METEOROLOGICAL APPLICATIONS
*Meteorol. Appl.* **23**: 165–181 (2016)
Published online 29 March 2016 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/met.1538

RMetS
Royal Meteorological Society

# Review

# Convection-permitting models: a step-change in rainfall forecasting

Peter Clark,[a] Nigel Roberts,[b] Humphrey Lean,[b] Susan P. Ballard[b] and Cristina Charlton-Perez[b*]

[a] *Department of Meteorology, University of Reading, UK*
[b] *MetOffice@Reading, Meteorology Building, University of Reading, UK*

**ABSTRACT:** Convection-permitting models (CPMs) have provided weather forecasting centres with a step-change in capabilities for forecasting rainfall. They are now used operationally to forecast precipitation in many parts of the world, including the UK. CPMs are models in which the dynamics of atmospheric convection is treated with sufficient accuracy in order to make it viable to switch off convection parametrization. This review describes the current state-of-the-art in operational CPM-based numerical weather prediction (NWP), primarily within the UK, and the historical development of CPMs. The characteristics of CPM systems and forecasts are highlighted and placed in an international context to recognize similar trends and highlight some differences. It is shown that the realism of CPM-based forecasts can provide improved subjective guidance on convection, and, when measured on appropriate scales, can improve rainfall forecasting skill compared to coarser-resolution NWP. Data assimilation techniques used with operational CPMs are reviewed and given historical context. Examples of new types of observations that may increase the skill of forecasts from improved initial conditions are discussed. CPM-based nowcasting systems are shown to provide considerable improvements in short-range forecasts of rapidly developing, intense systems. As a result, these CPM-based systems provide a new forecasting capability. Finally, the development of CPMs has also required new techniques to verify forecasts and define their skill. These have revealed that the lack of predictability of the smallest scales involving convection means that ensemble techniques are required to represent forecast uncertainty, resulting in a new capability to provide objective forecast probabilities of local precipitation.

KEY WORDS convection-permitting models; high-resolution NWP; rainfall forecasts/forecasting; explicit convection; high-resolution data assimilation; resolved processes; convective ensembles

*Received 19 December 2014; Revised 10 July 2015; Accepted 25 July 2015*

## 1. Introduction

This review article discusses the impact of the introduction of so-called convection-permitting numerical weather prediction (NWP) models (CPMs) on the rainfall forecasting process. The objective is to review the UK experience within the Met Office of developing and using forecasting systems over the UK. These developments are placed in the context of similar ones occurring elsewhere to emphasize that the trend towards use of CPMs is becoming common, but no attempt is made at a comprehensive review of the systems used outside the UK.

NWP can be broadly defined as the process of defining initial conditions based on observations of the current state of the atmosphere and then integrating forward in time using a numerical approximation to the physical laws governing atmospheric dynamics and thermodynamics, including approximate representations of physical processes such as cloud physics, turbulence and radiation. In general, in modern systems, the initial conditions use a previous short-range forecast as a 'first guess' in order to benefit from the impact of previous observations. This must be achieved much faster than real time to make useful forecasts. The fundamental concept goes back to the seminal work of Richardson (1922), and weather forecasts have been produced using NWP for over 50 years. However, historically, most of the focus of NWP research, development and application has been on synoptic and larger scales. It is well known that these scales are characterized by hydrostatic and geostrophic balance (with the exception, in the latter case, of the equatorial region). As a result, until recently, the majority of NWP models were based on the hydrostatic primitive equations (HPEs) (White *et al.*, 2005).

HPE models have been hugely successful in revolutionizing weather forecasting. However, they suffer from the restriction that, as a result of the hydrostatic approximation, they are unable to represent accurately the dynamics of many smaller-scale processes that contribute dramatically to the local weather. This particular restriction adversely affects the simulation of smaller-scale flows over hills and mountains as well as the dynamics of deep convective storms. Therefore, the use of HPE models should be restricted to scales significantly larger than those of deep convection, which means they have to represent these smaller-scale phenomena through some form of parametrization. This is discussed further in Section 2. At this stage it is sufficient to state that the explicit representation of many important weather phenomena requires both considerable

* Correspondence: C. Charlton-Perez, MetOffice@Reading, Meteorology Building, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK. E-mail: c.charlton-perez@metoffice.gov.uk

computer power, to run the appropriate model at high enough resolution, and models formulated using a form of non-hydrostatic dynamics and suitable physical parameterizations.

Non-hydrostatic, 3D numerical models emerged from the research community in the 1970s and the 1980s. This was in the form of models of individual convective clouds or convective systems (e.g. Miller and Pearce, 1974; Schlesinger, 1975; Klemp and Wilhelmson, 1978) as well as models designed for more general mesoscale studies, such as sea breezes and orographic flows including gravity wave propagation (e.g. Pielke, 1974; Tapp and White, 1976; Anthes and Warner, 1978; Tripoli and Cotton, 1980; Tripoli and Cotton, 1982; Tremback *et al.*, 1985; Pielke *et al.*,1992). These developments led to a vast mesoscale modelling literature, covering a wide range of topics from idealized process studies to quasi-operational forecasting.

The Met Office was one of the first National Meteorological Services to experiment with mesoscale NWP. This was developed from a model based on the Tapp and White (1976) non-hydrostatic dynamics. However, computer resources available to the Met Office at that time meant that operational forecasts with this model were restricted to a 15 km horizontal grid length and so, despite having non-hydrostatic dynamics, a parameterization of convection (Fritsch and Chappell, 1980; Golding, 1990) was still required. Similar constraints applied to other national weather services. Therefore, operational 'mesoscale' forecasting systems were generally limited to resolutions requiring convection parametrization, with CPM resolutions restricted to a research context, until beyond the end of the 20th Century.

Arguably, the first CPM-based forecasting system used in real time was developed at the Center for Analysis and Prediction of Storms (CAPS), established in 1989 at the University of Oklahoma, USA. Its ultimate vision was to make available a fully functioning storm-scale NWP system (Droegemeier, 1990; Lilly, 1990) for tornado prediction, which became known as the advanced regional prediction system (ARPS) (Xue *et al.*, 2003). In 1999, for example, it was run for a period of expected tornado outbreaks in a triple nested mode with the finest grid resolution at 3 km. In the early 1990s, the Tapp and White model was extended (Golding, 1992, 1993) to overcome the time-step restrictions that had been imposed by the original equation set that supported sound waves. It was modified to be semi-implicit and semi-Lagrangian and used for numerical investigation of tropical island thunderstorms at 3 km resolution without parametrized convection (Golding, 1993). Golding (1993) showed that a general purpose, mesoscale model could be configured to reproduce qualitatively the diurnal evolution of the Hector convective system (Crook, 2001) that forms over the Tiwi Islands, Northern Territory, Australia.

In spite of the challenging title of Lilly's (1990) paper, the Met Office realized that computing resources for UK-wide domain operational CPMs were not, in fact, affordable in the 1990s. Instead, the Met Office took the opportunity to rationalize its climate and NWP forecasting capability to use the same basic model formulation for both climate and NWP. Hence, the Met Office operational non-hydrostatic model was replaced with a version of the hydrostatic Unified Model (MetUM) (Cullen, 1993) and the Golding (1992) formulation was never used operationally. However, recognizing the future need for high-resolution models, the Met Office started developing a new non-hydrostatic, deep atmosphere dynamical core in 1995 (Davies *et al.*, 2005) that could be used for both large scale climate and kilometre or metre scale models. This formed the basis of a new version of the MetUM that produced its first routine forecasts in 2003.

The Met Office started development of physical parameterizations suitable for CPMs in 2000. Work on data assimilation (DA) specifically for CPMs followed soon after. Considerable effort went into evaluation and testing (and still does!). Included in this, as will be discussed further below, was the development of new verification methods to meet the considerable challenge of verifying forecasts from such models. The first CPM-based operational forecasts at the Met Office were made in 2005 using a 4 km configuration of the MetUM over the UK and surrounding waters, called the UK4 model. This was supplemented by a sub-UK regional 1.5 km 'on-demand' system in 2006. In 2009, a UK-wide model with variable resolution, called the UKV model (Tang *et al.*, 2013), having a 1.5 km grid over the UK, stretching to 4 km at the boundaries, was introduced. This replaced the 4 km model operationally in 2012.

Similar developments were happening elsewhere over a similar period, including the development of the Weather Research and Forecasting (WRF) model (Michalakes *et al.*, 2001), COnsortium for Small-scale MOdelling (COSMO) (Baldauf *et al.*, 2011), Applications of Research to Operations at MEsoscale (AROME) (Seity *et al.*, 2011) and the Japan Meteorological Agency (JMA) Non-hydrostatic Mesoscale Model (Saito *et al.*, 2006). Each of these model systems, which generally include data assimilation and associated observation processing, is a major development, and most systems, including the Met Office MetUM system, serve a large community of users and developers in many countries.

The design of a forecasting system, including model configuration, provision of initial conditions through assimilation of observations, provision of lateral and surface boundary conditions, forecast duration, forecast post-processing, dissemination and verification, depends upon the desired application. At present, all CPMs are too expensive to be run globally, so lateral boundary conditions must be provided by a larger scale (ultimately global) model. At longer forecast times the driving model (providing the lateral boundary conditions) begins to dominate the forecast. This timescale depends on the domain size and meteorological regime, but for the UK system is on the order of 1 day. For shorter forecasts than this, refinement of the initial conditions through high-resolution data assimilation has a significant positive impact on forecasts. For forecast times beyond that point, the system becomes essentially a non-linear down-scaling of the larger-scale forecast.

A pure down-scaling NWP system is one in which no data assimilation is done (apart from that done in the larger-scale model it is down-scaled from); thus, forecasts are 'spun-up' from larger-scale initial conditions. Although the larger-scale forecast constrains the down-scaled product, the down-scaling may still add considerable value in circumstances where small-scale variability is important, for example, where there is locally organized convection or orographic enhancement of rain (Roberts *et al.*, 2009). Given the same large scale and surface boundary conditions, those fine-scale features would be very difficult to predict using statistical down-scaling techniques. In CPMs, down-scaling alone can add value by non-linear spin-up of higher resolution features and also through the use of higher-resolution orography and surface characteristics.

The spin-up from smooth initial conditions is, of course, unphysical, and can have a deleterious impact on the forecast for a substantial time in unfavourable conditions (e.g. if the initial state should include highly active organized convection), but in most circumstances the spin-up period lasts just a few hours (depending on model, resolution and circumstances). This justifies down-scaling for longer forecasts. Unfortunately, a related problem occurs as the smooth flow provided by lateral boundary conditions spins up as it enters the domain throughout the forecast. A few hours spin-up time from smooth initial conditions is equivalent to air moving a distance of order 100 km at speeds of $\sim$10 m s$^{-1}$. This is the primary motivation for the use of variable resolution in the Met Office's UKV model. The boundaries can be extended much further away from the key area of interest without incurring excessive additional cost.

A traditional data assimilation approach for short-range forecasting over periods of 1–2 days has been extended for use with CPMs. This is used to refine intermediate scales, those between the cloud scale and those well-represented in the coarser model providing boundary conditions. Using a continuous assimilation scheme, the assimilation process at intermediate scales updates the high-resolution model state; this reduces, or removes, the period of 'spin-up' of smaller scales in the forecast and retains consistency with the intermediate and larger scales. Typically a 3 h assimilation cycle is used, new forecasts are run every 3 h which means that the 'first guess' forecast into which observational information is added is 3 h old.

However, now that the rain predicted by NWP models at CPM resolution appears much more realistic and similar to radar-derived rain rate composites, it has become feasible to make direct use of CPM for nowcasting (frequently updated high resolution forecasts for periods up to 6 h). Lorenz (1969) suggested that convective storms on scales of 10 km or less have predictable timescales of only a few hours or less. Nowcasting systems attempt to forecast deterministically at the convective-scale (or, at least, convective-cluster scale) and are designed to achieve predictability of rainfall on the convective-storm scale on the order of hours. This suggests the need for data assimilation at high-resolution scales in both space and time, such as hourly-cycling CPM data assimilation producing 6 h forecasts every hour. Naturally, the research focus in NWP-based nowcasting has been on making use of observational data available at these time and space scales. These data are primarily gathered from weather radar and, more recently, from geostationary satellites, though new observation types are beginning to be exploited (see Section 4).

The Met Office has employed all three strategies for use of CPMs: down-scaling (no data assimilation), 'traditional' data assimilation and nowcasting. The operational global model ($\sim$25 km resolution at UK latitudes and recently updated to 17 km resolution) provides the large-scale forecast environment for all of these strategies. The Met Office Global and Regional Ensemble Prediction System (MOGREPS) drives a 2.2 km down-scaling ensemble of runs over the UK called MOGREPS-UK. A purely down-scaling 4 km model (Euro4) is run over Western Europe. The operational 1.5 km UKV model currently operates traditional data assimilation with a 3-hourly 3D-Variational (3D-Var) analysis cycle with a forecast length of 36 h and is targeted at intermediate scales. An experimental 1.5 km 'Nowcasting Demonstration Project' (NDP) was run during 2012–2013 over the southern half of the UK using hourly cycling, 4D-Variational (4D-Var) (Golding *et al.*, 2014; Sun *et al.*, 2014; Ballard *et al.*, 2015) and UKV boundary conditions. The relative impacts of down-scaling and data assimilation for the UK4 and UKV models were investigated by Dow and Macpherson (2013).

