METEOROLOGICAL APPLICATIONS
Meteorol. Appl. **22**: 90–104 (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/met.1493

From months to minutes – exploring the value of high-resolution rainfall observation and prediction during the UK winter storms of 2013/2014

Huw Lewis,* Marion Mittermaier, Ken Mylne, Katie Norman, Adam Scaife, Robert Neal, Clive Pierce, Dawn Harrison, Sharon Jewell, Michael Kendon, Roger Saunders, Gilbert Brunet, Brian Golding, Malcolm Kitchen, Paul Davies and Charles Pilling

Met Office, Exeter, UK

ABSTRACT: As the societal impacts of hazardous weather and other environmental pressures grow, the need for integrated predictions that can represent the numerous feedbacks and linkages between sub-systems is greater than ever. This was well illustrated during winter 2013/2014 when a prolonged series of deep Atlantic depressions over a 3 month period resulted in damaging wind storms and exceptional rainfall accumulations. The impact on livelihoods and property from the resulting coastal surge and river and surface flooding was substantial. This study reviews the observational and modelling toolkit available to operational meteorologists during this period, which focusses on precipitation forecasting months, weeks, days and hours ahead of time. The routine availability of high-resolution (km scale) deterministic and ensemble rainfall predictions for short-range weather forecasting as well as weather-resolving seasonal prediction capability represent notable landmarks that have resulted from significant progress in research and development over the past decade. Latest results demonstrated that the suite of global and high-resolution UK numerical weather prediction models provided excellent guidance during this period, supported by high-resolution observations networks, such as weather radar, which proved resilient in difficult conditions. The specific challenges for demonstrating this performance for high-resolution precipitation forecasts are discussed. Despite their good operational performance, there remains a need to further develop the capability and skill of these tools to fully meet user needs and to increase the value that they deliver. These challenges are discussed, notably to accelerate the progress towards understanding the value that might be delivered through more integrated environmental prediction.

KEY WORDS hydro-meteorology; forecasting; hazards; seamless prediction; verification; environmental prediction; high impact weather; quantitative precipitation forecasting; user needs

Received 11 July 2014; Revised 14 November 2014; Accepted 25 November 2014

1. Introduction

Winter 2013/2014 in the United Kingdom was remarkable. The country was battered by at least 12 major winter storms over a 3 month period – assessed as the stormiest period that the United Kingdom has experienced for at least 20 years (Kendon and McCarthy, 2015). The series of storms resulted in the wettest winter in almost 250 years (according to the England and Wales precipitation series from 1766), significantly wetter than the previous wettest winter in 1914/1915. Winter rainfall accumulation records were broken for each of the respective national rainfall series from 1910 for the United Kingdom, England, Wales and Scotland, and it was the equal-wettest winter recorded in Northern Ireland (NCIC, 2013, 2014). Potential global drivers of these storms were reviewed by Huntingford *et al.* (2014) (see also Met Office and Centre for Ecology & Hydrology – CEH, 2014).

This study provides the first review of the current generation of operational prediction and observational tools, which underpinned the forecaster guidance and warnings issued to a range of users during winter 2013/2014, and presents a discussion on future research challenges and directions. User needs and the seamless suite of operational prediction systems and observations that aim to meet those are introduced in Section 2. The observation and model performance across timescales is discussed in Section 3, highlighting the particular challenges for verifying high-resolution precipitation forecasts. The future evolution of this operational toolkit is described in Section 4 and conclusions are drawn in Section 5.

1.1. Climatological context and impacts

Monthly mean rainfall anomaly maps (relative to the 1981–2010 climatology) shown in Figure 1 illustrate the UK rainfall distribution of winter 2013/2014. Much of central–southern England, parts of South Wales and eastern Scotland received around half a year's typical rainfall over this period, with a few locations approaching 75% of the annual average in just over 2 months (NCIC, 2014). These also highlight the fact that the December storms, including the 5 December storm which resulted in a significant storm surge along the east coast, mainly tracked across Scotland. December was Scotland's wettest calendar month in a series from 1910. From late December, the storm track shifted slightly southward, leading to the persistent rainfall during January and February, which particularly impacted southern England and Wales.

The time series of gauge and radar observations in Figure 2 illustrates that it was the persistence of storms in rapid succession rather than the intensity of rainfall for any specific period, which led to the high rainfall accumulations and records being broken. A notable exception occurred on 23 December 2013 (Figure 3), when frontal systems associated with an exceptionally deep area of low pressure to the west of Scotland brought particularly

* Correspondence: H. Lewis, Met Office, Exeter, UK.
E-mail: huw.lewis@metoffice.gov.uk
This article is published with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

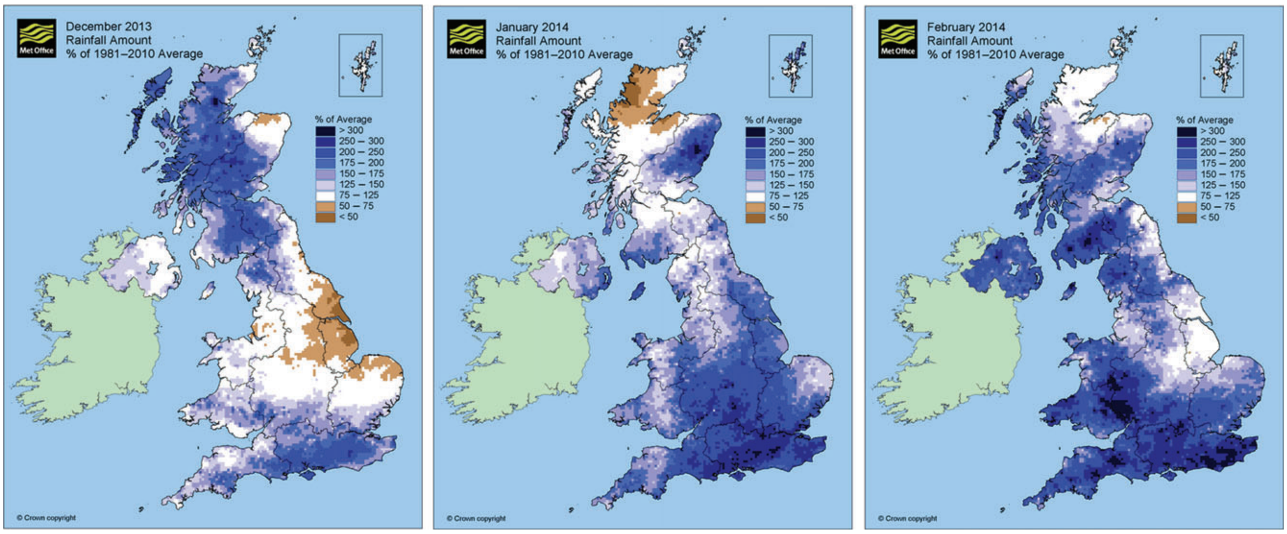

Figure 1. Gridded monthly observed rainfall totals during December, January and February 2014 (Source: Met Office National Climate Information Centre). These maps are based on 5 km resolution monthly gridded datasets at 5 km resolution, based on a network of more than 2500 rain gauges across the UK. The gridding methodology is described in Perry and Hollis (2005). Note that precipitation ranges in the figure legend include values greater than or equal to the lower bound and less than the upper bound indicated.

heavy rainfall to southern England and Wales. Typically, around two-thirds of the average rainfall amount in December was recorded over a 24 hour period in southern and south eastern England (Kendon and McCarthy, 2015), causing significant flooding problems (including at Gatwick Airport) over the Christmas period.

The impact of the severe winter storms on individuals, businesses and the government were substantial, including several fatalities, widespread power cuts and damaged infrastructure. These impacts were often mediated by other parts of the environment rather than directly by strong winds or high rainfall accumulations. On a number of occasions, large wind-driven waves and storm surge resulted in coastal flooding along large stretches of the eastern, southern and western coastlines of England and Wales. A prolonged storm surge affecting three successive high tides with strong winds and large waves along the east coast on 5–6 December 2013 (Figure 4) was the largest in the North Sea since the devastating event of 1953, and led to the highest coastal water levels ever recorded in many areas (Sibley *et al.*, 2015). The succession of low pressure systems producing rainfall over saturated ground led to extensive and protracted river and surface flooding in areas of southern England, most notably on the Somerset Levels and in sections of the River Thames, with groundwater flooding also reported from February 2014. Estimated outflows from rivers in the United Kingdom remained close to the highest levels ever recorded during late December and throughout most of January across large parts of England and Wales (Muchan *et al.*, 2015). The Environment Agency (covering England and Wales regions) issued 155 severe flood warnings during this period – over half the number ever issued since September 2000, compared with just four in 2012. In total, an estimated 7800 homes and nearly 3000 commercial properties were flooded from coastal, river, surface or groundwater sources (UK Government, 2014).

2. Protecting lives and livelihoods

2.1. Responsibilities and service provision

Protecting life, livelihoods and property is a fundamental responsibility of national governments. Weather-related hazards, including coastal and inland flooding, are high priorities in the UK government National Risk Register of Civil Emergencies (UK Government, 2013). The basic requirements of the UK government for providing weather information and for issuing warnings are delivered through the Public Weather Service (PWS), which draws on a broad range of directed research and development, both within the Met Office and across the wider UK and international environmental science community. For periods of high impact weather, the wealth of observational and model information is encapsulated in the National Severe Weather Warning Service (NSWWS). Warnings for rain, snow, wind, fog or ice are issued, together with a text assessment from the Chief Meteorologist, at a yellow (Be Aware), amber (Be Prepared) or red (Take Action) level. Since 2011, the warnings that the NSWWS issues depends on a combination of both the likelihood of the event happening and the impact the conditions may have, rather than any pre-defined meteorological thresholds (Neal *et al.*, 2013). A red warning is therefore only issued for high impact events when there is a high likelihood of occurrence.

