%)	37/34 (8%)	42/38 (10%)	44/42 (5%)

Table 3 shows again that there were more cycles when the forecasts from the 4D-Var system had better FSS scores than the forecasts from the 3D-Var system for the month of June 2012. The improvement in mean Lmin by the 4D-Var over the 3D-Var forecasts ranged from 5% at [T+5,T+6] to 16% at [T+1,T+2]. We may see that, by comparing Table 2 and 3, the case of 28th June 2012 is clearly a valid example illustrating the 4D-Var system's improved forecast skills compared to the 3D-Var system.

7. Concluding remarks

The dependence of the skill of an hourly cycling high resolution NWP-based nowcasting system on forecasting time scale of 0-6 hours on the method of generating its initial conditions has been investigated. The performance of the model skill in precipitation forecasts has been assessed from a sample of 720 forecasts for June 2012 by an objective, scale-dependent verification method that determines useful spatial scales the model produces in hourly rainfall accumulations.

The NWP-based 1.5km nowcasting system with hourly cycle 4D-Var is shown to produce convective rainfall forecasts of added benefit when compared with 3D-Var NDP. The improved forecasts from the 4D-Var is likely due to the fact that it exploits the time history of the observations and provides a direct link to the dynamics of precipitations by virtue of the linear perturbation forecast (PF) model. Successful application of 4D-Var in an operational NWP nowcasting system demonstrates the promise for improved prediction of severe weather.

Results indicate that 1.5 km version of the Unified Model (UM) has a good capability in predicting convective-dominated weather events. By assimilating latest high spatial and temporal observations, the model achieves spatial scales of rainfall predictions which are useful in flood forecasting. It is shown that a DA system based on hourly cycling 4D-Var and nudging in a 1.5 km version of the UM produces improved skills in spatial distribution of hourly rainfall when compared with a 1.5 km model with hourly 3D-Var. For the 90th percentile, the 4D-Var system has a mean acceptable skill of 15 km at $[T+0, T+1]$, and of 37 km at $[T+5, T+6]$. By comparison, the 3D-Var NDP produces acceptable scales of ~18 km and ~39 km at $[T+0, T+1]$ and $[T+5, T+6]$. These skills between 4D-Var and 3D-Var system translate to an improvement of ~17% and ~5%, respectively.

The 4D-Var NDP ran robustly and worked satisfactory at convective scales despite increasing non-linearity in the system at that resolution. When the same observations are used as the FGAT 3D-Var, the 4D-Var NDP has better forecasting skills in hourly rainfall accumulation both at absolute and relative thresholds. A clear benefit of using 4D-Var is evident to justify its higher running expense to make full use of high spatial and temporal observations. At convective scale resolution and with hourly assimilation, FGAT is still beneficial for 3D-Var using sub-hourly observations compared with traditional 3D-Var using only the observations closest to the nominal analysis time.

The results have also indicated that 1.5 km model has a positive model bias in the amounts of precipitation. The present model, like other convective scale models, assumes the resolution is fine enough so that convection can be resolved explicitly, whereas many of the actual convection events suggest that is not the case. The model has a tendency to produce too intense peaks at the centre of rain bands or convective storms, as well as too few small scale convective cells, leading to an over-prediction in the amounts of precipitation. However, good prediction of the occurrence of rainfall events and the actual locations of the rain, as shown earlier by the minimum scale for relative thresholds in the FSS verification, are of great importance to an end user if the model bias in the rainfall amounts is kept in mind. In this respect, the model provides a useful tool.

Although the NDP system has smaller correlation length scales in the background error covariances than was used in previous operational UK domain data assimilation systems, a more appropriate representation of the background error covariances may be required for convective scale models to utilise the full potential of 4D-Var. Balance constraints such as hydrostatic and geostrophic balance which are traditionally built into the covariance model in large scale models, have been used here. These assumptions are clearly less likely to apply in areas of active convection, and a new method has to be developed to define the balance relationship valid at high resolution so that appropriate background error covariance matrix can be modelled for convective scale models.

To make further use of high resolution spatial and temporal data in our system, work is underway to develop a method of assimilating radar reflectivity within 4D-Var. More time frequent conventional observations which are potentially available for use in the 1.5 km hourly nowcasting NWP system include GNSS water vapour and aircraft MODE-S temperatures and winds and sub-hourly surface observations. Success in assimilation of higher resolution (conventional and novel types) data and an improvement in the background error covariances are likely to utilise the potential of the 4D-Var further still.

Acknowledgements

We thank Professor Andrew Lorenc for his useful comments which improves the paper.

References

Ballard SP, Li Z, Simonin D, Caron J-F. 2015. Performance of 4D-Var NWP-based nowcasting of precipitation at the Met Office for summer 2012. *Q. J. R. Meteorol. Soc.*, (submitted)

Bloom SC, Takaks LL, Da Silva AM, Ledvina D. 1996. Data assimilation using incremental analysis updates. *Mon. Weather Rev.*, **124**: 1256-1271.

Bowler NE, Pierce CE, Seed AW. 2006. STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, *Q. J. R. Meteorol. Soc.*, **132**: 2127-2155.

Browning KA, Collier CG. 1989. Nowcasting of precipitation systems, *Rev. Geophysics*, **27**: 345-370.

Davies T, Cullen MJP, Malcolm AJ, Mawson MH., Staniforth A, White AA, Wood N. 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.*, **131**: 1759-1782.