High quality and comprehensive research observations provide essential information to help assess and improve model performance. The cost of gathering together a large number of instrumentation systems, often including ground-based radars, wind profilers, lidars, specialized research aircraft, high-density conventional surface observations and high time-frequency radiosonde ascents almost invariably means that they are collected during field campaigns of limited duration. Several have occurred over past decades, many in the USA, and their data remain valuable for years or even decades after collection. Over the last decade, a number of European campaigns have directly contributed to the development of CPMs (Blyth *et al.,* 2015). In general, observational campaigns provide too few cases to be useful for rigorous statistical analysis of CPM forecasting performance. However, case studies have been invaluable for separating systematic errors due to the convective-scale performance of models from forecast errors caused by larger scale errors in the atmospheric state. The larger-scale errors in convective rain are often associated with mesoscale features in the mid- and upper-level wind and temperature pattern (upper-level potential vorticity anomalies) or in the low-level thermodynamics, especially in the moisture distribution. Therefore, case studies gathered from an observational campaign can provide confidence that, given sufficiently accurate large-scale conditions, the CPM will provide a useful convective response.

This review article will concentrate on results from the 1.5 km and 2.2 km models used for NWP over the UK. Section 2 will discuss the representation of physical processes in CPMs and Section 3 will discuss the qualitative behaviour of CPMs with examples from cases in the UK. Data assimilation in the current Met Office CPMs will be discussed in Section 4 and quantitative verification will be addressed in Section 5, whilst use of the ensemble forecasting system to examine predictability will be discussed in Section 6. Section 7 looks towards the future of CPM use, and development and Section 8 provides a summary.

## 2. Convection-permitting model physics

The key difference between larger scale models and CPMs is that the latter allow structures recognizable as convective clouds to form, whilst the former represent the effects of convection through some form of parametrization. It is not always clear exactly what assumptions underlie so-called convective parametrizations, however a key idea is that all of the cloud 'circulation' is included in the parametrization, so no flow identifiable as convection develops in the model. In itself, this is a troublesome assumption as, in practice, it is not clear how the cloud circulation can be fully separated from the large-scale vertical motion. This aside, it is clear that the parametrization represents scales, at the very least, up to the inter-cloud spacing (i.e. the cloud up-draught and the local compensating subsidence).

Schemes stemming from the quasi-equilibrium ideas of Arakawa and Schubert (1974), including the MetUM's own Gregory and Rowntree (1990) scheme, average over an 'ensemble' of clouds sufficiently large that such the mean is independent of the sample of clouds taken. In practice, if the condition can be achieved at all, this requires an average over at least $100 \times 100\,\text{km}^2$. Schemes that are claimed to be designed for smaller grid squares, such as the Kain–Fritsch scheme (Kain and Fritsch, 1993), while differing greatly in their treatment of the cloud updraughts and downdraughts, mass-flux closure and, in particular, triggering, assume that one is averaging over both cloud updraughts and the subsiding environment. As a result, all these schemes can only predict the area-average rainfall. It may be possible to parametrize the spectrum of rainfall rates, but, even then, only on the assumption that the spectrum of rates is spatially uniform.

In general, current convection schemes are 'diagnostic' or 'first order', they assume that the timescale for the cloud field to adjust to a change in forcing (e.g. solar heating) is much shorter than the timescale of the change in forcing. Thus, they adjust instantaneously to changes in conditions, leading to unrealistic behaviour at places such as coasts (e.g. where showers over the sea are not advected inland) and unrealistic behaviour in time, in particular in simulating the diurnal cycle.

The purpose of convective parameterizations is to represent most, or all, of the convective-scale motion and thus prevent unstable growth of cloudy structures on the grid. CPMs do precisely the reverse in that they allow such unstable growth of convective clouds to take place in the model. Latent heat release through condensation may allow a region of the model to become buoyant with respect to its surroundings, such that explicit vertical circulations develop. In practice, such models often contain a parameterization of partial cloud cover, enabling condensation to occur when the 'resolved' flow is not yet saturated. The cloud flow field includes updraughts, cloud-scale downdraughts and larger-scale subsidence (equivalently larger-scale gravity waves). Initial cloud formation may be explicitly caused by features such as orography, coastal convergence lines and convergence due to land surface heterogeneity or convergence lines associated with gust fronts from earlier clouds. Each cloud has a life-cycle and this life-cycle interacts with that of other clouds, so the cloud field may evolve in time, perhaps leading to self-organization. CPMs are able to reproduce the positive feedback between precipitation formation and the development of deep convection *via* the formation of cold pools from evaporation of precipitation (Khairoutdinov and Randall, 2006), although the timescales on which this occurs may not be perfectly simulated because of limited spatial resolution.

So far, the discussion has implied a clear distinction between coarse model resolutions that require convective parametrization and CPMs that do not. In practice it is not that clear-cut. For example, a model with 4 km grid spacing should not be able to represent adequately the structure of small showers, but may be able to do a decent job for larger storms. For that reason, the UK4 model introduced by the Met Office in 2005 included the use of a convection parametrization scheme, but in a restricted way (Roberts, 2003a). The intention was to allow the parametrization to deal with some of the convection that cannot be resolved whilst leaving the model to remain convection-permitting for the larger showers that should be represented on that grid. This approach, now used in the Met Office European 4.4 km model, gives better results than running without convection parametrization and can adapt to the grid spacing. However, parametrization in this so-called grey-zone, where clouds are partially resolved (or large clouds are resolved, small ones not) is recognized as a major research challenge (see, for example, Gerard *et al.*, 2009; Yu and Lee, 2010; Arakawa and Wu, 2013).

CPMs require a fairly sophisticated treatment of cloud microphysics, sufficient to interact both with the updraught life-cycle and downdraught generation. The terminal velocity of hydrometeors (compared with horizontal wind velocities and grid scale) is a key factor controlling the sophistication required in a microphysics scheme. Particles that reach the surface in a time similar to the transit time of air across a grid box can be treated 'diagnostically'; that is, the horizontal advection of particles by the wind can be ignored and a local 1D budget can be used to calculate the vertical distribution of the hydrometeors. Thus, in large-scale models only cloud water needs some form of explicit (prognostic) treatment, which may include some measure of its variability in a grid box. Because mesoscale models were developed to run at higher resolution, ice and snow were added to the list of hydrometeor species treated prognostically. For example, Wilson and Ballard (1999) treated ice and snow together as one prognostic variable in the MetUM. A typical terminal velocity for snow (1 m $s^{-1}$) means that it can travel 50 km horizontally in a 10 m $s^{-1}$ wind, so the prognostic scheme is required at grid lengths of around 10 km. Similarly, CPMs must treat rain prognostically (5–10 m $s^{-1}$ terminal velocity leading to several km horizontal movement as it falls to the ground). Treatment of graupel may still be diagnostic in CPMs, but often it is not. In addition to more species, more variables are often used to describe the evolving size distribution, typically using the moments of the distribution.

As yet, no single microphysics scheme has been shown to yield the most accurate results in all circumstances, and some schemes may be particularly applicable to certain weather regimes. Some models include multiple schemes, which facilitate direct comparison; the WRF model, for example, contains at least 12 schemes (depending on how they are counted, as there is significant overlap between some schemes). Many sensitivity studies have shown that the simulation of convective storms is sensitive to microphysics, *via* numerous interacting mechanisms. However, demonstrating a systematic impact on forecast skill is much more challenging and some studies (e.g. Leoncini *et al.*, 2013) have shown that it can be difficult to distinguish systematic impact of microphysics from the random impact of turbulent fluctuations.

The Met Office system used an adaptation of the Wilson and Ballard (1999) scheme including prognostic rain and graupel (Wilson and Forbes, 2004) though the prognostic graupel was not found necessary for the majority of (UK) forecasts (Forbes and Halliwell, 2003). It has recently migrated to a more flexible, multi-moment scheme (Wilkinson *et al.,* 2011). Although prognostic graupel was not operational in 2012, the UKV model was re-run for 27 July 2012 using that scheme in order to produce a mass-mixing ratio of graupel for use in a lightning-flash rate diagnostic. This special lightning forecast gave good guidance to those forecasting for the Opening Ceremony of the 2012 Olympic Games who were concerned about the severity of the convective showers they were expecting on that day over London (Wilkinson and Bornemann, 2014).

Real clouds have important motions on scales much smaller than those resolved by the horizontal grid lengths (1–5 km) of current CPMs. Resolutions of 100 m or higher might be needed to 'properly' resolve deep convective clouds (Bryan *et al.*, 2003). In practice, CPMs need 3D parameterizations of cloudy turbulence. The effects of these parameterizations are poorly understood, but they serve to control the scales of motion (both in-cloud and between clouds) that develop. In their absence, cloud motion would still be controlled by the dissipation implicit in the model dynamics, which prevents energy accumulating on the grid scale (Skamarock, 2004).

A crude analogy is that CPMs develop clouds as if the atmosphere were much more viscous than it is (by perhaps eight orders of magnitude!). These model clouds tend to be larger, more widely spaced and slower growing than real clouds. The realism of the Met Office's CPM will be discussed in detail in Section 3. At this stage, however, it should be noted that one key attraction of CPMs lies in what they can explicitly resolve well; in particular, this is the organization of convection into mesoscale convective systems (MCSs). CPMs are capable of organizing convection because it would appear that the mechanisms leading to organization (initially, at least, propagation of gravity waves generated by convective heating and cooling) do not critically depend on the detail of the convective heating. Clark *et al.* (2014) report an example of a good forecast, in which relatively benign 'air mass' convection over the UK evolves into a small MCS which led to some localized flooding.

Interactions with the surface are also a key part of CPMs; in fact, the MetUM, since it is used as a climate model, already possesses a sophisticated surface and sub-surface scheme (Essery *et al.*, 2003), but even this has been refined for mesoscale application, particularly in its treatment of urban surfaces (Porson *et al.,* 2010).

Single-dimensional (1D) radiation schemes are generally still adequate, though some experimental work has been done using 3D schemes (which are very expensive). The main area where some enhancement to radiation schemes has been found beneficial is in its interaction with the surface. Oliphant *et al.* (2003) studied a particular case over the New Zealand Southern Alps in some detail and found a decreasing order of importance of slope aspect, slope angle, elevation, albedo, shading, sky view factor and leaf area index. The most important terms are relatively easy to understand. Resolved slope and aspect define the vector direction of the surface element and so have direct impact throughout the day from the solar short-wave (SW) direct beam. South-facing slopes receive more SW radiation than north facing slopes. On the other hand, shading and sky view factor have their main impact on SW at dawn and dusk when SW radiation is small anyway. In common with urban canyons, sky view factor is likely to have its largest impact on long-wave (LW) radiation.

Müller and Scherer (2005) reviewed the treatment of orographic impacts on radiation in a variety of mesoscale models. The majority do nothing (i.e. assume flat Earth). A significant minority include slope and aspect, and only ARPS (in the survey) include shadowing as well. The impact of resolved slope and aspect is relatively easy to retrofit to schemes; the MetUM now treats both slope and aspect. It is not so clear what benefit is actually gained in terms of forecast skill. Müller and Scherer (2005) showed 0.5–1 °C root mean squared (RMS) improvement in screen (2 m) temperature from an explicit sub-grid model, but this was over the Alps and may be extreme. It is also likely that the primary benefit was in the local temperature diagnosis, rather than in forecast evolution. Manners *et al.* (2012) developed a parameterization of the influence of topography on surface and atmospheric radiative fluxes and used this in the two-stream radiative transfer model that is part of the MetUM. By studying cases run on various convection permitting resolutions from 100 m to 1.5 km grid lengths, they concluded that surface slopes and terrain shadowing should be taken into account for the SW bands and sky-view factor should be considered for the LW bands.

## 3. Realism of convection-permitting models

Some of the major benefits of CPMs are illustrated in Figure 1, which shows rain rates forecast by the MetUM at three different model resolutions compared to radar-derived rain rates for a case of summer convective showers over the UK.

The global (then 25 km, now 17 km) resolution model produces convective rain from its convective parametrization scheme, which works independently at each grid point. The presence of convective rain in this context should be interpreted as ‘there will be convective showers in this general area’ rather than ‘this is what the showers are expected to look like’.

Although the global model forecast (Figure 1(b)) does indicate the general likelihood of showers and indicates their presence in roughly the right areas, it does not forecast the small-scale cells over southern Ireland or the localized bands of precipitation in northern England. It tends to predict widespread precipitation and has little indication of the areas which could expect larger rain rates; in fact, it does not give any indication of the rain over the Irish Sea.