These weather services also underpin the work of other partners, particularly the relevant national environment agencies, in delivering other life-critical environmental information such as flood warning and response. Operational hazard services are increasingly being delivered through collaborative approaches, recognizing the fact that hazards are invariably inter-dependent and require multi-disciplinary expertise and capability. In the United Kingdom, this is now embodied in the Natural Hazards Partnership, which is formed of 18 public agencies and departments to provide a common and consistent source of advice for civil contingencies planning and response. This includes the Flood Forecasting Centre (FFC) – a partnership between the Environment Agency and Met Office, which links meteorological and hydrological expertise and advice to provide common regional-scale guidance on flood risk. A similar cross-disciplinary function is performed by the Scottish FFC, a partnership between the Scottish Environment Protection Agency and the Met Office. These services are critically underpinned by collaboratively developed models of the atmosphere (led by the Met Office), ocean, waves and surge (led by the Met Office and National Oceanography Centre, NOC), land surface

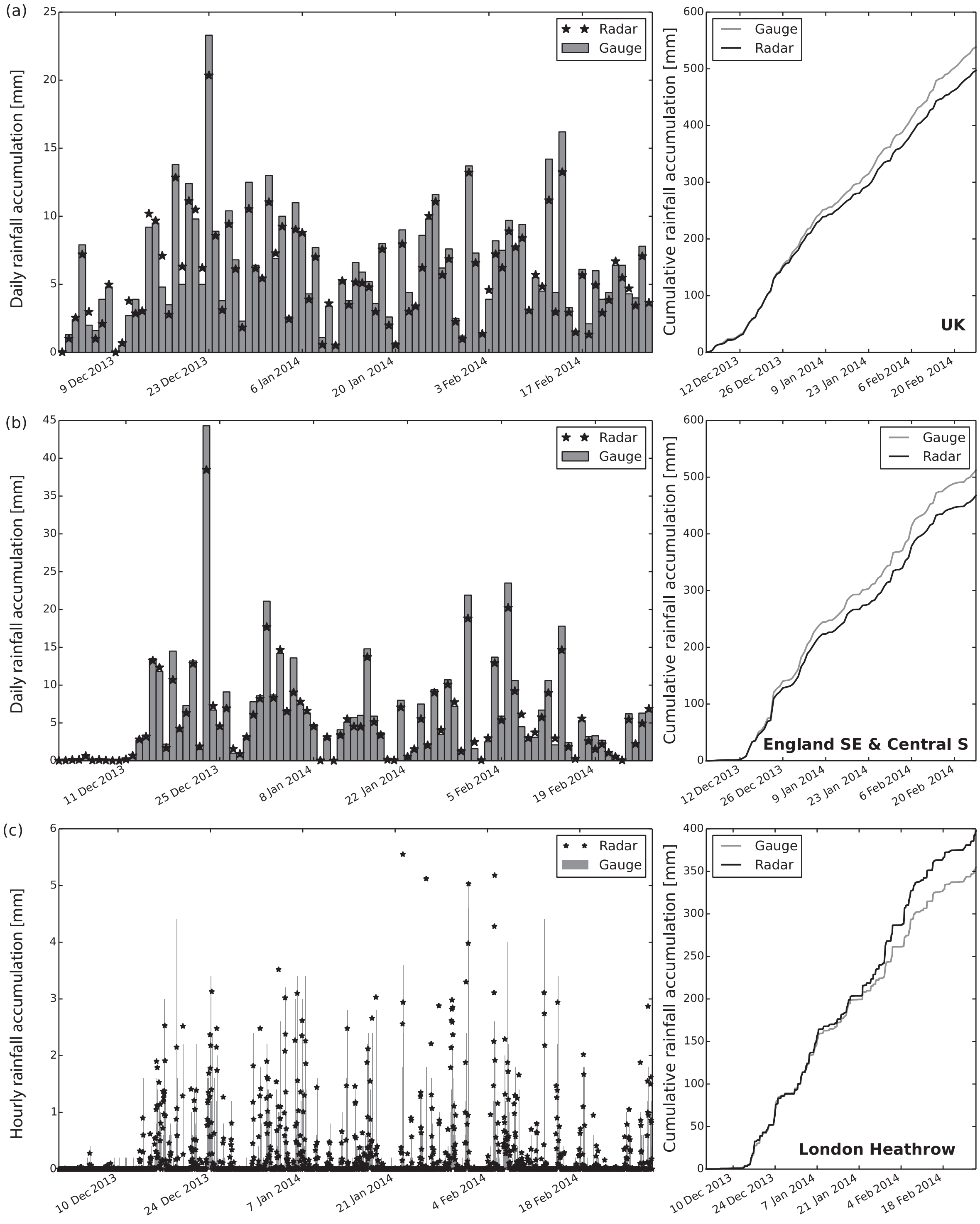

Figure 2. Time series illustrating persistent rainfall during winter 2013/2014 in the UK, as determined by daily precipitation accumulations from gauge and radar observations as (a) national average (UK), (b) regional average (England South East and Central South) and (c) at a specific point (London Heathrow).

(led by the Met Office and Centre for Ecology & Hydrology, CEH) and hydrology (led by CEH).

Advanced environmental warning delivers tangible benefits, enabling sufficient time for relevant authorities, businesses and individuals to take mitigating actions. Typically, these result in lives saved and direct costs of several billion pounds *per* year avoided (e.g. Lazo *et al.*, 2009). With regards to the warnings of high impact weather and other related hazards, the earlier that warnings can be issued, the greater their reliability or the geographic specification, the more effective that mitigating actions can be with greater associated benefits resulting. Although the impacts of the winter 2013/2014 storms were substantial, the number of lives, livelihoods and properties protected is noteworthy, particularly when assessed against similar previous events (e.g. 1953, 2000 and 2007). Much of this improved response and protection can be attributed to improved coastal and flood defences and more accurate and actionable forecast information and warnings communicated several days in advance. The

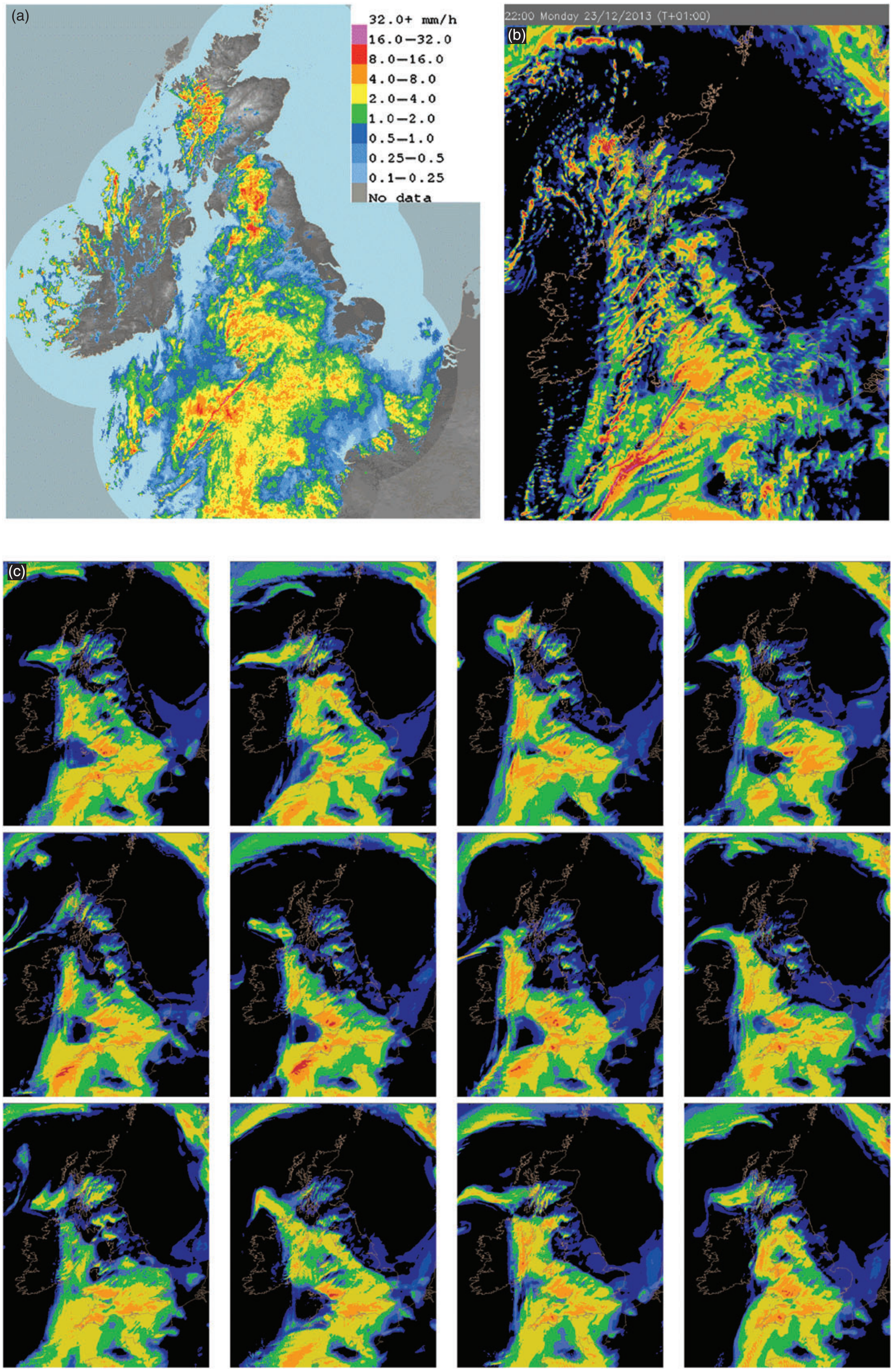

Figure 3. (a) Snapshot of UK rain radar surface rainfall rate for 2200 GMT on 23 December 2013, illustrating the extent of rainfall. This storm brought particularly heavy rain to southern England and Wales – typically around three-quarters of the December average rainfall amount in some places, causing significant flooding problems. (b) Equivalent rainfall rates for same time predicted ($T + 1$ h) by the deterministic UKV high-resolution model, running at 1.5 km, illustrating the very realistic and detailed rainfall distributions simulated. (c) Hourly rainfall accumulations predicted by each member of the UK ensemble (2.2 km), illustrating some variation in detail of the rainfall structure within the general frontal pattern. All plots are shown on the same colour scale.