Dixon M, Li Z, Lean H, Roberts N, Ballard S. 2009. Impact of data assimilation on forecasting convection over the United Kingdom using a high-resolution version of the Met Office Unified Model. *Mon. Weather Rev.*, **137**: 1562-1584.

Germann U, Zawadzki I. 2002. Scale-dependence of the predictability of precipitation from continental radar images. Part 1: Description of the methodology, *Mon. Weather Rev.*, **130**: 2859–2873.

Golding BW. 1998. Nimrod: a system for generating automated very short range forecasts, *Met. Apps.*, **5**: 1-16.

Golding BW. 2000. Quantitative precipitation forecasting in the UK. *J. Hyd.*, **239**: 286- 305.

Honda Y, Nishijima M, Koizumi K, Ohta Y, Tamiya K, Kawabata T, Tsuyuki T. 2005. A pre-operational variational data assimilation system for a non-hydrostatic model at the

Japan Meteorological Agency: Formulation and preliminary results, *Q. J. R. Meteorol. Soc.*, **131**: 3465-3475.

Ide K, Courtier P, Ghil M, Lorenc AC. 1997. Unified notation for data assimilation: Operational, Sequential and Variational *J. Meteorol. Soc. Japan*, Special issue "Data Assimilation in Meteorology and Oceanography: Theory and Practice." **75**, No. 1B: 181-189.

Lean HW, Clark PA, Dixon M, Roberts NM, Fitch A, Forbes R, Halliwell C. 2008. Characteristics of high-resolution versions of Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Weather Rev.* **136**: 3408-3424.

Li L, Schmid W, Joss J. 1995. Nowcasting of motion and growth of precipitation with radar over a complex orography. *J. Appl. Meteor.*, **34**: 1286-1300.

Ligda MG. 1953. The horizontal motion of small precipitation areas as observed by radar. Technical Report 21, Department of Meteorology, M.I.T., Cambridge Massachusetts, 60 pp.

Lorenc AC, Ballard SP, Bell RS, Ingleby NB, Andrews PLF, Barker DM, Bray JR, Clayton AM, Dalby T, Li D, Payne TJ, Saunders FW. 2000. The Met. Office global three-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.*, **126**: 2991-3012.

Lorenc AC, Rawlins F. 2006. Why does 4D-Var beat 3D-Var? *Q. J. R. Meteorol. Soc.* **131**: 3247-3257. doi: 10.1256/qj.05.85

Mueller C, Saxen T, Roberts R, Wilson J, Betancourt T, Dettling S, Oien N, Yee J. 2003. NCAR Auto-Nowcast system. *Wea. Forecasting*, **18**: 545-561.

Pierce C, Seed A, Ballard S, Simonin D, Li Z. 2012. Nowcasting, chapter in "Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications", Dr. Joan Bech (Ed.), ISBN: 978-953-51-0496-4, InTech, DOI: 10.5772/39054. Available from: <http://www.intechopen.com/books/doppler-radar->

[observations-weather-radar-wind-profiler-ionospheric-radar-and-other-advanced-applications/nowcasting](#)

Rawlins FR, Ballard SP, Bovis KR, Clayton AM, Li D, Inverarity GW, Lorenc AC, Payne TJ. 2007. The Met Office global four-dimensional data assimilation system. *Q. J. R. Meteorol. Soc.* **133**: 347-362.

Roberts NM, Lean HW. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Mon. Weather Rev.*, **136**: 78-97.

Saito K, Fujita T, Yamada Y, Ishida J, Kumagai Y, Aranami K, Ohmori S, Nagasawa R, Kumagai S, Muroi C, Kato T, Eito H, Yamazaki Y. 2006. The operational JMA nonhydrostatic mesoscale model, *Mon. Weather Rev.*, **134**: 1266-1298.

Seed AW, Pierce CE, Norman K. 2013. Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resources Research*, **49**: 6624-6641.

Seity Y, Brousseau P, Malardel S, Hello G, Benard P, Bouttier F, Lac C, Mason V. 2011. The AROME-France convective scale operational model. *Mon. Weather Rev.*, **139**: 976-991.

Simonin D, Ballard SP, Li Z, Caron, J-F. 2014. Doppler radar radial wind assimilation using an hourly cycling 3D-Var with an 1.5km resolution version of the Met Office Unified Model for Nowcasting, *Q. J. R. Meteorol. Soc.*, doi:10.1002/qj.2298

Stephan K, Klink S, Schraff C. 2008. Assimilation of radar derived rain rates into the convective scale model COSMO-DE at DWD. *Q. J. R. Meteorol. Soc.* **134**: 1315-1326.

Sun J, Xue M, Wilson JW, Zawadzki I, Ballard SP, Onvlee-Hooimeyer J, Joe P, Barker DM, Li P-W, Golding B, Xu M, Pinto J. 2014. Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges, *Bull. Amer. Meteor. Soc.*, **95**: 409-426. doi: 10.1175/BAMS-D-11-00263.1

Witt A, Johnson JT. 1993. An enhanced storm cell identification and tracking algorithm, *Preprints, 26th International Conference on Radar Meteorology*, pp. 514–521, Norman, OK, American Meteorological Society.

Met Office
FitzRoy Road, Exeter
Devon EX1 3PB
United Kingdom

Tel (UK): 0370 900 0100 (Int): +44 1392 885680
Fax (UK): 0370 900 5050 (Int): +44 1392 885681
enquiries@metoffice.gov.uk
www.metoffice.gov.uk