In contrast, the Euro4 model (Figure 1(c)) was able to produce lines of convection and in generally the right area. However, in this example and typically, the showers are generally too few, too intense (Lean *et al.*, 2008) and too organized. In Figure 1(c), the Euro 4 produces just one and not two bands in northern England. In this case, the UKV model (Figure 1(d)) manages to reproduce the two lines of showers over the north of England and also attempts to produce the lighter individual showers, such as over southern Ireland, Wales and southern England. The organized line of showers extending down the Irish Sea to the tip of Cornwall is better in the Euro4 model than in UKV, but close examination shows this is more the result of excessive cell size than genuine organization.

The showers in the UKV model often look very realistic when compared to the radar-derived rain rates, but note that due to the relatively coarse grid length of the model compared to the typical sizes of real convective plumes, the showers (especially small showers) are still under-resolved. This can lead to evolution on the wrong scales.

Another important feature of explicitly modelled convection is that, due to the finite grid-length of the model, the convection may initiate later than it should, because the model cannot reproduce the very small initial convective plumes. The delay in the initiation is seen to improve as the model grid-length is reduced, because it is essentially less viscous, and is much less of a problem at ∼1 km than at 4 km (Lean *et al.*, 2008).

The fact that the showers are produced explicitly leads to another important benefit. If the output in Figure 1 was animated forward in time, then the global model convection scheme would react to the local column instability without direct memory of what went before, leading to a flickering on and off of the rain. The CPMs, on the other hand, are able to advect showers realistically with the flow as well as reproducing phenomena such as the formation of daughter cells and back-building of storms.

The explicit representation of convective showers gives many physical benefits. An example, shown in Figure 2, is showers of frozen precipitation being advected inland from the coast in winter. This case emphasizes how surface features, mainly land/sea contrast and orography, can influence convection. The UKV can simulate the high accumulation bands of precipitation quite well because the explicitly represented and organized showers can penetrate inland before dying out. However, the higher snow accumulations in the 12 km North Atlantic and European (NAE) model (now retired) are limited to the sea and coastal strip. This is a well-known problem in larger scale models with parameterized convection because the convective parameterization responds to the instability caused by the cold airmass over the warm sea, and as soon as the air comes over the colder land the convection stops. On this occasion, and in other similar situations, the UKV gives a much-improved forecast of showers affecting coastal areas. Improved location is a clear advantage; it is too early to assess the skill at predicting the phase of precipitation, in part because in showers it is not always easy to obtain reliable ground truth in marginal situations.

Figure 3 shows another example of a CPM producing more realistic convection than a lower resolution model using a convective parameterization. In contrast with the scattered showers shown in Figure 1, this case was of a mesoscale convective system (MCS) occurring overnight over southern England that produced localized flooding and there were several reports of lightning damage. On comparing with the radar composite, it is clear

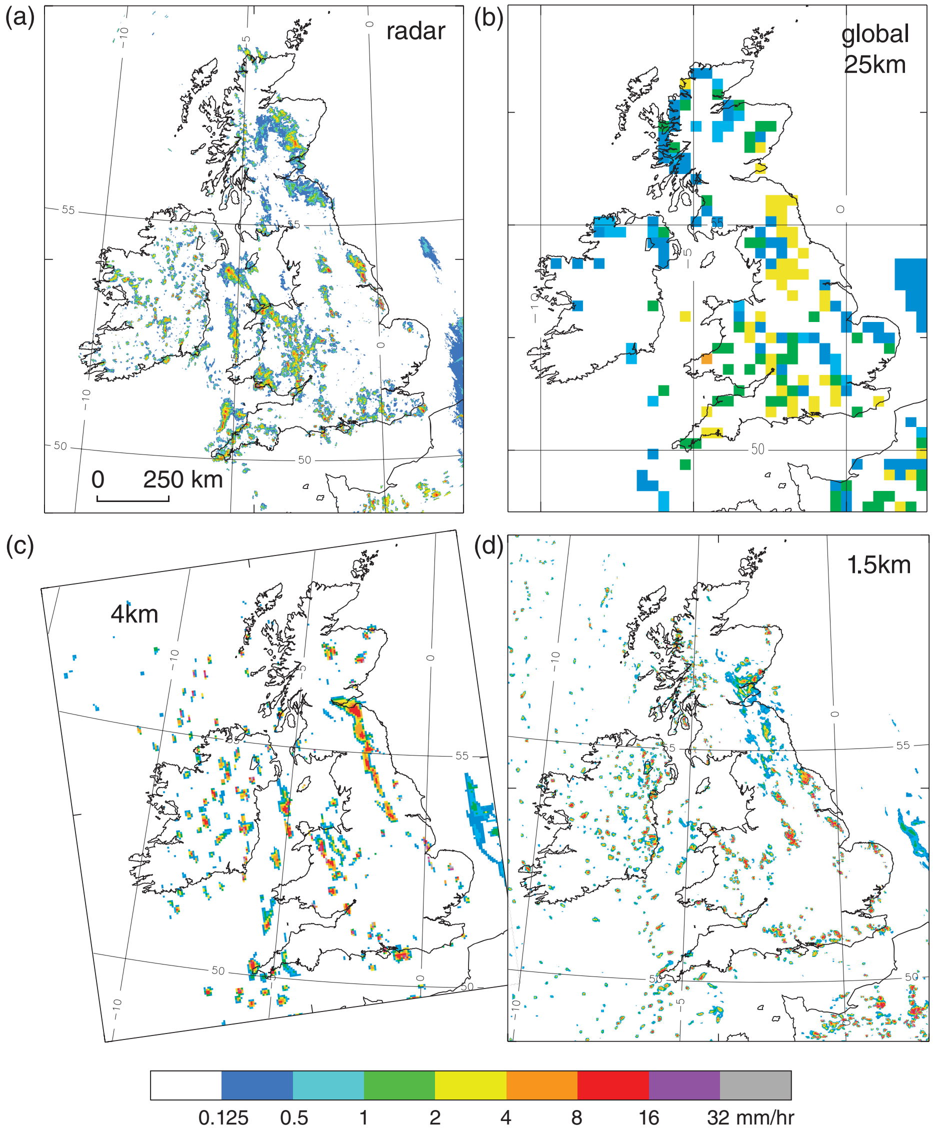


Figure 1. Case of convective showers over the UK on 8 July 2014 1200 UTC. (a) Radar-derived rainrates (mm $h^{-1}$) at 1 km resolution shown with instantaneous rainrates from the (b) MetUM Global model T + 12 forecast from 0000 UTC, (c) Euro 4 (4 km down-scaled) T + 12 forecast and (d) T + 9 UKV (1.5 km convection-permitting model) forecast from 0300 UTC.

that the global model shows no skill at representing this organized system of convective rainfall. There is no system present in the global simulation; just a few pixels of very light rainfall rates show up in the vicinity. However, the UKV simulates an area of organized convective rainfall developing from more scattered intense showers over central England during the evening and, once organized around midnight, propagating to the south at approximately the right speed. The magnitude of maximum rainfall rates are well forecast in generally the right region (perhaps 10–30 km too far east). This was a particularly good forecast and the ability of CPMs to replicate this type of storm system with this degree of realism holds the promise of more accurate quantitative precipitation forecasts of flood-producing rainfall events in future. Not all forecasts are this good (and forecasting MCSs remains one of the more challenging problems) and even the question of how the quality of a forecast from a CPM can be measured is difficult; verification of CPMs will be discussed later.

Although the examples given in this section have been focussed on the UK, similar considerations apply to the representation of convection in other parts of the world. The main difference is that the scales of interest may be different. For example, over the Great Plains in the USA the primary meteorology of interest is much larger convective systems that are much less common over the UK. Supercell convection with cloud scales of 10 km or more is much more common than over the UK, while MCSs such as squall lines that are ~100 km long are a regular occurrence. In that context, 4 km models are deemed to have a sufficiently fine grid (without the need for any convective parametrization). It is the view of many American scientists when assessing forecasts of such large convective systems that any benefits from using models with higher resolution than 4 km are outweighed by the additional cost (Kain *et al.*, 2008). This contrasts with experience in the UK where 1.5 km is clearly beneficial compared to 4 km resolution.

## 4. Data assimilation for convection-permitting models

Many National Meteorological Services, including the Met Office, are already using CPMs with some form of data

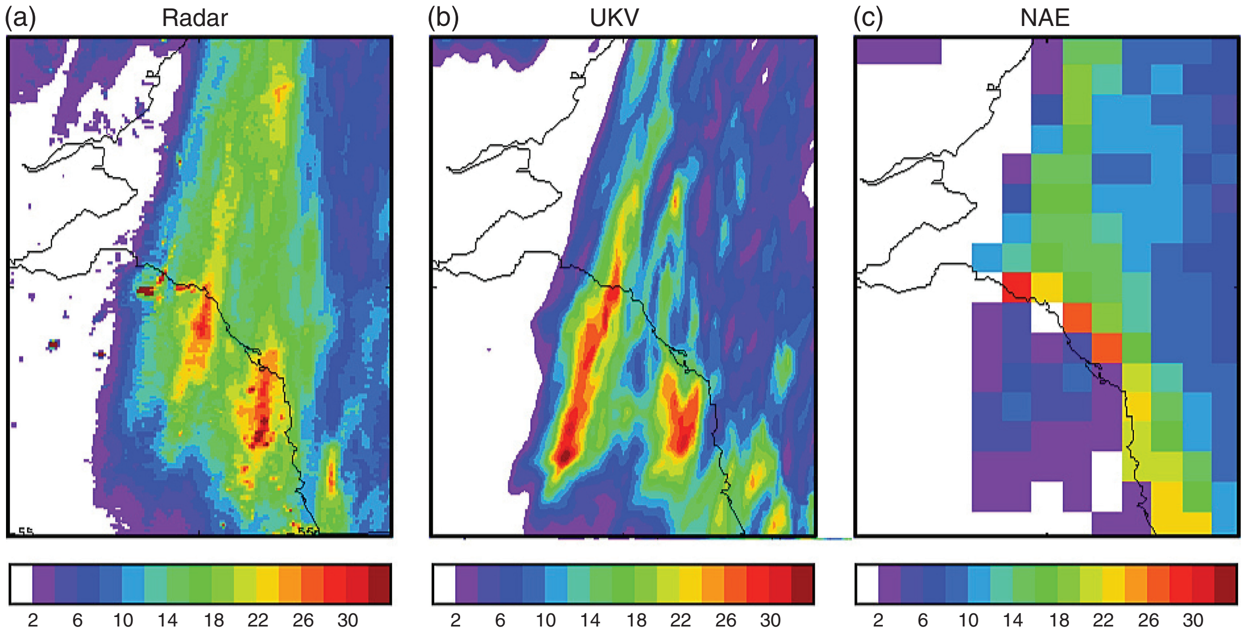


Figure 2. Snowfall penetrating inland over east coast of Scotland and northeastern coast of England in a northeasterly flow pattern. Images shown above are 24 h precipitation accumulations (mm) from 0300 UTC on 25 November to 0300 UTC on 26 November 2010. (a)–(c) The radar composite data at 5 km resolution, the UKV forecast from 0300 UTC 25 November 2010 and the NAE (12 km UM) forecast from 0000 UTC 25 November 2010.

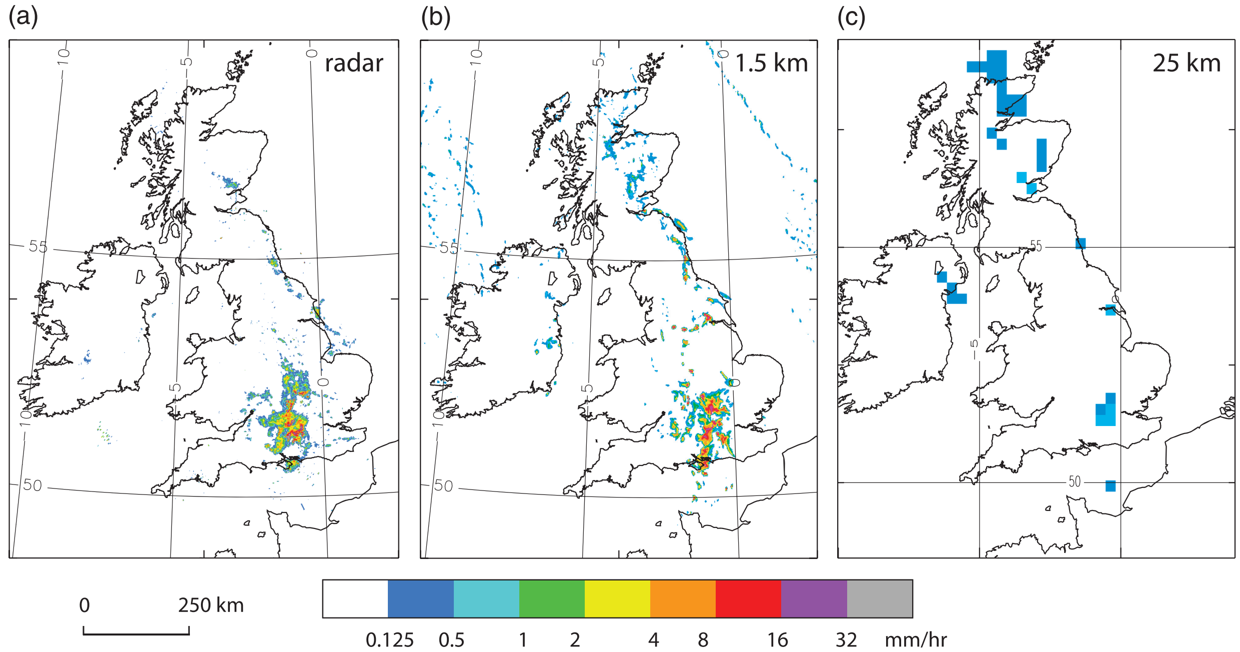


Figure 3. Mesoscale convective system (MCS) over the UK represented by instantaneous rainfall rates (mm h$^{-1}$) for 0000 UTC 14 June 2014. (a) Radar-derived rainrates at 1 km resolution are shown with (b) UKV T + 9 forecast started at 1500 UTC 13 June 2014 and (c) a global MetUM T + 12 forecast started at 1200 UTC 13 June 2014.