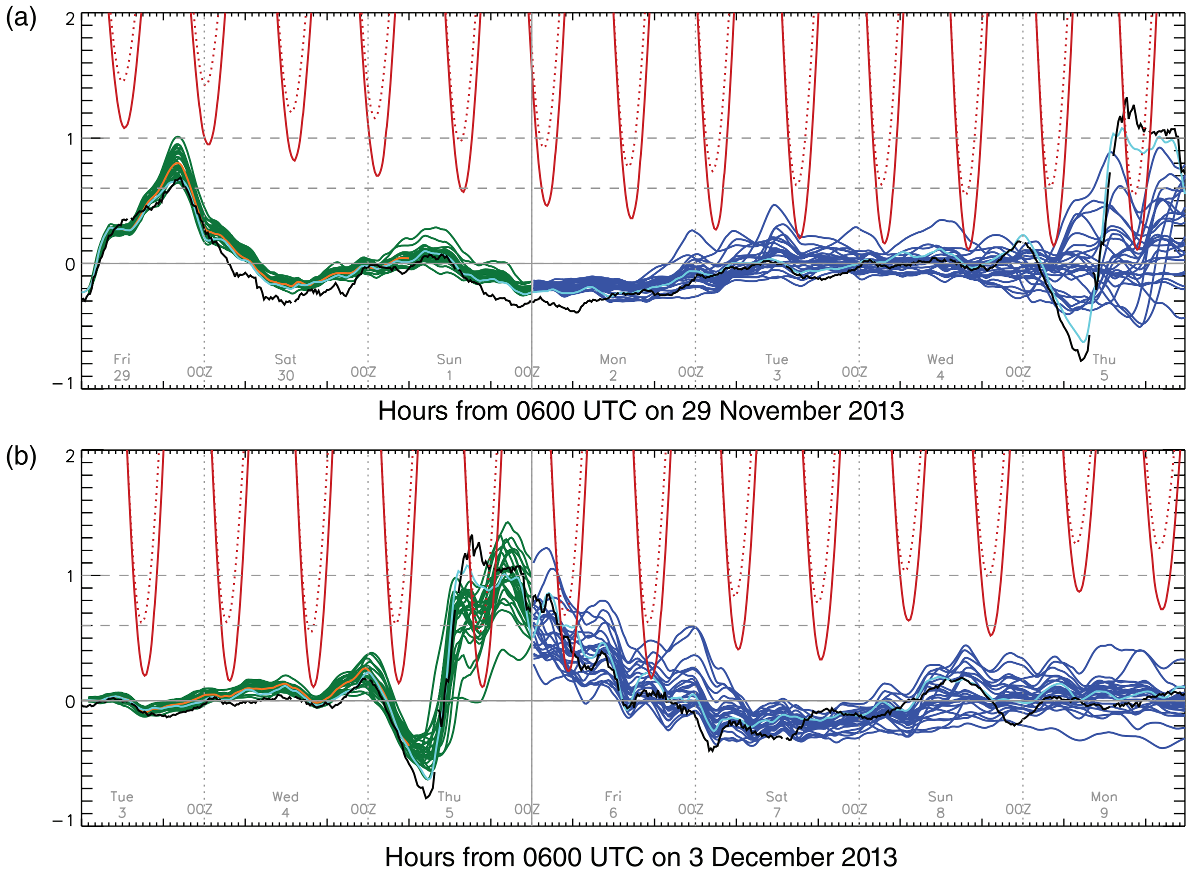

Figure 4. Ensemble forecast time series of surge level at the port of North Shields on the East Coast of England for the surge event of Thursday 5 December 2013 originated at (a) 0600 UTC on 29 November and (b) 0600 UTC on 3 December. Line colours indicate the NWP source providing the driving meteorology: MOGREPS-G (green), MOGREPS-15 (blue), deterministic (orange) and hindcast analyses (cyan). The black line illustrates the verifying observations. Red lines indicate the dynamic alert level, which depends on the astronomical tide, calculated using the optimal harmonic tide forecast (solid) and the model tide forecast (dashed).

Environment Agency estimated that over 1.3 million properties in England and Wales were protected throughout the winter by flood defences or by taking mitigating actions (e.g. UK Government, 2014). For example, the Environment Agency Thames Barrier was closed on 50 occasions, over one quarter of all closures since its completion in 1984.

2.2. An operational toolkit

The remainder of this study discusses the observational and modelling toolkit available to the operational meteorologists providing this national guidance, both through NSWWS in the hours and days ahead and through longer lead time outlooks. Table 1 summarizes the modelling components in this toolkit during winter 2013/2014, from the monthly to seasonal predictions through to nowcasting in near-real time.

This toolkit, and the value of services it provides to users, results from a long history over several decades of research that has progressed in atmospheric, ocean and land surface science, including better understanding of physical processes, improving the availability and use of observations and developing improved numerical prediction and post-processing systems. The Met Office strategy for improving short-range forecasts is built on moving towards the use of twin configurations of the Met Office Unified Model system (Davies *et al*, 2005): a global mesoscale coupled ensemble for daily to seasonal forecasting and a UK convective-scale coupled ensemble for very short-range forecasting to 2 days ahead.

2.2.1. Global observation and modelling – months, weeks and days ahead

Advances in global weather prediction over recent decades have dramatically improved the capability to provide early warnings of severe weather events. At the Met Office this has been realized in a seamless framework, recognizing the synergies between predictions on weather and climate scales (Brown *et al.*, 2012). For example, the Met Office has typically realized improvements in global numerical weather prediction (NWP) skill at a rate of around 1 day increase in lead time for comparable forecast accuracy for every decade of research and development effort. Based on headline metrics, forecasts of comparable skill to next-day forecasts in the mid-1980s might be available on an average 4 days ahead of time today. This lead time improvement translates to a significant improvement in the ability of users to respond effectively on the advice provided to them.

At lead times from one to several months ahead, ensemble predictions are provided by the Global Seasonal Forecast System (GloSea, Arribas *et al.*, 2011). Now in its 5th generation, this coupled atmosphere-land-ocean-sea-ice ensemble prediction system is built around a highest resolution version of the coupled climate model HadGEM3 (with an unprecedented ocean resolution of 0.25° and an atmospheric horizontal resolution of <1°). As commonly found in seasonal forecast systems, the

Table 1. Summary of Met Office operational weather forecasting configurations from months to minutes available during winter 2013/2014 to support forecast services.

Leadtime	Operational Met Office Unified Model configuration	Domain	Update frequency	Ensemble members	Horizontal resolution (vertical levels)	Complexity*	Primary additional info sources	Downstream partner systems / outputs	Example key users	Illustrative responses to information
Months ↓	Glosea5 (MacLachlan *et al.,*2014)	Global	Daily (two forecasts)	42 (lagged over 3 weeks + stochastic)	A: ~60 km (85 L) O: 0.25° (75 L)	A-L-O	WMO GPC; climate monitoring	e.g. UK Hydrological Outlook	Contingency planners	Enhanced monitoring
Weeks ↓	MOGREPS-15 (Bowler *et al.*, 2008)	Global	12 h (to 15 days)	24 (stochastic)	60 km (70 L)	A-L	ECMWF		NHP partners	Contingency/ responder preparedness
	Global NWP	Global	12 h	1	25 km (70 L)	A-L	ECMWF; satellite obs		National resilience (e.g COBR)	
Days ↓	MOGREPS-G (Bowler *et al.*, 2008)	Global	6 h (to 7 days)	12 (44–9 h, stochastic)	33 km (70 L)	A-L	ECMWF; other NMSs satellite obs	e.g. Storm surge ensemble		Public and business awareness
	Euro4	Europe	6 h	1	4.4 km (70 L)	A-L	ECMWF, other NMSs,		Business customers	
Hours ↓	MOGREPS-UK	UK	6 h	12 (lagged)	2.2 km (70 L)	A-L	satellite obs; surface obs; radar	e.g. Grid-to-grid hydrological modelling	General public	Alerts/ warnings
	UKV (Tang *et al.,* 2012)	UK	3 h	1	1.5 km 70 L)	A-L	satellite obs; surface obs; radar	e.g. NHP Hazard Impact Model	FFC warnings	Response deployment
Minutes	Nowcasting	UK	15 min/h	24 (STEPS)	2 km	A	*In situ* obs; satellite obs; radar; WoW	e.g. catchment-specific flood models	Emergency services	Evacuation/ rescue

Listed are the other primary sources of information used alongside model output to inform users, for example, downstream systems or products (typically by or with partners), and an illustration of key users and actions.
*The *complexity* metric summarizes where atmosphere, land, ocean and sea ice components are coupled as part of the current prediction system.

prediction skill for ENSO is high in GloSea5 (MacLachlan *et al.,* 2014). The high-resolution climate model at the core of GloSea5 has also been shown to reproduce accurately the frequency of Atlantic blocking (Scaife *et al.*, 2011) and shows significant skill for winter extratropical flow and the North Atlantic Oscillation, allowing skilful winter forecasts of extreme winter events such as winter wind speed and storminess (Scaife *et al.*, 2014a).

For NWP timescales, typically out to two weeks ahead, the global deterministic and ensemble configurations are currently run as atmosphere-land simulations with a grid resolution of around 25 km at mid-latitudes. The model is initialized using a hybrid ensemble/4D-Var (4D variational) data assimilation system, exploiting a 44-member ensemble to provide the statistics on flow-dependent background error (Clayton *et al.*, 2013). Ocean surfaces are represented by the persistence of observed sea-surface temperature (SST) anomalies. Global ocean conditions are forecast using separate ocean and wave models driven by wind and pressure fields from the atmosphere forecasts.

The ensemble system also provides a 24-member time-lagged ensemble forecast four times *per* day to 3 days at 33 km resolution (MOGREPS-G) and a 24-member forecast twice a day to 15 days at 60 km resolution (MOGREPS-15) (Tennant and Beare, 2014; Bowler *et al.*, 2008). A first-guess warning tool for severe weather (EPS-W, Neal *et al.*, 2013) routinely post-processes ensemble forecasts from MOGREPS as well as the ECMWF (European Centre for Medium-range Weather Forecasts) using the low, medium and high NSWWS impact thresholds for rainfall and other warning parameters. Wind and pressure fields from each member of the MOGREPS global ensembles are used to drive a storm surge ensemble out to 7 days ahead (Figure 4; Flowerdew *et al.*, 2010, Flowerdew *et al.*, 2012). This is the main tool currently used for coastal flood forecasting in the FFC, provided by running the CS3X storm surge model developed by NOC (Horsburgh and Wilson, 2007).

One of the key factors contributing to the improvement in forecast skill has been the expanding use of observations from satellites in global (and UK) NWP models in variational assimilation systems. Satellite-based and *in situ* (surface and upper air) observations are received typically within 3 h of the measurement time. Their relative impact in the Met Office global model has been assessed using an adjoint-based sensitivity method (Joo *et al.*, 2013), which showed that infrared and microwave radiances from polar orbiting satellite missions have the greatest impact on the short-range forecast accuracy, with radiosonde and aircraft profiles giving the next biggest impact. Satellite observations of atmospheric humidity have an important influence on improving precipitation forecasts. These include top of atmosphere radiance observations from both infrared and microwave sensors and total zenithal delay from ground-based GPS measurements. The direct assimilation of precipitation products from satellites remains a challenge, although the recently launched Global Precipitation Measurement mission (Smith *et al.*, 2007) will provide a new global dataset to enhance research in this area.

2.3. High-resolution local-scale observation and modelling – days and hours ahead, impact and detail

Arguably the most significant revolution in public weather forecasting over the past decade has been the development and operational implementation of km-scale, convection permitting forecast configurations. Rainfall fields exhibit important sources of variability over a wide range of spatial scales and these must be adequately represented if models are to provide skilful forecasts of heavy and extreme rainfall. Higher resolution, along with corresponding upgrades to relevant physics parameterization schemes has enabled the intense and localized weather features, such as intense rainfall that cause so many of the highest impact natural hazards, to be better captured.

For high-resolution modelling of the United Kingdom, a variable resolution deterministic version of the Met Office Unified Model is applied (UKV), with a 1.5 km × 1.5 km resolution grid inner domain covering the United Kingdom and a transition region out to 4 km × 1.5 km (with 4 km × 4 km in the corners) towards the edges to nest directly into the global model (Tang *et al.*, 2012). In 2013, a corresponding 2.2 km ensemble configuration (MOGREPS-UK) was implemented operationally, after an initial demonstration as part of the service provision for the London 2012 Olympics (Golding *et al.,* 2014).

Recent verification statistics of surface weather parameters suggest that high-resolution atmosphere models provide significant additional skills over the latest global models, equating to around a decade's worth of research and development. This now means that useful forecasts can be made for a few hours or even days ahead of the location, timing and intensity of some high impact weather events, permitting the use of a 'warn-on-forecast' approach to respond to these hazards in place of the traditional 'warn-on-observe' approach. The use of the 1.5 km resolution convection permitting configuration to better understand potential future climate scenarios is also now being explored (Kendon *et al.*, 2014).

Much of the benefit from the assimilation of observations of the UK models is currently provided by their impact on the global model skill at the UK model boundaries. The benefits of convective-scale data assimilation (using 3D-Var) and observing systems is currently found to be related to assimilating satellite and *in situ* observations at higher resolution than in the global model (Dow and Macpherson, 2013), although radar-derived winds, temperature information from roadside sensors and cloud fraction information from the Meteosat satellite are additionally assimilated in the UK model. This suggests that there is further work to be done to better use the information content derived from high-resolution observations in these systems.

While the realism presented by convective activity in a particular forecast simulation of the current convection-permitting systems is remarkable (e.g. Figure 3(b)), the need for probabilistic approaches and appropriate post-processing is vital to provide a useful representation of likely extremes (Figure 3(c)). The life-cycles and, hence, predictability timescales, of individual showers are short, of the order 1–3 h. This means that local detail is inherently uncertain and there is a rapid error growth in deterministic predictions of rainfall at convective scales. A neighbourhood processing technique is therefore used to estimate the probability of precipitation at any point. In this approach convective precipitation predicted at a particular location may be equally likely to occur at nearby locations (Roberts, 2003b; Theis *et al.,* 2005). While this technique may be applied to the single 1.5 km resolution deterministic forecast, it is far more powerful when applied to MOGREPS-UK. It is currently applied on a 2 km post-processing grid with a neighbourhood of 7 grid-lengths (14 km) for precipitation giving 15 × 15 grid-points, thereby effectively increasing the ensemble size from 12 members to 2700 (although not fully independent). The disadvantage of neighbourhood processing is that some useful resolved detail, such as orographically enhanced precipitation or well-resolved frontal precipitation, can be erroneously distributed to probabilities at neighbouring locations; so, it must be used with caution. Future enhancements include an adaptive neighbourhood scheme that adapts the spatial scale of the neighbourhood according to the degree of spatial agreement in the fields, or restricts the grid-points used to those with similar characteristics such as elevation or land/sea.

The quality of these high-resolution systems are also highly dependent on the quality of the forcing information at the boundaries, that is provided by the relevant global configuration. The dual approach to maintain and develop world-class global and local scale modelling capability is, therefore, particularly important. During the winter of 2013/2014, the value of the UKV predictions were in determining the local distribution of rainfall within the frontal structures, which were generally well predicted and established by the global scale capability.

2.4. Very short-term predictions – nowcasting

Nowcasting remains a powerful tool at very short lead-times (0–6 h), which for precipitation is still based on extrapolating radar and satellite estimates of surface precipitation (Wilson *et al.*, 1998; Pierce *et al.*, 2012). Ballard *et al.* (2012) reported on new developments in using NWP-based nowcasting. An hourly cycling nowcast configuration of the Met Office Unified Model, exploiting 4D-Var was demonstrated during the London 2012 Olympics. Sun *et al.* (2014) recently reviewed these developments across a number of operational centres.

The Met Office UKPP post-processing system generates high-resolution (2 km) ensemble nowcasts of surface precipitation rate and accumulation on a 15 min cycle using the Short-Term Ensemble Prediction System (STEPS, Bowler *et al.*, 2006; Seed *et al.*, 2013). Each nowcast is generated from a scale-selective blend of extrapolated observations with the most recent high-resolution NWP forecast and a time series of synthetically generated precipitation fields (noise), which has space-time statistical properties inferred from weather radar. The noise component aims to account for uncertainties in the evolution of the extrapolation and NWP forecast components and also to downscale the NWP forecast.

More recently, STEPS has been applied to generate 2 km resolution blended ensemble precipitation forecasts out to 32 h ahead, incorporating an extrapolation nowcast and MOGREPS-UK ensemble precipitation forecasts. The aim is to improve the calibration of MOGREPS-UK by generating additional ensemble members that allow a more comprehensive representation of the forecast uncertainties at convective scales. These are now used to drive the operational distributed Grid-to-Grid river flow model developed by CEH (Bell *et al.*, 2007) to support the FFC.

2.5. High-resolution precipitation observations – radar

The Met Office collates observations from around 4000 daily reporting storage rain gauges across the United Kingdom, many operated by partner organizations, which provide the most accurate measurement of long-term accumulated precipitation (e.g. Figure 1). However, many gauge observations are not necessarily available in real time and the network may not capture small scale variation in precipitation intensity. Measurements are also prone to errors, for example as a result of frozen precipitation, debris blocking the gauge or under estimation in strong winds. (e.g. Habib *et al.*, 2013).

The weather radar network is therefore a critical component of the United Kingdom observing capability in providing a unique characterization of precipitation on scales of the order 1 km every 5 min (e.g. Figure 3(a)). The Met Office Radarnet processing system attempts to derive the best quality national 5 min composite surface precipitation estimate from radar reflectivity measurements (Harrison *et al.*, 2012). Doppler capability also allows winds to be estimated in precipitation with similar spatial and temporal resolution, enabling the detection of hazards such as wind shear. The assimilation of these observations has been shown to improve rain and wind significantly in NWP-based nowcasts (Simonin *et al.*, 2014).

Experiments are in progress to assimilate radar reflectivity measurements directly into high-resolution NWP models, which will allow information from the full 3D radar sampled volume to be more fully used. The radars' signal processing has also been recently modified to extract information on atmospheric refractivity on scales of about 5 km. These data will only be available over about a quarter of the UK land area, but when assimilated alongside complementary data (e.g. from ground-based global navigation satellite system receivers) may enable regions of moisture convergence and convective storm initiation to be more accurately located.

The UK weather radar network is currently undergoing complete renewal. Each radar is being sequentially replaced over a 5 year period with Doppler and dual-polarization systems developed in-house.

3. Evaluating performance

The current operational toolkit was sternly tested throughout the winter of 2013/2014. This section presents and discusses an evaluation of the deterministic and ensemble systems across scales, focussing in particular on the longer range global scale monthly predictions and the short-range km-scale predictions. These represent the most recent developments in the seamless toolkit and highlight particular challenges for verifying and demonstrating the value of these capabilities for predicting rainfall.

3.1. Performance of global-scale long-range predictions (GloSea5)

A large set of independent events, rather than a single period such as winter 2013/2014, is required to determine skill scores for ensemble seasonal forecasts. It is therefore only appropriate to examine whether longer-range forecasts for winter 2013/2014 contained predictable signals and how these signals compared to observed events.

Three out of four of the summary longer-range forecasts covering winter 2013/2014, produced each month from October to January, indicated an increased risk of more westerly conditions than normal in the Atlantic and hence enhanced the risk of some combination of mild/stormy/wet conditions for northern Europe (e.g. Figure 5). For example, the summary temperature forecast issued to government contingency planners in October 2013 indicated that a mild winter was about twice as likely as a cold winter and that enhanced storminess was likely over the coming three months. Forecasts also suggested increased risk of northerly advection over North America, as indeed occurred. This was associated with a persistent pattern of a high pressure anomaly over the Gulf of Alaska and a low pressure anomaly over North America in successive forecasts. Only the November forecasts suggested that a different outcome was likely for northern Europe when the signal for high pressure temporarily moved north over the United Kingdom (in both GloSea5 and other real time forecast systems such as monthly forecasts from ECMWF).