assimilation in operational NWP systems for short-range forecasting up to about T + 36 h, every 3 or 6 h. A variety of data assimilation methods are used or being adapted for CPMs. The Met Office system is discussed in more detail below. The German National Meteorological Service (DWD) use nudging at present (Stephan *et al.*, 2008), but are working on development of a local ensemble transform Kalman filter (LETKF) data assimilation system. The Met Office and Meteo France (Brousseau *et al.,* 2011; Seity *et al.,* 2011; Ballard *et al.,* 2012a, 2012b) use 3D-Var. In April 2015, Meteo France introduced an hourly cycling 3D-Var system with a 1.3 km resolution version of their AROME model. JMA now use a non-hydrostatic model at 5 km resolution with 4D-Var and a 2 km version with 3D-Var (Honda *et al.,* 2005; Saito *et al.,* 2006). Schwartz and Liu (2014) give evidence for the superiority of hybrid data assimilation over the ensemble Kalman filter (EnKF) and 3D-Var in the WRF at 4 km resolution and Li *et al.* (2015) have shown the benefit of hourly-cycling 4D-Var over 3D-Var in a 1.5 km MetUM system.

The capability to accurately forecast local weather such as fog, frost, cloud and precipitation is one of the key aims of convective-scale NWP. Therefore, it is important to initialize the model fields used to predict them, in particular the cloud, aerosols and surface and near-surface temperatures. Many operational NWP systems do not do this, especially within global model systems. However, the Met Office has found benefit from using surface data as part of the data assimilation strategy for UK NWP ever since the original non-hydrostatic, mesoscale model system (Golding, 1990; Ballard *et al.*, 1991) with its Interactive Mesoscale Initiation (IMI) system was introduced (Wright and Golding, 1990; Ballard *et al.*, 1991). Therefore, the Met Office has developed methods within its UK data assimilation system to make use of any available observations containing information on those fields, such as surface weather reports, radar-derived surface rain rates and Meteosat cloud-top images.

It is important to ensure dynamical consistency in the initial conditions, especially when trying to correct locations of convective precipitation. The operational replacement of the UK non-hydrostatic mesoscale model, with a version of the original hydrostatic MetUM (Cullen and Davies, 1991; Cullen, 1993) in 1992, had the advantage of introducing the use of an analysis correction scheme (Lorenc *et al.*, 1991) for the UK model. This provided a data assimilation scheme that ensured dynamical consistency. However, it lacked the use of satellite and radar data to adjust cloud and moisture fields which in turn adjust the precipitation forecasts. Therefore, schemes were developed to exploit cloud and radar-derived surface precipitation rate analyses (produced for use in the operational nowcasting system (Golding, 1998)) by assimilating them using moisture and latent heat nudging (LHN) in the Met Office 17 km UK model (Macpherson *et al.,* 1996; Jones and Macpherson, 1997).

This work in the 1980s and 1990s anticipated what is required to produce good initial conditions for CPM forecasts of cloud and precipitation: an ability to modify the initial conditions to add cloud and precipitation where it is absent in the previous forecast and to remove it where it is in the wrong place. At that time, when convection was parametrized rather than resolved, the moisture and LHN performed better for frontal situations than for convective systems. Over the years as the model resolution and formulation have changed so has the data assimilation method; however, the LHN (and until recently the moisture nudging) has been retained.

In 1999, a 3-hourly continuously cycling 3D-Var system, based on that used in the global model (Lorenc *et al.*, 2000), was introduced for use with the UK model, which was then a 12 km model. The analysis increments were gradually added to the forecast over a period of 3 h surrounding the analysis time (Bloom *et al.*, 1996) and the cloud/relative humidity and LHN schemes were retained to update the model along with the nudging of the 3D-Var analysis increments. With the move to operational convection-permitting NWP, using the UK4 and UKV models, 3-hourly cycling 3D-Var was retained. The UK4 model used a 4 km resolution analysis grid and the UKV model uses a 3 km resolution analysis grid. As mentioned above, they both retained LHN of hourly radar-derived rain rates. Whilst initially these systems also used moisture nudging (Dixon *et al.*, 2009) to exploit a 3-hourly 3D cloud cover analysis, by summer 2012 both of these systems were assimilating cloud directly in the variational assimilation systems *via* equivalent humidity (Renshaw and Francis, 2011).

The operational use of CPM models has allowed the radar rain rate and satellite cloud-top temperature information to be used at much higher horizontal resolution, closer to the resolution of the source data and with less smoothing. This provides better resolution of convective storms in the observations. Some of the assumptions underlying the LHN, such as latent heat release and precipitation at the same horizontal location, are expected to break down at convective scale. However Dixon *et al.* (2009), in experiments at 4 km resolution, showed that LHN was still beneficial to the forecasts with only slight modification. It appears that its use with explicit microphysics and dynamics for the convection, rather than a convection parametrization scheme, has outweighed any expected problems.

The UKV system also includes direct variational assimilation of Spinning Enhanced Visible and InfraRed Imager (Meteosat SEVIRI) radiances. The upper tropospheric water vapour channels 5 and 6 are assimilated for clear skies and over low cloud whilst window channels are assimilated for clear skies over sea only. The use of the SEVIRI data provides additional information on humidity fields which helps to constrain the cloud in the model.

Figure 4 shows the comparison of a 4 h down-scaled 4 km Euro4 forecast with a 7 h and 1 h forecast from the UKV model compared with the radar-derived surface rain rate composite (Figure 4(c)). The 4 km resolution forecast (Figure 4(d)) is missing the thunderstorms over England and the precipitation over Wales is too far south. The 7 h forecast from the UKV model (Figure 4(b)) has a better position for the precipitation in Wales and forecasts some precipitation in the band from mid-Wales to southeast England, but is missing the forward west-to-east line of storms. The forward band is picked up in the 1 h forecast (Figure 4(a)) from an analysis done once the storms are present. The 1 h forecast has benefitted from the LHN as well as the 3D-Var data assimilation of other observation types. On the whole, statistically, data assimilation improves the CPM forecasts, but not every case will show such a dramatic and positive benefit as seen in the case illustrated in Figure 4.

Whilst the operational 3D-Var for the UKV can exploit the higher spatial resolution of the radar-derived rain rates and satellite imagery, the UK operational radar network actually provides observations of reflectivity and Doppler radial velocity from the precipitation at high temporal and spatial resolution in the horizontal and vertical. The quality of the reflectivity data is being improved by the roll-out of dual-polarization across the network. Furthermore, geostationary satellite radiances, wind profilers, global navigation satellite system (GNSS) integrated column water and surface observations as well as radar data are available sub-hourly (Ballard *et al.,* 2015) and MODE-S aircraft temperatures and winds (de Haan 2011; de Haan and Stoffelen, 2012) are potentially available at high temporal resolution. Additional information can be extracted from the time variation of observations using 4D-Var, by providing a constraint on the evolution of the forecast over a fixed time-window. Representing rapid temporal changes becomes important for forecasting convection.

The Met Office uses 4D-Var for its global NWP (Rawlins *et al.,* 2007) at 25 km resolution (now 17 km resolution). Until the model was dropped from operational use, the Met Office also used 4D-Var for its North Atlantic and European (NAE) 12 km resolution limited-area model. Therefore, it is essential to establish whether 4D-Var is beneficial at convective scales in order to exploit the time variation of sub-hourly observations (the UKV model uses just hourly observations). Once operational use of CPMs was achieved, it was also a natural next step to test whether they could be used for nowcasting (6 h forecasts) using hourly as opposed to 3-hourly cycling NWP. Therefore, experiments were undertaken for a domain covering southern England and Wales with hourly-cycling 3D-Var (Ballard *et al.,*

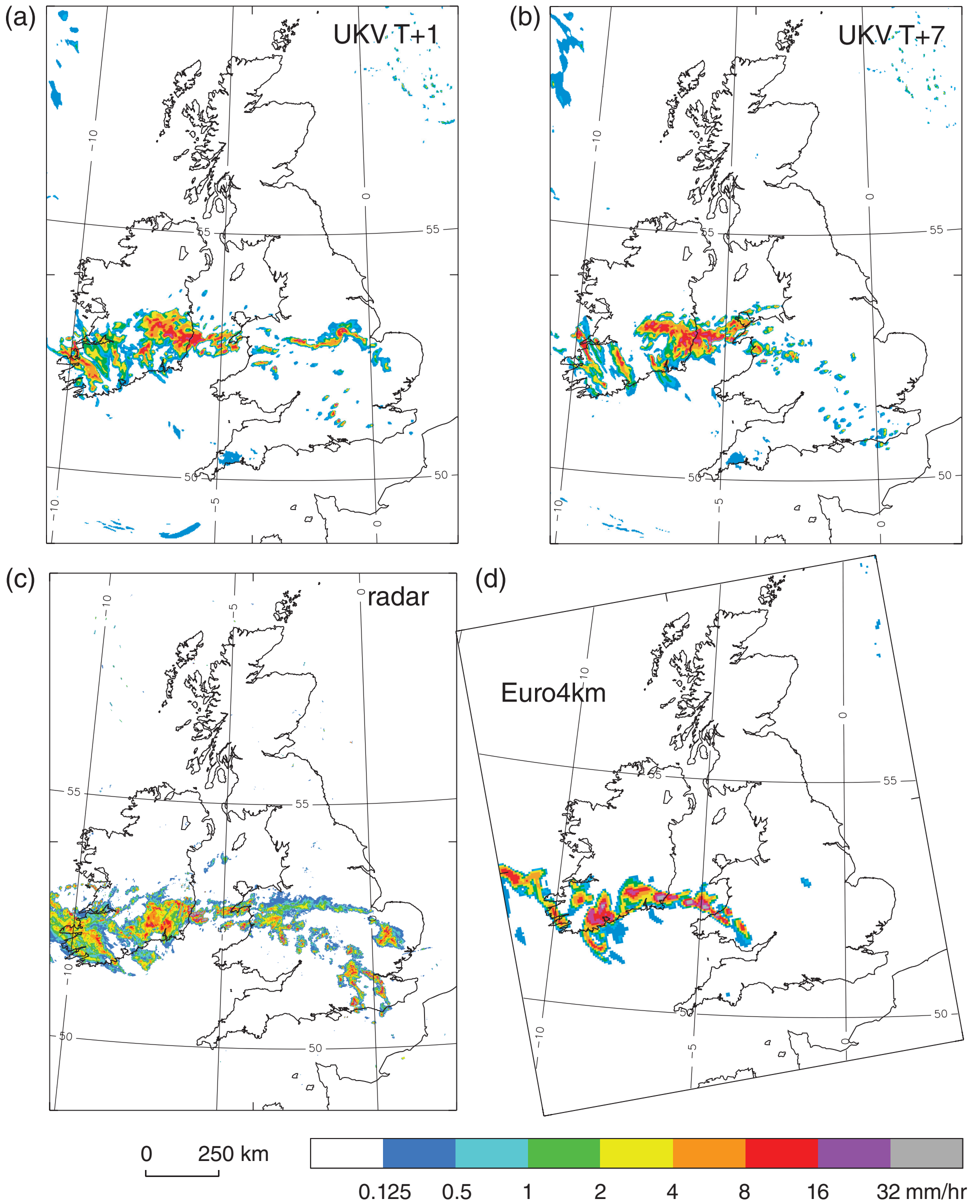


Figure 4. Forecasts for 0400 UTC 18 July 2014 compared to (c) quality controlled radar derived surface rain rate composite at 1 km resolution. Rainrate forecasts (mm h$^{-1}$) given by (a) T + 1 UKV from 0300 UTC, (b) T + 7 UKV from 2100 UTC, (d) T + 4 Euro4 4 km model down-scaled from global 25 km 0000 UTC analysis.

2012a, 2012b) and then with 4D-Var (Golding *et al.*, 2014; Sun *et al.*, 2014; Ballard *et al.*, 2015). This system was extended to allow assimilation of radar Doppler radial winds (Simonin *et al.*, 2014), which were then introduced operationally in the UKV model in 2011. Work is underway to develop direct assimilation of radar reflectivity (Hawkness-Smith and Ballard, 2013) and ceilometer backscatter as well as radar refractivity (Nicol *et al.*, 2014).