Forecast circulation patterns over the winter of 2013/2014 therefore repeatedly suggested that predictable signals were contained in the ensembles of forecast members. Ensemble mean anomalies of several hPa were predicted (e.g. Figure 5) and these exceed the uncertainty due to a finite ensemble size in the extratropics (of approximately 1 hPa). This suggests that external drivers of the winter circulation anomalies in and around the Atlantic basin were present. At least two candidates have been identified as discussed by Huntingford *et al.* (2014). The anomalous hemispheric-scale wave pattern in Figure 5, with low pressure over the west Pacific, high pressure over the Gulf of Alaska and low pressure over North America is commonly seen in response to strong tropical convection in the West Pacific (Palmer, 2014). In addition, strong westerly flow over the North Atlantic is often seen in response to westerly phases of the tropical quasi-biennial oscillation (QBO), which is predictable in this forecast system (Scaife *et al.*, 2014b). Both enhanced west equatorial Pacific rainfall and westerly QBO were predicted in these forecasts and the associated tropical to extratropical teleconnections mean that winter 2013/2014 conditions around the Atlantic basin were at least partly predictable on seasonal timescales.

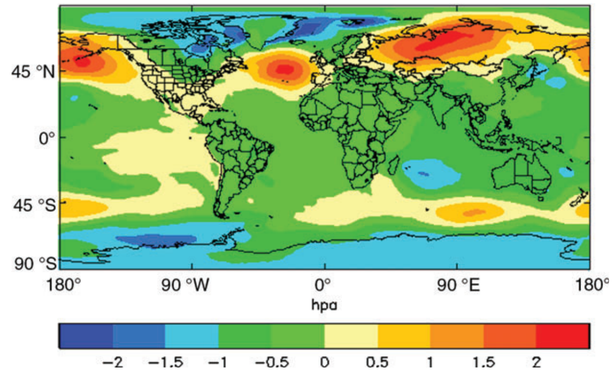

Figure 5. Seasonal forecast of winter mean (December, January, February) surface pressure from forecasts starting in October. Low pressure over the Arctic and high pressure to the south suggested increased risk of stronger than normal westerly flow over the Atlantic, and enhanced probability of mild UK conditions for early winter. Forecasts also indicated increased risk of higher than normal pressure over the Gulf of Alaska – a key region for influencing events downstream, and northerly advection over North America.

3.2. Performance of national-scale high-resolution deterministic predictions (UKV)

Table 2 summarizes the performance of 6 h UKV precipitation forecasts at 24 h lead time using the equitable threat score (ETS) metric during winter 2013/2014, with a comparison to the previous winter seasons and long-term average scores for the period between December 2011 and August 2014. Jolliffe and Stephenson (2011) provide for a comprehensive list of contingency table-based metrics such as the ETS.

Results are calculated comparing forecasts against all verifying UK rain gauge sites for three different exceedance thresholds (4, 8 and 12 mm accumulations in 6 h). Scores for winter 2013/2014 are marginally higher than for the previous two winters, and considerably better than the long-term average scores. This might be expected to some extent as, due to the typically frontal nature of UK winter precipitation, skill scores tend to be higher than for other times of the year, although winter forecasting can present its own challenges such as wintertime convection associated with cold air outbreaks leading to snow showers. The observed frequency of occurrence (or base rate) of each accumulation threshold across the UK rain gauge network during each period is also provided in Table 2. Note that scores for the higher accumulation thresholds where the frequency of occurrence is less than 1% are very difficult to interpret statistically, and cannot be considered robust, especially over a 3

Table 2. Summary of UKV precipitation forecast performance at 24 h lead time based on the equitable threat score (ETS) calculated for different 6 h accumulation thresholds for the three most recent winters (December to February), and for a longer assessment period (reference).

Time window	≥4 mm 6 h^{-1}	≥8 6 mm 6 h^{-1}	≥12 mm 6 h^{-1}
December to February 13/14 ($t+24$ h)	0.35 (10%)	0.31 (3%)	0.27 (1%)
December to February 12/13 ($t+24$ h)	0.36 (6%)	0.27 (2%)	0.25 (<0.6%)
December to February 11/12 ($t+24$ h)	0.33 (5%)	0.31 (1%)	0.3 (0.2%)
December 2011 to August 2014 ($t+24$ h)	0.29 (5%)	0.22 (1.7%)	0.17 (<0.6%)

Indicated in brackets is the representative observed frequency of occurrence of events (base rate).

month window, although the long-term average scores for 12 mm 6 h^{-1} might be considered more stable due to the size of the sample.

Early work on km-scale modelling showed that whilst the detail of precipitation features looked realistic (e.g. Figure 3(b)), they were often in the wrong place or occurred at the wrong time. Traditional metrics that rely on a precise matching of the forecast and observation at a location will treat these situations as both a false alarm and a miss (a 'double penalty') (e.g. Mittermaier, 2014). Note that when conditions are as wet as during winter 2013/2014, the double-penalty effects on verification are somewhat tempered.

The double penalty effect, and the corresponding challenge to show that increasing horizontal resolution leads to an improvement in precipitation forecast skill, led to the development of many spatial verification methods (e.g. Gilleland *et al.*, 2009; Mittermaier and Roberts, 2010). A new probabilistic verification framework that uses a neighbourhood around an observing site is replacing the traditional gauge-based verification (Mittermaier, 2014). Results show that the skill in the UKV forecast against persistence is improved when considering even a ~4 km (3 × 3 grid points) neighbourhood over using the nearest grid point to an observing site. The improvements can be further increased by using a larger neighbourhood.

To reduce gauge sampling issues, radar data have also been used extensively to verify high-resolution km-scale NWP for at least a decade (e.g. Roberts 2003a; Roberts 2008; Mittermaier 2006, 2007, 2008; Mittermaier *et al.* 2013). The need for the spatial detail that radar data can provide is a strong requirement for the verification of high-resolution NWP, particularly in terms of showing the benefit of high-resolution over coarser model grids. Figure 6 shows a comparison of the long-term rainfall accumulations for winter 2013/2014 from the radar and the UKV forecasts. Totals in excess of 1000 mm are apparent over the highlands of northwest Scotland in both the radar and the model, though the UKV also has local maxima in excess of 1000 mm over the Lake District in northwest England and the high ground in Wales and southwest England. Differences between model and radar accumulations are shown in Figure 6(c). This highlights the considerable reduction in the area that can be assessed when periods of incomplete radar data are screened out of the analysis. Some regions of very large accumulation differences (±250 mm) are also evident. This may in part be related to the likely underestimation of rainfall by radar over upland areas, but these are good examples of the double-penalty issue where a precise comparison of matched grids shows that if the forecasts have slightly misplaced the precipitation with respect to the truth, large differences at the local scale will become visible. Despite these large local differences, the main precipitation maxima in terms of the pattern and distribution are generally in remarkable agreement.

Roberts and Lean (2008) introduced the fractions skill score (FSS) as a more objective method to assess high-resolution precipitation forecasts against a gridded observation field. A square 'neighbourhood' of a particular size around each grid point is defined and the fraction of forecast and observed pixels exceeding a threshold in that neighbourhood is compared. Mittermaier *et al.* (2013) recommended the use of frequency thresholds (percentiles of the distribution) rather than absolute thresholds to retain useful information on the spatial extent and positioning of precipitation. Figure 7 summarizes FSS of the upper 5th percentile threshold of 6 h UKV forecast accumulations, relative to 1 km gridded radar observations. Results for winter 2013/2014 are shown as a function of neighbourhood size. A skilful spatial scale, above which the model forecast might be considered to have useful skill, is defined where FSS exceeds $0.5 + f/2$, where f is the observed frequency of an event (Roberts and Lean 2008). From Figure 7 it can be seen that for this period the skilful scale ranges between 30 and 70 km, increasing with increasing lead time. For 21 h lead time forecasts, the value of about 50 km is around 10 km better than that quoted in Mittermaier *et al.* (2013), obtained when evaluating the previous 4 km resolution version of the operational UK model.

The time series of daily FSS values in Figure 8(a) demonstrates clearly the variability in forecast skill during winter 2013/2014. Figure 8(b) shows the accumulation threshold which corresponded to the upper 5th percentile of the daily rainfall distribution during this period. Values were consistently above 5 mm accumulation in 6 h, and occasionally nearer 20 mm, demonstrating just how unusual winter 2013/2014 was. In general, 6 h accumulations with thresholds above 4 mm are rarely assessed for UK verification (especially against gauges). As indicated in Table 2, such totals are typically in the top 1–2% of the long-term climatology for many locations across the United Kingdom, which can lead to serious sampling issues when calculating statistics.

In summary, the UKV forecast performance over the entire winter period in terms of precipitation (and wind, though not discussed here) was quite remarkable in its accuracy and consistency, and is a strong testament to the significant research progress made in developing and implementing these systems over the past decade.

3.3. Performance of national-scale high-resolution ensemble predictions (MOGREPS-UK)

Objective verification of the MOGREPS ensemble precipitation forecasts over a 12 month period, which included the winter storms (mid-April 2013 to mid-April 2014) is shown in Figure 9. For the first time, this presents a reliability diagram for probability forecasts from both the new convective scale MOGREPS-UK (with neighbourhood post-processing) and global MOGREPS-G systems. Forecasts of 3 h precipitation accumulations exceeding 10 mm have been verified against a 2 km resolution UK analysis. Forecast probabilities are calculated for each of the 147 county regions rather than on the model grid, providing a relatively

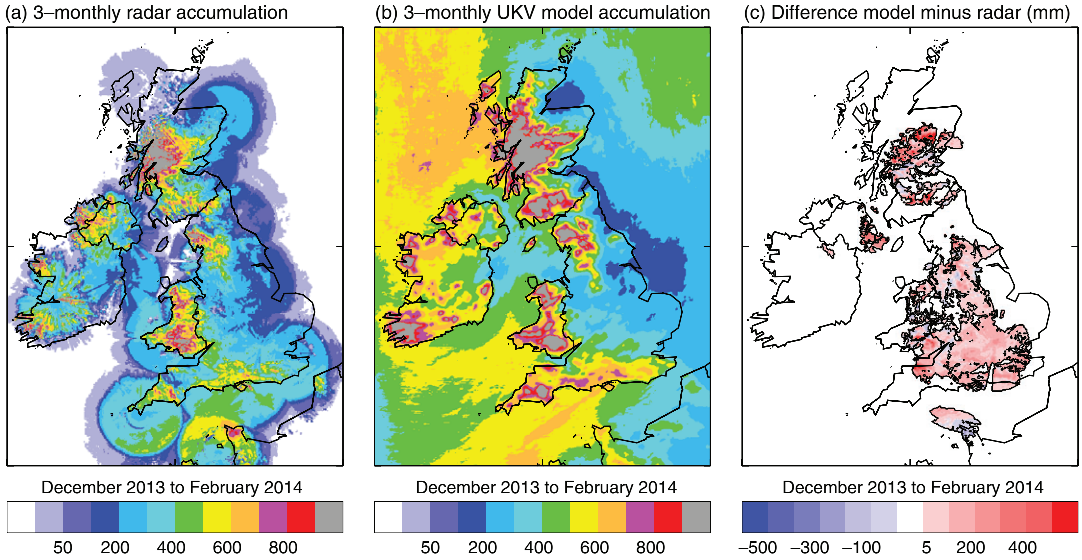

Figure 6. Precipitation accumulations in mm for the period 1 December 2013 to 28 February 2014 as derived from (a) radar observations and (b) UKV forecasts from the 21Z initializations ($t + 3$ h to $t + 27$ h). (c) Difference between radar and forecast accumulations in millimetres, applying a strict radar data quality and availability criterion before computing statistics.

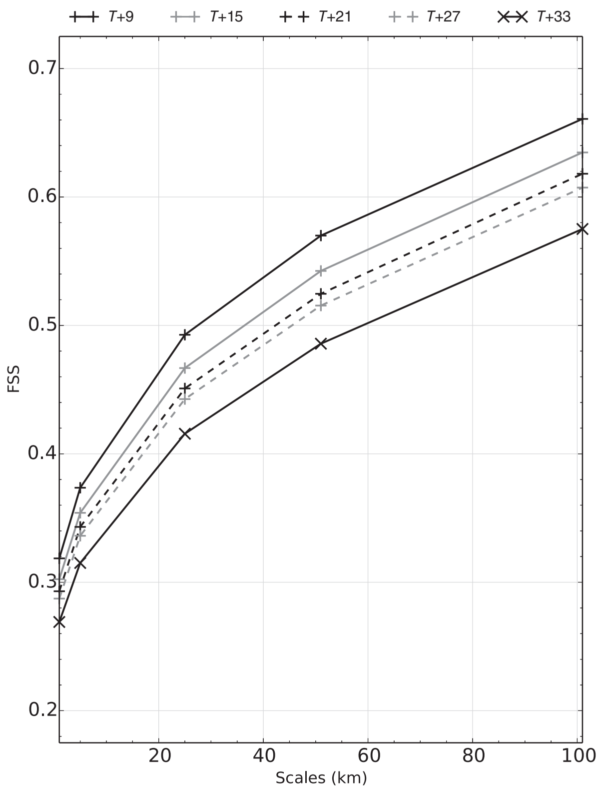

Figure 7. Fractions skill score (FSS) of the upper 5% percentile threshold 6 h precipitation accumulations from the UKV (against radar) as a function of lead time for the period 1 December 2013 to 28 February 2014.

broad-scale neighbourhood verification. A forecast event need only to be observed within the county area boundary to score a hit, giving allowance to minor spatial errors in the model and for some inaccuracies in the observed analysis. Observed frequency is plotted against forecast probability for 12 equally distributed probability bins, excluding all 0% forecasts. Reliability is indicated by the proximity of the plotted lines to the diagonal.

Figure 9 shows that MOGREPS-UK has demonstrably improved forecast reliability compared to the global MOGREPS-G system, with results lying closest to the diagonal for most probability bins. While MOGREPS-G shows under-forecasting of lowest probability events, MOGREPS-UK shows a slight over-forecasting. MOGREPS-G shows significant over-forecasting for middle and higher probability bins, which is much improved with the MOGREPS-UK system. The over-forecasting shown for both ensemble configurations may be a result of under-dispersive ensembles, whereby the precipitation fields in different members are too similar resulting in probabilities which are unrealistically high. It may be possible to improve reliability for MOGREPS-UK by adjusting the neighbourhood size. However, this will need to be considered with care to ensure that the ensemble continues to maintain its resolution (as discussed by Murphy, 1993) in providing forecasts that correctly distinguish situations with distinctly different frequencies of occurrence.

The sample sizes for each probability bin in Figure 9 show a large number of low probability forecasts and very few high probability forecasts. As is common with rare (and severe) events, the number of samples drops off dramatically after the first few probability bins. MOGREPS-UK has a greater number of samples in the lower probability bins compared to MOGREPS-G, which is likely a result of the neighbourhood processing. The low number of samples, particularly in the higher probability bins, makes generating robust verification results for severe events very difficult and therefore these results should be treated with some caution. However, the benefit of the higher resolution ensemble forecasts now available in the operational toolkit to underpin guidance to users is clearly demonstrated.

3.4. Performance of the UK weather radar network

During the winter of 2013/2014, the Cobbacome Cross radar in southwest England was in the process of being upgraded to dual polarization capability. Figure 6 reflects the resulting poorer

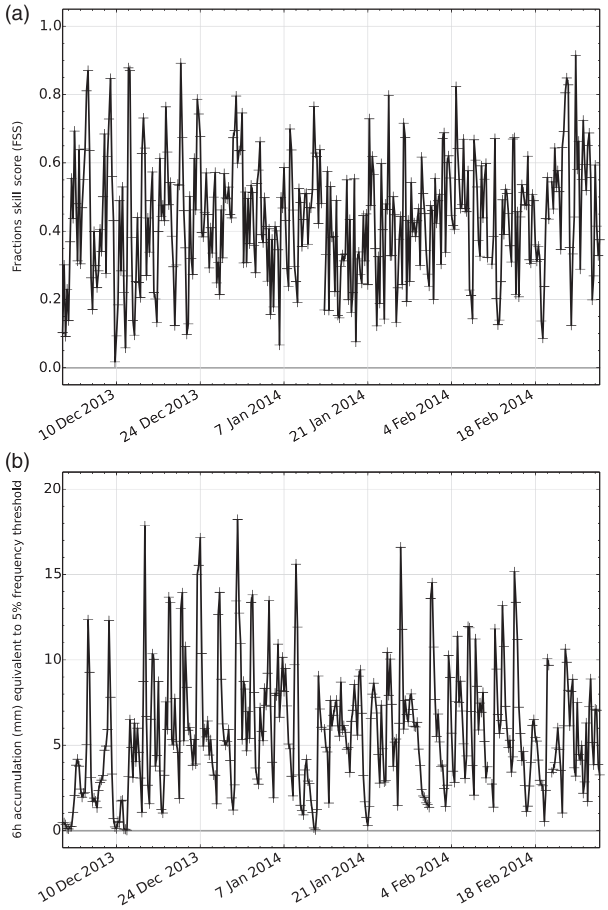

Figure 8. Time series of (a) the fractions skill score (FSS) for 6 h precipitation accumulations exceeding the upper 5th percentile threshold at $t+24$ h for the 25 km neighbourhood. In (b) the equivalent physical threshold in millimetres is plotted to show that for a large proportion of the period the 6 h accumulations were well above the typically used 4 mm/6 h threshold (which is the typically the 98th + climatological threshold for most of the United Kingdom).

coverage in the UK radar composite over southwest England. This radar was temporarily re-introduced into the network over the Christmas and New Year period, in time for the forecast event on 23 December.

Radar precipitation estimates are routinely verified against hourly tipping bucket rain gauge measurements. The root mean square factor (Golding, 1998) comparing radar with gauges was consistently between 2 and 3 for each week of winter 2013/2014, which is typical of the long-term performance of the UK radar network.

Time series of radar- and gauge-based daily accumulations averaged at the UK and regional scales are shown in Figure 2. These show that the radar and gauge precipitation estimates were generally in good agreement, but the radar underestimated precipitation on days with the highest totals. This is particularly evident on 23 December 2013, for example, when the average rainfall based on gauges for the Central South and South East England region was close to 45 mm compared with a radar-based value 38 mm.

This underestimation of daily rainfall accumulations is likely to have been influenced by two main factors, partial radar beam blocking by trees and masts close to certain sites and the current dependence on a climatological relationship between radar reflectivity and rain rate, which tends to underestimate intense precipitation. Both of these issues are the focus of current development activities, for which additional parameters obtained from the new dual-polarizations capability is expected to be of benefit.

4. Discussion

The assessment of the operational toolkit presented here demonstrates the current state of NWP across scales for high impact weather. While this represents a considerable scientific and technical achievement, there remain a number of challenges and opportunities to derive greater value from the predictions for end users. A few of these are discussed here.

4.1. Application of ensemble information ahead of high impact weather

Global NWP performance was particularly highlighted in the media and public imagination ahead of a severe windstorm at the end of October 2013, due to the formation of the storm system in the mid-Atlantic in the simulation several days ahead of it forming in reality and becoming visible on satellite imagery. Nevertheless, the development of such systems several days ahead is frequently sensitive to small analysis and modelling errors, and ensemble predictions provide both measures of confidence and estimation of the risks and uncertainties in the forecast, particularly for extreme and high impact events. Ensemble forecasts are thus a critical part of the toolkit, allowing the issue of alerts for high impact events much earlier than would be achievable with only deterministic model forecasts.

This approach was well illustrated with the storm surge event of 5 December 2013 (Sibley *et al.*, 2015). The surge forecast on 29 November 2013 (e.g. Figure 4(a)), driven by MOGREPS wind and pressure fields, indicated a low probability of a very significant surge event coinciding with a high astronomical tide at a large number of ports on both the east and west coasts. This was sufficient for the FFC to identify the risk in the Guidance provided 6–7 days ahead of the event. The level of risk was steadily raised during the following days with a Yellow issued in the Guidance on Monday, 2 December; Amber on Wednesday, 4 December; and Red from 1030 on Thursday, 5 December 2013. By 2–3 days ahead, the ensemble forecast was indicating very high confidence for an extremely hazardous event, with almost all ensemble members predicting a significant exceedance of the alert level (e.g. Figure 4(b)).