The trial of the 4D-Var, hourly-cycling NDP (see Section 1) for the summer of 2012 showed the benefit of frequently updating forecasts using more recent observations within the NWP model in the prediction of convective storms causing flash flooding (Ballard *et al.,* 2015). It showed some significant objective improvements over the latest available forecasts from the UKV model (updated only every 6 h) and the operational advection nowcasting system called the Short Term Ensemble Prediction System (STEPS) nowcast (Bowler *et al.,* 2006; Seed *et al.*, 2013). Li *et al.* (2015) have shown the benefit of hourly-cycling 4D-Var over 3D-Var in the NDP for precipitation forecasts during June 2012, which saw a number of significant flooding events in the UK (Z. Li *et al.,* 2015; personal communication).

It was also clear from the NDP experiment that the STEPS nowcast, the NDP forecasts and the UKV forecasts all had problems forecasting light, scattered convection. Although STEPS would have good locations from the radar at the initial time it would lose them in the forecast. In contrast, they could be missing at the initial time in the NWP forecasts, but then develop in roughly the right locations a short time into the forecast, but at too large a scale with too few individual cells. Whilst this could indicate a need for improvements in the data assimilation, in this situation it is also a symptom of the limitations of the model resolution. The model cannot resolve this type of convection at 1.5 km resolution so that either higher resolution or improved model parametrizations are required.

Dow and Macpherson (2013) investigated the benefit of convective scale data assimilation and the impact of the various observation types used operationally. They compared the relative impacts of higher model resolution, continuous data assimilation and the use of extra observations in the convection-permitting systems for the UK (on top of those used in the global 6-hourly cycling 4D-Var system). UK4 and UKV configurations were run as down-scaled forecasts from the 25 km resolution global analyses and also with continuous 3D-Var 3-hourly data assimilation cycles. The UKV and UK4 continuous data assimilation were run with and without the extra observations used in the UK4 and UKV systems and not the global system: MOPS cloud fraction analysis; radar rain rate for latent heat nudging; Doppler radial winds; OpenRoad Roadside 2 m temperature and 2 m relative humidity at Highways Agency sites; and visibility.

In experiments with the UK4 model, it was seen that the impact of data assimilation on the first 12 h of forecast was almost twice as great as the impact on the full 36 h of forecast. Observation network data denial experiments were undertaken for the UK4 system. A modest benefit was seen in precipitation verification through the use of the LHN. It was also seen that the precipitation forecasts benefitted from all the separate observation types of surface, satellite, upper air, aircraft and radar to broadly the same magnitude. With the proviso that the UKV model was run with a later version of the system than the UK4 model, the UKV seemed to gain more benefit from the full higher resolution data assimilation than the UK4. The 6 h precipitation accumulation forecasts gained benefit from both the higher resolution data assimilation and the use of additional observations. The benefit of full data assimilation compared with continuous cycling without data assimilation and also with down-scaling was investigated in the UKV model. For precipitation, full data assimilation provided the most benefit. Again the benefit measured over just the first 12 h forecast period was much greater than when it was measured over the full 36 h forecast period. It was also found that down-scaling (i.e., using the global analysis as initial conditions) provided better forecasts than continuous cycling without data assimilation.

The real challenges for convection-permitting data assimilation are just beginning. The Met Office operational system still uses the same balance constraints and control variables as in the global data assimilation systems. Additional or different control variables are likely to be needed to get the greatest benefit from observations of hydrometeors (Vetra-Carvalho *et al.*, 2012). The observation errors are assumed to be uncorrelated. These latter two issues are being investigated in the Natural Environment Research Council (NERC) Flooding from Intense Rainfall programme which aims to improve the accuracy of the initial conditions and hence forecasts of intense rainfall. This programme is also investigating the use of radar refractivity in data assimilation and developing improvements to the quality of radar reflectivity data. There is a need for capability to analyse the small-scale details, vital in storm initiation, which may be in conflict with the synoptic-scale flow forced through the lateral boundaries of the model. Climatological background errors are still used rather than flow-dependent, synoptically varying background errors. Research is ongoing at the Met Office to try to introduce synoptically varying, background errors through the use of ensemble NWP and data assimilation.

## 5. Verification

Although the introduction of CPMs has provided a step-change in the realism of convective rainfall in NWP models, it does not necessarily follow that more realistic-looking forecasts lead to greater forecast skill. It is possible that a forecast rainfall field, which may be almost indistinguishable from radar in visual impression, can still have the rain in the wrong place; therefore, according to a traditional grid point comparison metric a realistic-looking forecast might be assessed to have low skill. Lorenz (1969) argued that the ability to resolve smaller scales would result in forecast errors growing more rapidly and if this is the case it should be expected that CPM forecast skill would degrade more quickly. Several studies using traditional verification metrics have found that higher resolution does not deliver more skill (Done *et al.,* 2004; Weisman *et al.*, 2008; Mittermaier *et al.*, 2013), because the benefit of using a high-resolution model is counteracted by the unpredictable nature of the smaller scales that the model is designed to represent.

In other words, it is difficult to get a convective storm in exactly the right place at the right time and it becomes more difficult with longer-range forecasts. This raises the problem for objective verification that if the rain is even slightly in the wrong place, a forecast will score poorly, even if visually it looks very good and, in practise, is proving to give useful guidance to forecasters. A forecaster can make allowances for possible timing or spatial errors where a point-by-point score cannot. This inconsistency between subjective impression and the traditional objective scores limits the possibility of a meaningful objective evaluation of the performance of CPMs. For this reason, the whole philosophy of precipitation forecast verification has had to change to adapt to the introduction of these models.

Instead of the traditional approach to verification in which a variety of measures are used to compare forecast values at points (or pixels) against observations, newer methods have been developed that relax the condition that the rain has to be exactly correct in order to give a good score. Numerous verification methods have now been developed and many have been documented in Ebert (2008, 2009) and Gilleland *et al.* (2010). Examples are the wavelet approach of Casati *et al.* (2004), or the ‘fuzzy’ approach using the Fractions Skill Score (FSS) (Roberts, 2008; Roberts and Lean, 2008), or an image morphing technique (Keil and Craig, 2009). Another approach is to view the rainfall field as objects, such as the individual showers, and compare the characteristics of those objects with those observed. Two examples are MODE (Davis *et al.,* 2009) and SAL (Wernli *et al.*, 2008; Hanley *et al.,* 2015). Most of these methods require an observed rainfall field with good spatial coverage and therefore use radar-derived rainfall as ‘truth’. Results from these verification approaches have indeed shown the benefit of convection-permitting resolution both in terms of spatial accuracy and the characteristics of the rainfall (Lean *et al.*, 2008; Roberts*,* 2008; Schwartz *et al.*, 2009; Clark *et al.*, 2010; Mittermaier *et al.*, 2013).

Examples of verification results are shown in Figure 5, adapted from Roberts and Lean (2008), Figure 5(a) shows 1 km forecasts having higher skill than 12 km forecasts for convective events. The impact is greatest at scales larger than 10–20 km and for the higher rainfall totals. The intercept of the lines with the line FSS = 0.5 shows a reduction in the spatial error of the forecasts from around 70 km for the 12 km model to around 40 km (width of square) for the 1 km model for the 4 mm total in 4 h threshold, but less impact for the lighter 1 mm threshold. Figure 5(b) (taken from Mittermaier *et al.,* 2013) shows that the scale (neighbourhood size) at which a 4 km model has a useful level of skill is around 10 km less than a 12 km model taken over a 2 year period.

Dow and Macpherson (2013) report work by Mittermaier (personal communication) which confirms that over the past 3 years,

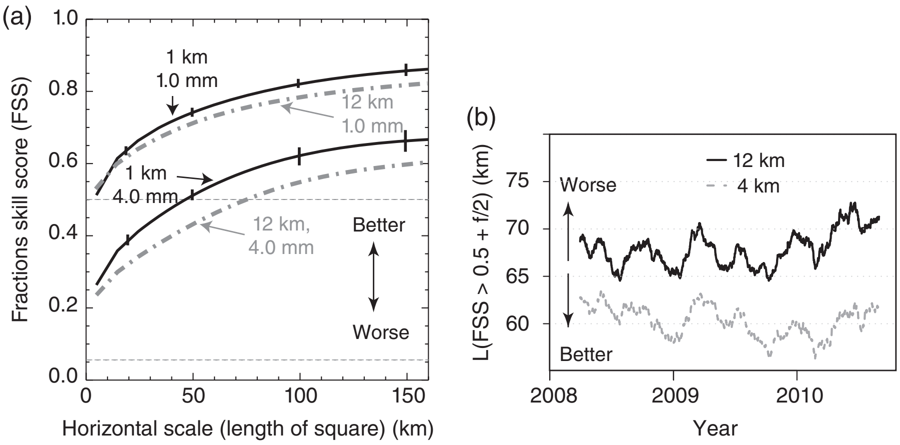


Figure 5. (a) Graph of the spatial skill of 1 km and 12 km forecasts of 4 h rainfall accumulations for a set of convective events (Roberts and Lean, 2008). The black lines are for the 1 km forecasts and the dashed for the 12 km. The upper curves are for a 1 mm threshold and the lower curves for a 4 mm threshold. (b) Graph of the neighbourhood size at which 4 km and 12 km forecasts have useful skill for an accumulation threshold of 4 mm in 6 h (Mittermaier *et al.,* 2013). (Copyright © 2011 British Crown Copyright, the Met Office.)

the UKV forecast system gives benefit over the global forecast system for prediction of surface weather. This was despite the fact that the UKV forecasts were 3 h older than the comparable global forecasts, and driven by 6 h older global lateral boundary conditions, reflecting the operational availability of forecasts over that period.

## 6. Predictability and ensembles

CPMs give more spatially accurate forecasts on average than coarser-resolution models. However, the scales at which CPM forecasts have good skill, on average, remain stubbornly large compared with the grid spacing of the models and the typical convective rain extent. Mittermaier *et al.* (2013) and Schwartz *et al.* (2009) used the FSS to show that reasonable forecast skill is achieved at scales of several tens to over a hundred kilometres (depending on forecast length) for models with grid spacings of 4 and 2 km. Clark *et al.* (2010) used a neighbourhood equitable threat score (ETS) with similar findings. Of course, forecasts can be very accurate on some occasions, especially for more predictable situations such as orographically-enhanced rainfall (Roberts 2008), but in general these results raise serious doubts about whether CPMs should be interpreted in a deterministic way, especially for point locations. Instead, there is a need to take a more probabilistic approach and consider the spatial uncertainty in the forecast.

The standard way to generate probabilistic forecasts is to run an ‘ensemble’ of several forecasts and produce probabilities based on how much those different forecasts agree. This approach, pioneered at ECMWF (Molteni *et al.*, 1996), has been used successfully for years at coarser resolutions, but until very recently was not possible for CPMs because of the computational cost. As an alternative, Roberts (2003b) and Theis *et al.* (2005) identified the use of ‘neighbourhood processing’ as a means of generating probabilities from single forecasts. This is really no different from the ‘fuzzy’ approach used for verification. It involves searching for alternative forecasts within the vicinity (or neighbourhood) of the pixel of interest with the assumption that each of those alternatives is equally likely. Probabilities are produced by taking the fraction of occurrences exceeding a particular threshold within the neighbourhood of each pixel. Alternatively, a Gaussian smoother can be used as in Sobash *et al.* (2011). These methods give smooth probability fields and have been shown to be more skilful than using the raw output (Theis *et al.*, 2005; Roberts and Lean, 2008; Ben Bouallègue and Theis, 2014).

In the last few years, increased supercomputer capability has allowed National Meteorological Services to start running CPM ensembles (Gebhardt *et al.*, 2008; Clark *et al.*, 2011; Bouttier *et al.*, 2012; Duc *et al.*, 2013; Golding *et al.*, 2014). Ensemble members are obtained by either varying the initial conditions to account for the uncertainties at the beginning of the forecast or by varying the representation of physical processes, systematically or randomly, to take account of the uncertainties associated with those processes. The objective is to produce an ensemble of forecasts that represents the true forecast uncertainty.

The methodology for constructing CPM ensembles is still very much a new area of research. Several approaches have been developed in recent years that include embedding the CPM within a coarser-resolution ensemble (Gebhardt *et al.*, 2008; Golding *et al.*, 2014), perturbing the CPM initial conditions by using an ensemble transform Kalman filter (Caron*,* 2013) and perturbing parameters in physics schemes randomly (Bouttier *et al.*, 2012), systematically (Gebhardt *et al.*, 2008, 2011; Leoncini *et al.*, 2013) or by using alternative schemes (Schwartz *et al.,* 2010). Studies have suggested that it is the information that comes into the domain from the coarser-resolution driving model or ensemble that has the biggest influence after around 6–12 h with physics perturbations or initial perturbations on the CPM grid being more influential before that (Gebhardt *et al.*, 2011; Kühnlein *et al.*, 2014). This means that the behaviour of the coarser-resolution ensemble (in which the CPM ensemble is embedded) is also crucial for the behaviour of the CPM ensemble.

An example of the benefit of a convection-permitting ensemble is given by Figures 6 and 7. Figure 6 shows a deterministic rainfall forecast (from the Met Office UKV model) that gave a good indication of the main areas of rain and showers, but there are location errors. In particular, a mesoscale area of rain in the south of England enclosed by the red dashed line was largely misplaced too far southeast in the forecast. In Figure 7, an ensemble of alternative forecasts for the same time contains several members with more rain in the dashed region than the deterministic forecast and therefore provides a better indication that rain is a possibility in that area.

 

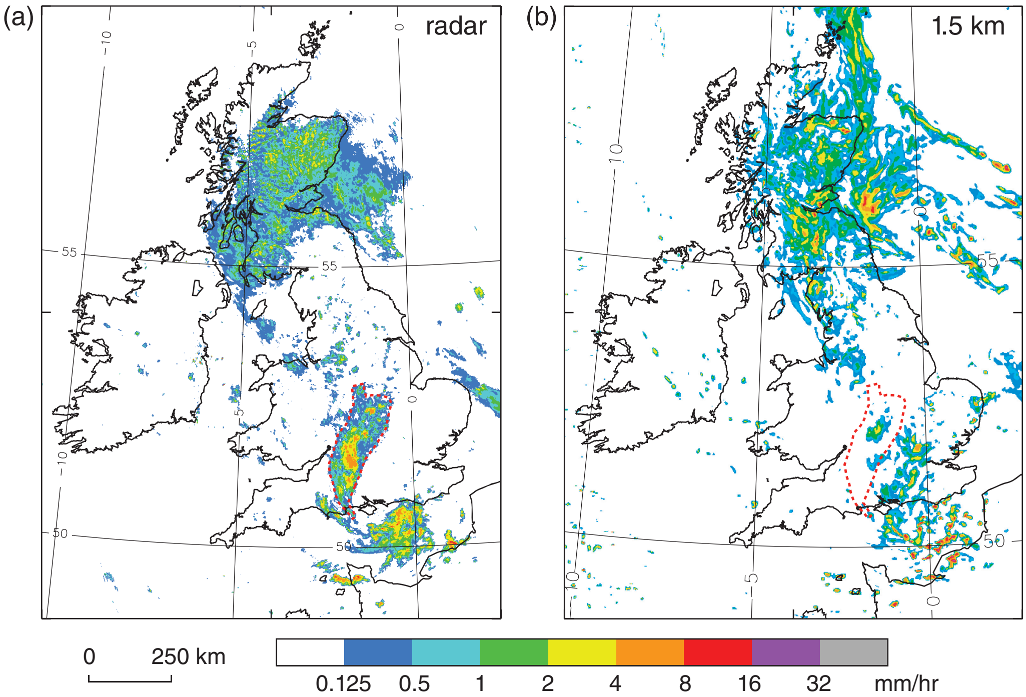


Figure 6. Mesoscale Convective System (MCS) event over the UK at 2100 UTC on 4 June 2014. Rainfall rates (mm h$^{-1}$) (a) as derived from the radar network and (b) as simulated by the UKV 1.5 km T + 18 forecast started at 0300 UTC on 4 June 2014. Red dashed line outlines the MCS on the radar picture and is reproduced on the forecast picture in order to illustrate that the MCS was not properly simulated in this particular deterministic forecast.

Although the neighbourhood method and the CPM ensemble have been treated as two alternative ways of producing probabilistic forecasts, they should really be used in conjunction. It has been shown that CPM ensembles using neighbourhood post-processing give more skilful probabilities (Schwartz *et al.,* 2010; Clark *et al.,* 2011; Duc *et al.,* 2013; Ben Bouallègue and Theis, 2014). Just as for deterministic forecasts, there will be an optimal neighbourhood size for any given ensemble forecast, which needs to be determined using spatial verification and this adds an extra complexity to both the verification and the post-processing of CPM ensembles. Upscaling methods are also being introduced (e.g. Ben Bouallègue and Theis, 2014) in which probabilities are presented for areas larger than the model grid square to provide a more appropriate and skilful product for severe rainfall events.

Another benefit of CPM ensembles is that they provide a useful scientific framework for improving the understanding of meteorological processes, predictability and relationships between the local weather and the larger-scale environment. Hanley *et al.* (2011, 2013) and Keil *et al.* (2014) showed how the predictability of convection can be related to aspects of the larger-scale environment. Barrett *et al.* (2015) also used a convection-permitting ensemble to gain understanding of the meteorological factors involved in the formation of a stationary rain band. The use of CPM ensembles should prove beneficial in investigating CPM model biases and for testing systematic effects of new model formulations.

## 7. Future work

Today CPMs are viable tools for simulating and forecasting rainfall at very high resolutions and meteorologists are beginning to rely upon CPMs to make forecasts on which critical decisions depend. Multiple paths for the continued improvement of today's CPMs could be considered. There are some obvious advances that can be made in the next few years as even more powerful, faster computers become available. For example, model resolution could be increased further, the model domain enlarged, ensemble sizes could be increased, more sophisticated DA could improve initial conditions, and longer or more frequent forecasts could be run. Each of these strategies for improvement comes at a cost and work is required to determine the relative benefits.

The design and study of very high resolution models has become an active area of research which has been given a boost by the trend towards running such models over areas of interest that have relatively small and computationally affordable domain sizes. Examples are the Weymouth 333 m model, which was run by the Met Office for the 2012 Olympics (Golding *et al.*, 2014), and the 250 m model over Toronto being developed for the Pan American Games (Leroyer *et al.,* 2014). Benefits from such models are not confined, of course, to their forecasts of convection.

Much convection in the UK is under-resolved by models at 1.5 km grid length (Bryan *et al.*, 2003; Hanley *et al.*, 2015). However, there is evidence that even at 100 m grid length there are issues with the representation of convection that are probably related to partially resolving turbulence (Hanley *et al.*, 2015). Thus, it appears that increasing the resolution should go hand in hand with further investigation and adaptation of physical parametrizations. Even given this assumption, it does not automatically follow that going to higher resolution will lead to significant forecast improvements for the cost (Schwartz *et al.*, 2009).

New data assimilation strategies are being developed in order to provide CPMs with the best possible initial conditions on the appropriate grid scale. The Met Office experiment with a 1.5 km resolution model and hourly-cycling 4D-Var (NDP) to

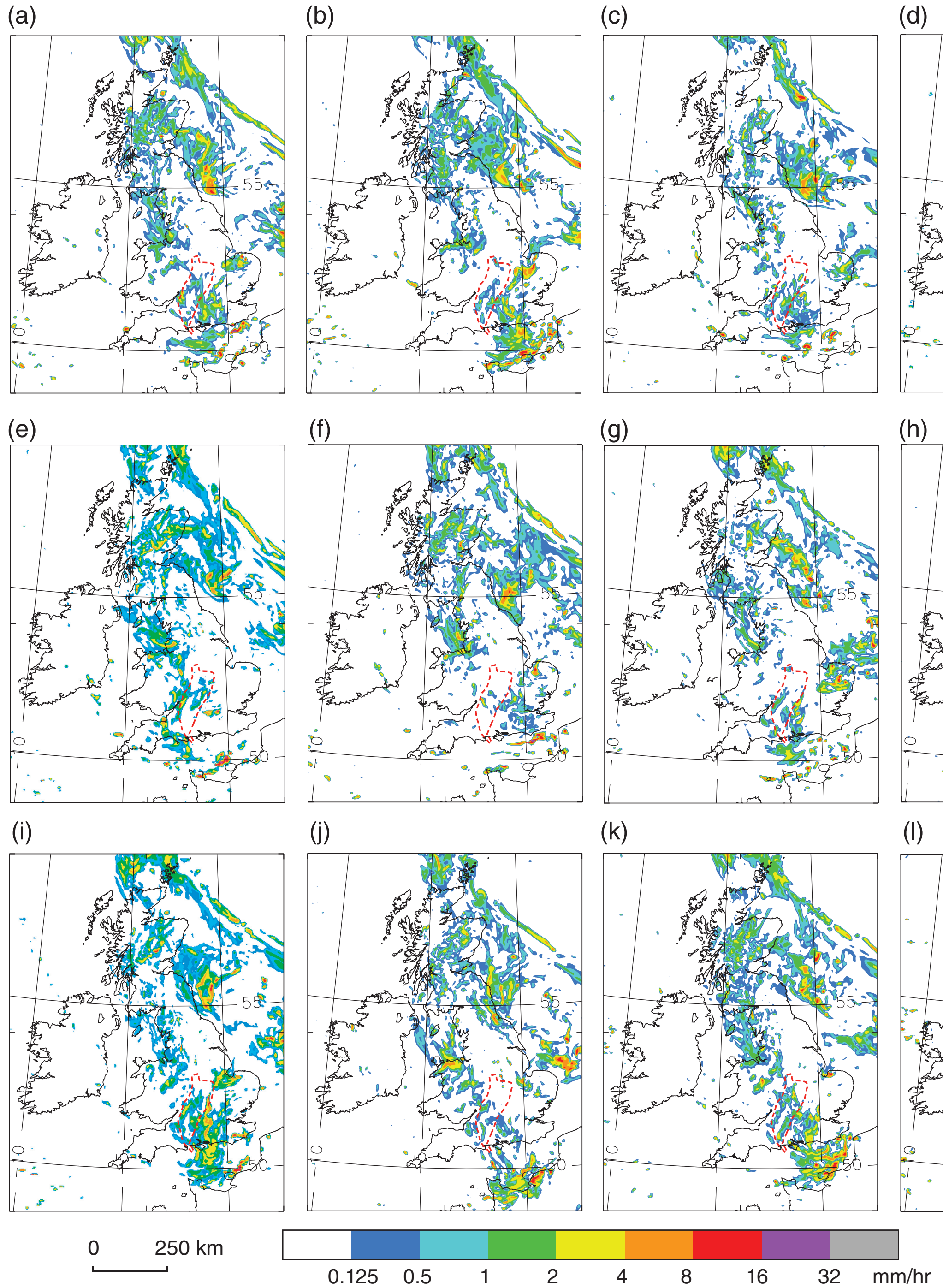


Figure 7. Rainfall rates (mm h$^{-1}$) given by MOGREPS UK ensemble members for 2100 UTC on 4 June 2014. The T + 18 forecast was started at 0300 UTC on 4 June 2014. Model resolution is 1.5 km. Red dashed line as in Figure 6.

produce 6 h forecasts (‘nowcasts’) every hour resulted in some deterministic forecasts of the location of systems producing flash flooding which would contribute to a major enhancement to the flood-warning system (Golding *et al.*, 2014; Sun *et al.,* 2014; Ballard *et al.*, 2015). Overall, the skill of predicting precipitation locations over 3 months from June to August 2012 exceeded that of the operational STEPS nowcast system from T + 2 onwards. As would be hoped, more frequent updating and fresh forecasts using the latest data, resulted in improved skill compared to the latest, but older, UKV forecasts produced only every 6 h. However, a few cases clearly needed improved data assimilation methods to ensure that both the 3-hourly cycling UKV and hourly-cycling NDP short-range forecasts were better than longer-range UKV forecasts.

Other forecasts suffered from errors in the synoptic-scale forcing through the boundary conditions, a particular problem for the very small NDP domain. This suggests a need for larger domain sizes and better observation coverage over the oceans and seas close to the UK as well as illustrating the dependency of high resolution limited area model forecasts on data assimilation in, and skill of, the forecast model providing the boundary conditions. This serves to illustrate the inter-connected nature of the forecast system, with improvements to one aspect having potential to change (and, hopefully, improve) the performance of other aspects.

The current aim in the Met Office is to implement hourly-cycling 4D-Var in the UKV model. One issue to be resolved is whether the same system can produce UK-wide forecasts up to T + 36 and frequently updated precipitation nowcasts. The 36 h forecasts need to produce good forecasts of temperatures, wind, visibility and low cloud as well as precipitation. This may require improved background errors and control variables to represent relationships between variables at the convective scale. Forecasts should also benefit from exploitation of ensemble or hybrid data assimilation methods at convection-permitting scale to enable representation of errors of the day or synoptically varying background errors. Research continues to exploit novel observations such as radar reflectivity and refractivity, ceilometer backscatter and future geostationary sounders in order to improve the initial conditions for the forecasts.

Beyond (and even within) the nowcasting forecast range, it is becoming accepted wisdom that because convection-permitting resolution does not lead to sufficient skill at the grid scale, it is not possible to rely only on deterministic forecasts and that a probabilistic approach to forecasting rainfall is needed. This has led to the development of CPM ensemble systems, which, at the moment, tend to follow the methodologies used at coarser resolution, and will need to be developed further.