Extreme events such as this are very rarely predictable with high confidence at longer lead times since they are sensitively dependent on many factors coming together to generate them, especially when focussing on the risk at a specific location. This case illustrates how judicious use and communication can provide an effective early alert of a low risk for a potentially very high impact event, progressing to a more confident and localized warning closer to the event. This enables civil contingencies staff to be prepared well in advance and to steadily ramp up their level of preparedness as the confidence increases, including, for example, evacuating vulnerable members of the population, which may require many hours of advance notice. Ensemble forecasts of other weather elements, notably rainfall amounts, were also used during winter 2013/2014 to support planning and modelling of the ongoing flooding levels.

It is clear that the information provided by ensembles will continue to become more deeply integrated in the operational toolkit over the coming years, particularly at high resolution (see

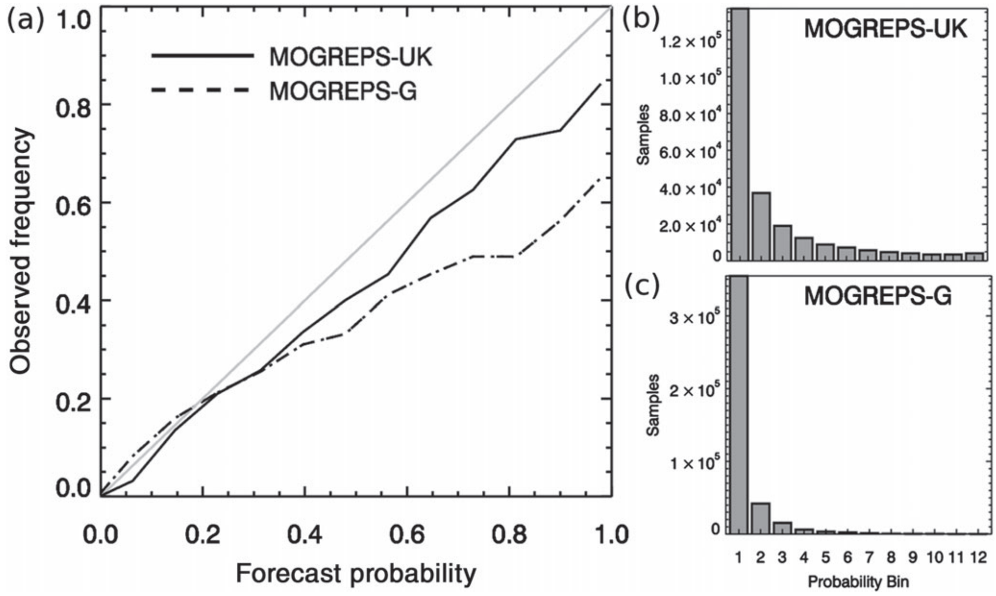

Figure 9. (a) Reliability diagram comparing MOGREPS-UK to MOGREPS-G for 3 h precipitation accumulation forecasts ≥10 mm. Sharpness diagrams showing the number of samples in each probability bin are shown in (b) for MOGREPS-UK and (c) for MOGREPS-G. The verification period for all plots covers the 12 months from 18 April 2013 to 17 April 2014.

Figure 9) and for deriving the most useful post-processed data and guidance to support effective warning and decision-making. There remain significant communication challenges to ensuring effective use of this information, which requires a continuing dialogue with users, and further development of the scientific and interpretive methods to ensure users get the most benefit from these systems. Over the next decade at least, it seems likely that there will be a continuing role for parallel deterministic control forecasts within the overall suite. Unperturbed control forecasts continue to provide the best deterministic representation of evolving weather and of local details.

4.2. The evolving toolkit

Being able to meet user needs requires a continuous evolution of the operational toolkit, underpinned by long-term directed research and development. Future step-changes will build on the existing capability and potential and be facilitated by significant enhancement in high-performance computing over the next few years.

The focus for improving global forecasting continues to be on high impact weather up to 1 week ahead. In spring 2014, a new dynamical core for the Met Office Unified Model named ENDGAME (Wood *et al.,* 2013) was implemented. This provides more active development and an improved climatology of intense storm systems and more realistic precipitation systems with less spatial diffusion. The MOGREPS-G ensemble has also been extended to provide forecasts to 7 days ahead (from 3) at 33 km resolution. Further ahead, on the next high-performance computer system, it is planned to increase grid resolution to around 12 km in the deterministic model and 25 km in MOGREPS-G and GloSea, with a vertical resolution increase to 120 or more levels. Global systems will also be coupled to high-resolution ocean circulation models to better represent ocean–atmosphere interactions (e.g. Shelly *et al.*, 2014). Model physics enhancements are planned to improve the diurnal cycle and sub-synoptic organization of convective showers, while inclusion of prognostic aerosols aims to improve visibility prediction and air quality modelling. Beyond that, work is already underway to develop a new global model framework with grid-points more evenly distributed globally (named GungHo) to provide efficient scalability on next generation computers with many more processors and to enable running ensembles of earth-system coupled models at around 10 km resolution.

Major enhancements to the Met Office km-scale prediction systems are also planned for the forthcoming high-performance computing upgrade. An increase in vertical resolution to over 100 levels in all configurations and potentially increased horizontal resolution, for MOGREPS-UK in particular, is likely, together with an increased number of ensemble members. MOGREPS-UK currently employs a simple downscaling approach, where both initial conditions and lateral boundary conditions are interpolated from the global system. High-resolution 4D-Var UK data assimilation updated hourly will provide enhanced initial conditions, and stochastic perturbations to model physics will be introduced to address uncertainties due to model errors in the ensemble. There will also be continued use of significantly higher resolution models run deterministically, often over smaller areas, as research and development platforms improve model performance. For example, the Met Office is currently experimenting with a 333 m resolution model over southeast England as a potential tool to improve details such as fog and low cloud around London area and its airports.

The upgrade of the UK weather radar network to dual polarization is expected to improve accuracy significantly, particularly in moderate and intense precipitation. This will be enabled by better discrimination between precipitation and non-precipitation targets, better correction for signal attenuation in rain and use of variable relationships between precipitation rates and measured parameters. Research into the assimilation of radar reflectivity, refractivity and Doppler radial winds is ongoing. Assimilation of radar refractivity in providing near-surface estimates of humidity is likely to have a large impact on the diagnosis of localized areas of convection with short lead times.

Weather radar will continue to provide crucial information on precipitation over short timescales for nowcasting and surface water flood forecasting, and increasingly for direct assimilation into NWP. However, the best surface precipitation estimates over longer accumulation periods (hours to days) are likely to be derived by blending radar with other sources of data, such as gauges, and possibly data from microwave links (e.g. Goudenhoofdt and DeLobbe, 2009). For daily or weekly accumulations, there is still a need to understand how to best combine these data sources and establish the most accurate antecedent conditions for hydrology, calibrating hydrological models and validating precipitation forecasts.

4.3. From weather to environmental prediction

Though it has long been understood, the experiences of winter 2013/2014 illustrated well the complex and interdependent nature of the environment, and the value an integrated approach provides to its prediction. For example, guidance on the severe storm surge in early December 2013 required a detailed prediction and synthesis of both the atmospheric (pressure, wind) and ocean state (tides, waves). The flooding events that followed equally illustrated that high impact weather need not in itself be severe to cause significant impacts given the strong preconditioning of the land surface. Useful guidance on the evolving flooding situation required a detailed prediction and synthesis of the atmospheric (rainfall), land surface (soil moisture, runoff) and river state (flow, level).

Recent advances in the skill, resolution and information content (e.g. on uncertainty) of meteorological modelling now make it more relevant to directly integrate or couple forecast models of the atmosphere, oceans, land surface (including hydrology) and other aspects of the environment. This is particularly the case for wind and precipitation, where local geographic or meteorological details can have a significant effect on quantitative prediction. Shapiro *et al.* (2010) discussed the need to accelerate progress in earth system prediction across all scales (climate and weather, global and local). While the need to represent feedbacks between different components of the environment (atmosphere, land, ocean, sea-ice) is well understood and considered mature for climate prediction, the use of coupled approaches is relatively less well developed on shorter timescales. At the Met Office, development towards short-range global coupled prediction and data assimilation has demonstrated modest improvements in coupled atmospheric and ocean forecast skill as compared to uncoupled skill (e.g. Shelly *et al.*, 2014).

For regional high-resolution prediction, there is already evidence of the benefit of coupled prediction for improving weather forecast skill. For example, coupled atmosphere-ice-ocean forecasts are now operational at the Canadian Meteorological Centre for the Gulf of St Lawrence region, with evaluation demonstrating significant improvement in the skill of both atmospheric and ice forecasts (Pellerin *et al.,* 2004; Smith *et al.,* 2013). The development of a flexible and collaborative modelling framework for coupled land-surface and hydrological models in this system was key to improved understanding of the behaviour of different land surface and streamflow forecasts to improve the representation and accuracy of the regional water budget (Pietroniro *et al.*, 2007; Deacu *et al.*, 2012).

Coupled regional prediction systems have also been applied in research mode to improve the representation of air–sea interactions on Bora winds (e.g. Pullen *et al.,* 2006), the evolution of Mediterranean storms (e.g. Renault *et al.*, 2012), hurricane development (e.g. Chen *et al.*, 2010; Sandery *et al.,* 2010; Warner *et al.*, 2010) and on suppressing the urban heat island effect (e.g. Pullen *et al.,* 2007).

In the United Kingdom, the Met Office, CEH and NOC are working with others to develop the foundations of a coupled high-resolution probabilistic forecast system that links the predictions of the atmosphere, coastal ocean, land surface and hydrology. The potential value for this UK Environmental Prediction system is significantly increased by existing linkages between weather, ocean and hydro-meteorological service providers, through collaborations such as the Natural Hazards Partnership. The challenge now is to realize the potential of integrated regional coupled prediction in the UK context. This will require a unified approach across traditional disciplinary boundaries to better understand and represent the interactions between the relevant bio-geophysical systems, and to better observe, initalize and verify these processes. If successful, improved predictions could galvanize research and development effort over the next decade, and should increase the value and use of UK environmental science capability and investment to society.