The small ensembles that can be run now are insufficient to capture the uncertainty in rainfall forecasts and more ensemble members are required, especially for longer forecasts (Clark *et al.*, 2011). The difficulty will be in developing perturbation strategies that provide a set of ensemble members that truly represent the forecast uncertainty, including the challenge of dealing with model error and initial condition uncertainty at convection-permitting resolutions. This is very much an active area of research (Baker *et al.,* 2014) and is likely to continue to be in the future. Stensrud *et al.* (2009) presented a vision for a 'warn-on' forecast system that uses ensemble forecasts at the convective scale to forecast high-impact weather in an objective, probabilistic framework. In particular, the NWP-based nowcasting technique is discussed with emphasis on assimilation of 'in-storm' observations to track the ongoing convection. However, the problem of not having enough ensemble members for forecasting purposes is likely to remain. This means that the use of neighbourhood (or similar) post-processing methods will continue to be necessary. These post-processing methods will require further development and optimization in order to give well-calibrated probabilities.

Regional climate modelling researchers are moving towards using CPMs. Initial experiments took place using configurations of the MetUM based on those above to down-scale analyses (Wang *et al.*, 2013) and climate projections (Kendon *et al.*, 2012, 2014). It is likely that CPMs will become popular tools in the future of regional climate modelling research. This will provide valuable information about model performance in a wide range of conditions, information that is difficult to obtain in a purely NWP context.

The discussion above has focussed on the development of the modelling system, and highlights that there is much that can be done that could push forward the usefulness of the CPM as a tool to forecast rainfall both qualitatively and quantitatively at high resolution. However, there is little doubt that the quality of quantitative precipitation forecasts will be limited by an understanding of the key physical processes in convection, including boundary-layer turbulence, turbulent entrainment, cloud microphysics and its interaction with aerosols, downdraught generation and cold pool dynamics, combined with the ability of the models to represent these processes accurately (Fritsch and Carbone, 2004). There is still much fundamental research required, including detailed research observations of these processes.

Ultimately, the aim is to provide probabilistic forecasts for both 'ordinary' and 'extreme' precipitation events at the finest possible scientifically credible resolution which are useful for decision-making purposes. To do this well will require investment in the development of both generic and customer-specific products that extract the scientifically acceptable information from the CPM ensemble outputs without losing sight of customer needs. In the future, CPMs are expected to be included in a more holistic forecasting approach that fully couples with ocean and hydrological models to provide a better physical representation of the complete hydrological system. This challenging plan must be accompanied by an open dialogue between the scientists who are developing the modelling capability and the scientists or customers who are using the output in order to maintain clear expectations about the high-resolution rainfall products.

## 8. Summary

Convection-permitting models (CPMs) have provided operational weather forecasting centres with a step-change in their capabilities to forecast rainfall. The increasing power of computers and the development of new, non-hydrostatic models with more detailed representation of physical processes have given National Meteorological Services the ability to run regional numerical weather prediction (NWP) models at resolutions high enough to eliminate the need for convective parameterizations, providing hitherto unavailable guidance on the nature of severe convective storms. CPM-based nowcasting systems are beginning to show the promise of short-range forecasts of rapidly developing intense systems, which were previously impossible. Some centres are even running ensembles of CPMs to provide forecasters with a measure of uncertainty in the forecast in order to be able to issue objective, quantitative probabilistic rainfall forecasts and warnings. The production of probabilistic warnings for local intense rainfall was simply not possible with coarser-resolution models and has provided a completely new capability in weather forecasting.

The present study has described the current state of the art in operational CPM-based NWP, primarily from a UK perspective, starting with the historical context that led to the development of practical forecasting systems in which the physics of atmospheric convection is treated much more completely and directly than was possible in models using a parametrization of convection. The resulting forecasts are much more realistic than those relying on parametrization of convection, but present both an opportunity and a challenge when compared with detailed observations such as radar data. Various data assimilation techniques used with operational CPMs have been reviewed, along with a historical context and with examples of the new types of observed data that may increase the skill of forecasts due to improved initial conditions. The development of CPMs has also brought with it the need to develop new techniques to verify forecasts and define their skill. Finally, the lack of predictability of the smallest scales involving convection means that ensemble techniques are required to assess forecast confidence. The current state of the art represents a beginning, not a conclusion, and it is anticipated that many advances in various directions will be possible in the future.

## References


Anthes RA, Warner TT. 1978. Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Mon. Weather Rev.* **106**: 1045–1078.

Arakawa A, Schubert WH. 1974. Interaction of a cumulus cloud ensemble with large-scale environment, part 1. *J. Atmos. Sci.* **31**: 674–701.

Arakawa A, Wu C-M. 2013. A unified representation of deep moist convection in numerical modelling of the atmosphere – part 1. *J. Atmos. Sci.* **70**(7): 1929–1953.

Baker L, Rudd A, Migliorini S, Bannister R. 2014. Representation of model error in a convective-scale ensemble prediction system. *Nonlinear Processes Geophys.* **21**: 19–39.

Baldauf M, Seifert A, Förstner J, Majewski D, Raschendorfer M, Reinhardt T. 2011. Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Mon. Weather Rev.* **139**: 3887–3905.

Ballard SP, Golding BW, Smith RNB. 1991. Mesoscale model experimental forecasts of the Haar of North East Scotland. *Mon. Weather Rev.* **119**: 2107–2123.

Ballard SP, Li Z, Simonin D, Buttery H, Charlton-Perez C, Gaussiat N, Hawkness-Smith L. 2012a. Use of radar data in NWP-based nowcasting in the Met Office. In *Proceedings of Weather Radar and Hydrology*, Vol. 352. 18–21 Exeter, UK; 336–341.

Ballard SP, Macpherson B, Li Z, Simonin D, Caron J-F, Buttery H, Charlton-Perez C, Gaussiat N, Hawkness-Smith L, Piccolo C, Kelly G, Tubbs R, Dow G, Renshaw R. 2012b. Convective scale data assimilation and nowcasting. In *Proceedings of ECMWF Seminar on Data Assimilation for Atmosphere and Ocean,* 6–9 September 2011, Shinfield Park, Reading.

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*. DOI: 10.1002/qj.2665.

Barrett AI, Gray SL, Kirshbaum DJ, Roberts NM, Schultz DM, Fairman JG. 2015. Synoptic versus orographic control on stationary convective banding. *Q. J. R. Meteorol. Soc.* **141**: 1101–1113.

Ben Bouallègue Z, Theis SE. 2014. Spatial techniques applied to precipitation ensemble forecasts: from verification results to probabilistic products. *Meteorol. Appl.* **21**: 922–929.

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

Blyth AM, Bennett LJ, Collier CG. 2015. High-resolution observations of precipitation from cumulonimbus clouds. *Meteorol. Appl.* **22**: 75–89.

Bouttier F, Vie B, Nuissier O, Raynaud L. 2012. Impact of stochastic physics in a convection-permitting ensemble. *Mon. Weather Rev.* **140**: 3706–3721.

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

Brousseau P, Berre L, Bouttier F, Desroziers G. 2011. Background-error covariances for a convective-scale data-assimilation system: AROME-France 3D-Var. *Q. J. R. Meteorol. Soc.* **137**: 409–422.

Bryan GH, Wyngaard JC, Fritsch JM. 2003. Resolution requirements for the simulation of deep moist convection. *Mon. Weather Rev.* **131**: 2394–2416.

Caron J-F. 2013. Mismatching perturbations at the lateral boundaries in limited-area ensemble forecasting: a case study. *Mon. Weather Rev.* **141**: 356–374.

Casati B, Ross G, Stephenson DB. 2004. A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl.* **11**: 141–154.

Clark PA, Browning KA, Forbes RM, Morcrette CJ, Blyth AM, Lean HW. 2014. The evolution of an MCS over southern England. Part 2: model simulations and sensitivity to microphysics. *Q. J. R. Meteorol. Soc.* **140**: 458–479.

Clark AJ, Gallus WA Jr, Weisman ML. 2010. Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF model simulations and the operational NAM. *Weather Forecast.* **25**: 1495–1509.

Clark AJ, Kain JS, Stenstrud DJ, Xue M, Kong F, Coniglio MC, et al. 2011. Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble. *Mon. Weather Rev.* **139**: 1052–1081.

Crook NA. 2001. Understanding Hector: the dynamics of island thunderstorms. *Mon. Weather Rev.* **129**: 1550–1563.

Cullen MJP. 1993. The unified forecast/climate model. *Meteorol. Mag.* **122**: 81–94.

Cullen MJP, Davies T. 1991. A conservative split-explicit integration scheme with fourth-order horizontal advection. *Q. J. R. Meteorol. Soc.* **117**: 993–1002.

Davies T, Cullen MJP, Malcolm AJ, Mawson MH, Staniforth A, White AA, et al. 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.

Davis CA, Brown BG, Bullock R, Halley-Gotway J. 2009. The method for object-based diagnostic evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring program. *Weather Forecast.* **24**: 1252–1267.

de Haan S. 2011. High-resolution wind and temperature observations from aircraft tracked by Mode-S air traffic control radar. *J. Geophys. Res.* **116**: D10111, DOI:10.1029/2010JD015264.

de Haan S, Stoffelen A. 2012. Assimilation of high-resolution mode-S wind and temperature observations in a regional NWP model. *Weather Forecast.* **27**: 918–937.

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.

Done J, Davis CA, Weisman ML. 2004. The next generation of NWP: explicit forecasts of convectioning the weather research and forecasting (WRF) model. *Atmos.Sci. Lett.* **5**: 110–117.

Dow G, Macpherson B. 2013. Benefit of convective-scale data assimilation and observing systems in the UK Models. Forecasting Research Technical Report no 585. Met Office: Exeter, UK.

Droegemeier KK. 1990. Toward a science of storm scale prediction. Preprint, *16th Conference on Severe Local Storms.* American Meteorological Society: Kananaskis Park, Alberta, Canada.

Duc L, Saito K, Seko H. 2013. Spatial-temporal fractions verification for high-resolution ensemble forecasts. *Tellus A* **65**: 18171–18193.

Ebert EE. 2008. Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorol. Appl.* **15**: 51–64.

Ebert EE. 2009. Neighbourhood verification: a strategy for rewarding close forecasts. *Weather Forecast.* **24**: 1498–1510.

Essery RLH, Best MJ, Betts RA, Cox PM, Taylor CM. 2003. Explicit representation of subgrid heterogeneity in a GCM land surface scheme. *J. Hydrometeorol.* **4**: 530–543.

Forbes R, Halliwell C. 2003. Assessment of the performance of an enhanced microphysics parametrization scheme in the Unified Model at 1 km resolution. Met Office internal report. Met Office: Exeter, UK.

Fritsch JM, Carbone RE. 2004. Improving quantitative precipitation forecasts in the warm season: a USRWP research and development strategy. *Bull. Am. Meteorol. Soc.* **85**: 955–965.

Fritsch JM, Chappell CF. 1980. Numerical prediction of convectively driven mesoscale pressure systems. Part I: convective parametrization. *J. Atmos. Sci.* **37**: 1722–1733.

Gebhardt C, Theis S, Krahe P, Renner V. 2008. Experimental ensemble forecasts of precipitation based on a convection-resolving model. *Atmos. Sci. Lett.* **9**: 67–72.

Gebhardt C, Theis SE, Paulat ME, Ben Bouallègue Z. 2011. Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmos. Res.* **100**: 168–177.

Gerard L, Piriou J-M, Brožková R, Geleyn J-F, Banciu D. 2009. Cloud and precipitation parameterization in a meso-gamma-scale operational weather prediction model. *Mon. Weather Rev.* **137**(11): 3960–3977.

Gilleland E, Ahijevych DA, Brown BG, Ebert EE. 2010. Verifying forecasts spatially. *Bull. Am. Meteorol. Soc.* **91**: 1365–1373.

Golding BW. 1990. The Meteorological Office mesoscale model. *Meteorol. Mag.* **119**: 81–96.

Golding BW. 1992. An efficient non-hydrostatic forecast model. *Meteorol. Atmos. Phys.* **50**: 89–103.

Golding BW. 1993. A numerical investigation of tropical island thunderstorms. *Mon. Weather Rev.* **121**: 1417–1433.

Golding BW. 1998. Nimrod: a system for generating automated very short range forecasts. *Meteorol. Appl.* **5**: 1–16.

Golding BW, Ballard SP, Mylne K, Roberts N, Saulter A, Wilson C, Agnew P, Davis LS, Trice J, Jones C, Simonin D, Li Z, Pierce C, Bennett A, Weeks M, Moseley S. 2014. Forecasting Capabilities for the London 2012 Olympics. *Bull. Amer. Meteor. Soc.* **95**: 883–896.

Gregory D, Rowntree PR. 1990. A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Weather Rev.* **118**: 1483–1506.

Hanley KE, Kirshbaum DJ, Belcher SE, Roberts NM, Leoncini G. 2011. Ensemble predictability of an isolated mountain thunderstorm in a high-resolution model. *Q. J. R. Meteorol. Soc.* **137**: 2124–2137.

Hanley KE, Kirshbaum DJ, Roberts NM, Leoncini G. 2013. Sensitivities of a squall line over Central Europe in a convective-scale ensemble. *Mon. Weather Rev.* **141**: 112–133.
Hanley KE, Plant RS, Stein TH M, Hogan RJ, Nicol JC, Lean HW, Halliwell C, Clark P A. 2015. Mixing-length controls on high-resolution simulations of convective storms. *Q. J. R. Meteorol. Soc.* **141**: 272–284.
Hawkness-Smith L, Ballard SP. 2013. Assimilation of radar reflectivity data in the Met Office convective-scale forecast system. In Proceedings of 36th Conference on Radar Meteorology, 19 September 2013, Breckenridge, CO, USA. American Meteorological Society: Boston, MA.
Honda Y, Nishijima M, Koizumi K, Ohta Y, Tamiya K, Kawabata T, et al. 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.
Jones CD, Macpherson B. 1997. A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. *Meteorol. Appl.* **4**: 269–277.
Kain JS, Fritsch JM. 1993. Convective parameterization for mesoscale models: the Kain-Fritsch scheme. The representation of cumulus convection in numerical models. *Meteorol. Monogr.* **24**: 165–170.
Kain JS, Weiss SJ, Bright DR, Baldwin ME, Levit JJ, Carbin GW, et al. 2008. Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Weather Forecast.* **23**: 931–952.
Keil C, Craig GC. 2009. A displacement and amplitude score employing an optical flow technique. *Weather Forecast.* **24**: 1297–1308.
Keil C, Heinlein F, Craig GC. 2014. The convective adjustment time-scale as indicator of predictability of convective precipitation. *Q. J. R. Meteorol. Soc.* **140**: 480–490.
Kendon EJ, Roberts NM, Fowler HJ, Roberts MJ, Chan S, Senior CA. 2014. Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nat. Clim. Change* **4**: 570–576.
Kendon EJ, Roberts NM, Senior CA, Roberts MJ. 2012. Realism of rainfall in a very high-resolution regional climate model. *J. Clim.* **25**: 5791–5806.
Khairoutdinov M, Randall D. 2006. High-resolution simulation of shallow-to-deep convection transition over land. *J. Atmos. Sci.* **63**: 3421–3436.
Klemp JB, Wilhelmson RB. 1978. The simulation of three-dimensional convective storm dynamics. *J. Atmos. Sci.* **35**: 1070–1096.
Kühnlein C, Keil C, Craig GC, Gebhardt C. 2014. The impact of downscaled initial condition perturbations on convective-scale ensemble forecasts of precipitation. *Q. J. R. Meteorol. Soc.* **140**: 1552–1562.
Lean HW, Clark PA, Dixon M, Roberts NM, Fitch A, Forbes R, et al. 2008. Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Weather Rev.* **136**: 3408–3424.
Leoncini G, Plant RS, Gray SL, Clark PA. 2013. Ensemble forecasts of a flood-producing storm: comparison of the influence of model-state perturbations and parameter modifications. *Q. J. R. Meteorol. Soc.* **139**: 198–211.
Leroyer S, Dorval EC, Canada QC, Bélair S, Husain Z, Vionnet V. 2014. Urban, water bodies, and orography induced circulations and their impact on extreme weather forecasting. In *Proceedings of 11th Symposium on the Urban Environment*, 2–6 February 2014, Atlanta, GA, USA. American Meteorological Society: Boston, MA.
Li Z, Ballard SP, Simonin D. 2015. 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. Met Office: Exeter, UK
Lilly DK. 1990. Numerical prediction of thunderstorms – has its time come? *Q. J. R. Meteorol. Soc.* **116**: 779–798.
Lorenc AC, Ballard SP, Bell RS, Ingleby NB, Andrews PLF, Barker DM, et al. 2000. The Met. Office global 3-Dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.* **12**: 2991–3012.
Lorenc AC, Bell RS, Macpherson B. 1991. The Meteorological Office analysis corrective data assimilation scheme. *Q. J. R. Meteorol. Soc.* **117**: 59–89.
Lorenz EN. 1969. The predictability of a flow which possesses many scales of motion. *Tellus* **21**: 289–307.
Macpherson B, Wright BJ, Hand WH, Maycock AJ. 1996. The impact of MOPS moisture data in the U.K. Meteorological Office mesoscale data assimilation scheme. *Mon. Weather Rev.* **124**: 1746–1766.
Manners J, Vosper SB, Roberts N. 2012. Radiative transfer over resolved topographic features for high-resolution weather prediction. *Q. J. R. Meteorol. Soc.* **138**: 720–733.
Michalakes J, Chen S, Dudhia J, Hart L, Klemp J, Middlecoff J, et al. 2001. "Development of a next generation regional weather research and forecast model" in developments in teracomputing. In *Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology*. World Scientific: Singapore.
Miller MJ, Pearce RP. 1974. A three-dimensional primitive equation model of cumulonimbus convection. *Q. J. R. Meteorol. Soc.* **100**: 133–154.
Mittermaier M, Roberts NM, Thompson SA. 2013. A long term assessment of precipitation forecast skill using the fractions skill score. *Meteorol. Appl.* **20**: 176–186.
Molteni F, Buizza R, Palmer TN, Petroliagis T. 1996. The ECMWF ensemble prediction system: methodology and validation. *Q. J. R. Meteorol. Soc.* **122**: 73–119.
Müller M, Scherer D. 2005. A grid- and subgrid-scale radiation parameterization of topographic effects for mesoscale weather forecast models. *Mon. Weather Rev.* **133**: 1431–1442.
Nicol JC, Illingworth AJ, Bartholomew K. 2014. The potential of 1 h refractivity changes from an operational C-band magnetron-based radar for numerical weather prediction validation and data assimilation. *Q. J. R. Meteorol. Soc.* **140**: 1209–1218.
Oliphant AJ, Spronken-Smith RA, Sturman AP. 2003. Spatial variability of surface radiation fluxes in mountainous terrain. *J. Appl. Meteorol.* **42**: 113–128.
Pielke RA. 1974. A three-dimensional numerical model of the sea breezes over south Florida. *Mon. Weather Rev.* **102**: 115–139.
Pielke RA, Cotton WR, Walko RL, Tremback CJ, Lyons WA, Grasso LD, et al. 1992. A comprehensive meteorological modeling system – RAMS. *Meteorol. Atmos. Phys.* **49**: 69–91.
Porson A, Clark PA, Harman IN, Best MJ, Belcher SE. 2010. Implementation of a new urban energy budget scheme in the MetUM. Part I: description and idealized simulations. *Q. J. R. Meteorol. Soc.* **136**: 1514–1529.
Rawlins FR, Ballard SP, Bovis KR, Clayton AM, Li D, Inverarity GW, et al. 2007. The Met Office global 4-Dimensional data assimilation system. *Q. J. R. Meteorol. Soc.* **133**: 347–362.
Renshaw R, Francis PN. 2011. Variational assimilation of cloud fraction in the operational Met Office Unified Model. *Q. J. R. Meteorol. Soc.* **137**: 1963–1974.
Richardson LF. 1922. *Weather Prediction by Numerical Process*. Cambridge University Press.
Roberts NM. 2003a. The impact of a change to the use of the convection scheme to high resolution simulations of convective events (stage 2 report from the storm scale numerical modelling project). Met Office Technical Report, 407; 31 (available online).
Roberts N M. 2003b. Precipitation diagnostics for a high resolution forecasting system. Met Office Technical Report, 423; 45 (available online).
Roberts NM. 2008. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorol. Appl.* **15**: 163–169.
Roberts NM, Cole SJ, Forbes RM, Moore RJ, Boswell D. 2009. Use of high-resolution NWP rainfall and river flow forecasts for advance warning of the Carlisle flood, north-west England. *Meteorol. Appl.* **16**: 23–34.
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, et al. 2006. The operational JMA nonhydrostatic mesoscale model. *Mon. Weather Rev.* **134**: 1266–1298.
Schlesinger RE. 1975. A three-dimensional numerical model of an isolated deep convective cloud: preliminary results. *J. Atmos. Sci.* **32**: 934–957.
Schwartz CS, Kain JS, Weiss SJ, Xue M, Bright DR, Kong F, et al. 2009. Next-day convection-allowing WRF model guidance: a second look at 2-km versus 4-km grid spacing. *Mon. Weather Rev.* **137**: 3351–3372.
Schwartz CS, Kain JS, Weiss SJ, Xue M, Bright DR, Kong F, et al. 2010. Toward improved convection-allowing ensembles: model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Weather Forecast.* **25**: 263–280.
Schwartz CS, Liu Z. 2014. Convection-permitting forecasts initialized with continuously cycling limited-area 3DVAR, ensemble Kalman filter, and "hybrid" variational-ensemble data assimilation systems. *Mon. Weather Rev.* **142**: 716–773.
Seed AW, Pierce CE, Norman K. 2013. Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resour. Res.* **49**: 6624–6641.

Seity Y, Brosseau P, Malardel S, Hello G, Bénard P, Bouttier F, et al. 2011. The AROME-France convective-scale operational model. *Mon. Weather Rev.* **139**: 976–991.
Simonin D, Ballard SP, Li Z. 2014. Doppler radar radial wind assimilation using an hourly cycling 3D-Var with a 1.5 km resolution version of the Met Office Unified Model for nowcasting. *Q. J. R. Meteorol. Soc.* **140**: 2298–2314.
Skamarock WC. 2004. Evaluating mesoscale NWP models using kinetic energy spectra. *Mon. Weather Rev.* **132**: 3019–3032.
Sobash RA, Kain JS, Bright DR, Dean AR, Coniglio MC, Weiss SJ. 2011. Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Weather Forecast.* **26**: 714–728.
Stensrud DJ, Wicker LJ, Kelleher KE, Xue M, Foster MP, Schaefer JT, et al. 2009. Convective-scale warn-on-forecast system. *Bull. Am. Meteorol. Soc.* **90**: 1487–1499.
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 J, Zawadzki I, Ballard SP, Onvlee-Hooimeyer J, et al. 2014. Use of NWP for nowcasting precipitation: recent progress and challenges. *Bull. Am. Meteorol. Soc.* **95**: 409–426.
Tang Y, Lean H, Bornemann J. 2013. The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorol. Appl.* **20**: 417–426.
Tapp MC, White PW. 1976. A non-hydrostatic mesoscale model. *Q. J. R. Meteorol. Soc.* **102**: 277–296.
Theis SE, Hense H, Damrath U. 2005. Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. *Meteorol. Appl.* **12**: 257–268.
Tremback CJ, Tripoli GJ, Cotton WR. 1985. A regional scale atmospheric numerical model including explicit moist physics and a hydrostatic time-split scheme. In *Proceedings of 7th AMS Conference on Numerical Weather Prediction*, June 17–20, Montreal, Quebec, Canada. American Meteorological Society: Boston, MA.
Tripoli GJ, Cotton WR. 1980. A numerical investigation of several factors contributing to the observed variable intensity of deep convection over south Florida. *J. Appl. Meteorol.* **19**: 1037–1063.
Tripoli GJ, Cotton WR. 1982. The Colorado State University three-dimensional cloud/mesoscale model – 1982. Part I: general theoretical framework and sensitivity experiments. *J. de Rech. Atmos.* **16**: 185–220.
Vetra-Carvalho S, Dixon M, Migliorini S, Nichols NK, Ballard SP. 2012. Breakdown of hydrostatic balance at convective scales in the forecast errors in the Met Office Unified Model. *Q. J. R. Meteorol. Soc.* **138**: 1709–1720.
Wang C, Jones R, Perry M, Johnson C, Clark P. 2013. Using an ultrahigh-resolution regional climate model to predict local climatology. *Q. J. R. Meteorol. Soc.* **139**: 1964–1976.
Weisman ML, Davis C, Wang W, Manning KW, Klemp JB. 2008. Experiences with 0-36 hour explicit convective forecasts with the WRF-ARW model. *Weather Forecast.* **23**: 407–437.
Wernli H, Paulat M, Hagen M, Frei C. 2008. SAL – a novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Weather Rev.* **136**: 4470–4487.
White AA, Hoskins BJ, Roulstone I, Staniforth A. 2005. Consistent approximate models of the global atmosphere: shallow, deep, hydrostatic, quasi hydrostatic and non-hydrostatic. *Q. J. R. Meteorol. Soc.* **131**: 2081–2107.
Wilkinson JM, Bornemann FJ. 2014. A lightning forecast for the London 2012 Olympics opening ceremony. *Weather* **69**: 16–19.
Wilkinson J, Wilson D, Forbes R. 2011. The large-scale precipitation parametrization scheme. UM documentation paper 26, Model version 8.1. Met Office: Exeter, UK.
Wilson DR, Ballard SP. 1999. A microphysically based precipitation scheme for the UK Meteorological Office unified model. *Q. J. R. Meteorol. Soc.* **125**: 1607–1636.
Wilson D, Forbes R. 2004. The large-scale precipitation parametrization scheme. UM documentation paper 26, Model version 5.5. Met Office: Exeter, UK.
Wright BJ, Golding BW. 1990. The interactive mesoscale initialization. *Meteorol. Mag.* **119**: 234–244.
Xue M, Wang D-H, Gao J-D, Brewster K, Droegemeier KK. 2003. The advanced regional prediction system (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteorol. Atmos. Phys.* **82**: 139–170.
Yu X, Lee T-Y. 2010. Role of convective parameterization in simulations of a convection band at grey-zone resolutions. *Tellus A* **62**: 617–632.