5. Conclusions

The prolonged period of high impact weather experienced in the United Kingdom during the winter of 2013/2014 was very well forecast by the operational tools available across space and time scales. It serves as a reminder of the remarkable progress in numerical weather prediction, particularly for high-resolution precipitation forecasting. This was underpinned throughout by robust operational computing and modelling and observational infrastructure and expertise, and routinely delivered by forecast experts to a broad range of users and responder communities. It is notable that so much of this activity might now be regarded as 'business as usual'.

This period also highlighted a continued need to derive greater value for users from these capabilities, and their potential future evolution. On monthly to seasonal scales, there is a need to improve the characterization and resolution of potential weather regimes, with future advances most likely through further increasing resolution and improved coupled initialization strategies. On weather forecasting scales, further improvements to the suite of models, and the ability to derive effectively the relevant information contained in high-resolution deterministic and ensemble systems provide significant challenges. Furthermore, the need to develop a more complete and integrated picture of the complex and interdependent environment has been discussed. In the United Kingdom, the potential benefits of this approach will be investigated through UK Environmental Prediction activities.

These developments and others offer potential for further significant advances to the toolkit over the next decade. Equally, they set new scientific and technical challenges for the environmental science modelling and observational communities across disciplines to meet in order to derive greater value for users.

References

Arribas A, Cusack S, Glover M, Maidens A, Peterson K, Gordon M, *et al.* 2011. The GloSea4 ensemble prediction system for seasonal forecasting. *Mon. Weather Rev.* **139**: 1891–1910.

Ballard SP, Li Z, Simonin D, Buttery H, Charlton-Perez C, Gaussiat N, *et al.* 2012. Use of radar data in NWP-based nowcasting. In *Weather Radar and Hydrology*, Moore RJ, Cole SJ, Illingworth AJ (eds), Vol. **351**. IAHS Publications: Wallingford; 336–341.

Bowler NE, Arribas A, Mylne KR, Robertson KB, Beare SE. 2008. The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.* **134**: 703–722.

Bowler N, Pierce CE, Seed A. 2006. STEPS: a probabilistic precipitation forecast scheme which merges an extrapolation nowcast with downscaled NWP. *Q. J. R. Meteorol. Soc.* **620A**: 2107–2125.
Bell VA, Kay AL, Jones RG, Moore RJ. 2007. Development of a high resolution grid-based river flow model for use with regional climate model output. *Hydrol. Earth Syst. Sci.* **11**(1): 532–549.
Brown A, Milton S, Cullen M, Golding B, Mitchell J, Shelly A. 2012. Unified modelling and prediction of weather and climate, a 25-year journey. *Bull. Am. Meteorol. Soc.* **93**: 1865–1877.
Chen S, Campbell TJ, Jin H, Gaberšek S, Hodur RM, Martin P. 2010. Effect of two-way air-sea coupling in high and low wind speed regimes. *Mon. Weather Rev.* **138**: 3579–3602.
Clayton AM, Lorenc AC, Barker DM. 2013. Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. *Q. J. R. Meteorol. Soc.* **139**: 1445–1461.
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 modeling of the atmosphere. *Q. J. R. Meteorol. Soc.* **131**: 1759–1780.
Deacu D, Fortin V, Klyszejko E. 2012. Predicting the net basin supply to the Great Lakes with a hydrometeorological model. *J. Hydrometeorol.* **13**: 1739–1759.
Dow G, Macpherson B. 2013. Benefit of convective-scale data assimilation and observing systems in the UK models. Met Office Technical Report 585, Met Office: Exeter.
Flowerdew J, Horsburgh K, Wilson C, Mylne K. 2010. Development and evaluation of an ensemble forecasting system for coastal storm surges. *Q. J. R. Meteorol. Soc.* **136**: 1444–1456.
Flowerdew J, Mylne K, Jones C, Titley H. 2012. Extending the forecast range of the UK storm surge ensemble. *Q. J. R. Meteorol. Soc.* **139**: 184–197.
Gilleland E, Ahijevych D, Brown BG, Casati B, Ebert EE. 2009. Intercomparison of spatial forecast verification methods. *Weather Forecast.* **24**(5): 1416–1430.
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, *et al.* 2014. Forecasting capabilities for the London 2012 Olympics. *Bull. Am. Meteorol. Soc*. **95**: 883–896.
Goudenhoofdt E, DeLobbe L. 2009. Evaluation of radar-gauge merging methods for quantitative precipitation extimates. *Hydrol. Earth Syst. Sci.* **13**: 195–203.
Habib E, Lee G, Kim D, Ciach GJ. 2013. Ground-based direct measurement. In *Rainfall: State of the Science*, Testik FY, Gebremichael M (eds). American Geophysical Union: Washington DC.
Harrison DL, Norman K, Pierce C, Gaussiat N. 2012. Radar products for hydrological applications. *Water Manage.* **165**: 89–103.
Horsburgh KJ, Wilson C. 2007. Tide–surge interaction and its role in the distribution of surge residuals in the North Sea. *J. Geophys. Res. - Oceans* **112**: C08003.
Huntingford C, Marsh T, Scaife A, Kendon E, Hannaford J, Kay A, et al. 2014. Potential influences in the United Kingdom's floods of winter 2013–14. *Nat. Clim. Change* **4**: 769–777.
Jolliffe IT, Stephenson DB. 2011. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley: New York, NY.
Joo S, Eyre J, Marriot R. 2013. The impact of Metop and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method. *Mon. Weather Rev.* **141**: 3331–3342.
Kendon M, McCarthy M. 2015. The UK's wet and stormy winter of 2013/2014. *Weather*, DOI: 10.1002/wea.2465 (in press).
Kendon EJ, Roberts NM, Fowler HJ, Roberts MJ, Chan SC, Senior CA. 2014. Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nat. Clim. Change* **4**: 570–576.
Lazo JK, Morss RE, Demuth JL. 2009. 300 Billion served. *Bull. Am. Meteorol. Soc.* **90**: 785–798.
MacLachlan C, Arribas A, Peterson D, Maidens A, Fereday D, Scaife AA, *et al.* 2014. Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Q. J. R. Meteorol. Soc.* DOI: 10.1002/qj.2396.
Met Office, CEH. 2014. The recent storms and floods in the UK. http://www.metoffice.gov.uk/research/news/2014/uk-storms-and-floods (accessed 25 November 2014)
Mittermaier MP. 2006. Using an intensity-scale technique to assess the added benefit of high-resolution model precipitation forecasts. *Atmos. Sci. Lett.* **7**(2): 35–42.
Mittermaier MP. 2007. Improving short-range high-resolution model precipitation forecast skill using time-lagged ensembles. *Q. J. R. Meteorol. Soc.* **133**: 1487–1500.
Mittermaier MP. 2008. Introducing uncertainty of radar-rainfall estimates to the verification of mesoscale model precipitation forecasts. *Nat. Hazards Earth Syst. Sci.* **8**: 445–460.
Mittermaier MP. 2014. A strategy for verifying near-convection-resolving forecasts at observing sites. *Weather Forecast.* **29**(2): 185–204.
Mittermaier MP, Roberts N. 2010. Inter-comparison of spatial forecast verification methods: identifying skillful spatial scales using the fractions skill score. *Weather Forecast.* **25**: 343–354.
Mittermaier MP, Roberts N, Thompson SA. 2013. A long-term assessment of precipitation forecast skill using the fractions skill score. *Meteorol. Appl.* **20**: 176–186.
Muchan K, Lewis M, Hannaford J, Parry S. 2015. The UK winter of 2013/2014 – hydrological and hydrogeological responses and impacts. *Weather*, DOI: 10.1002/wea.2469 (in press).
Murphy AH. 1993. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather Forecast.* **8**: 281–293.
NCIC (National Climate Information Centre). 2013. Winter storms, December 2013 to January 2014. http://www.metoffice.gov.uk/climate/uk/interesting/2013-decwind (accessed 25 November 2014)
NCIC (National Climate Information Centre). 2014. Winter storms, January to February 2014. http://www.metoffice.gov.uk/climate/uk/interesting/2014-janwind (accessed 25 November 2014)
Neal RA, Boyle P, Grahame N, Mylne K, Sharpe M. 2013. Ensemble based first guess support towards a risk-based severe weather warning service. *Meteorol. Appl.* **21**: 563–577, DOI: 10.1002/met.1377.
Palmer TN. 2014. Record-breaking winters and global climate change. *Science* **344**: 803–804.
Pellerin P, Ritchie H, Saucier FJ, Roy F, Desjardins S, Valin M, et al. 2004. Impact of a two-way coupling between an atmospheric and an ocean-ice model over the Gulf of St Lawrence. *Mon. Weather Rev.* **132**: 1379–1398.
Perry MC, Hollis DM. 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. *Int. J. Climatol.* **25**: 1041–1054.
Pierce C, Seed A, Ballard SP, Simonin D, Li Z. 2012. Nowcasting. In *Doppler Radar*, Bech J, Chau JL (eds). InTech: Rijeka, Croatia.
Pietroniro A, Fortin V, Kouwen N, Neal C, Turcotte R, Davison B, et al. 2007. Development of the MESH modelling system for hydrological ensemble forecasting of the Laurentian Great Lakes at the regional scale. *Hydrol. Earth Syst. Sci.* **11**: 1279–1294.
Pullen J, Doyle J, Signell RP. 2006. Two-way air-sea coupling: a study of the Adriatic. *Mon. Weather Rev.* **134**: 1465–1483.
Pullen J, Holt T, Blumberg A, Bornstein R. 2007. Atmospheric response to local upwelling in the vicinity of New York – New Jersey Harbor. *J. Appl. Meteorol.* **46**: 1031–1052.
Renault L, Chiggiaro J, Warner JC, Gomez M, Vizoso G, Tintoré J. 2012. Coupled atmosphere–ocean-wave simulations of a storm event over the Gulf of Lion and Balearic Sea. *J. Geophys. Res.* **117**: C09019.
Roberts NM. 2003a. The impact of a change to the use of the convection scheme in high-resolution simulations of convective events. Met Office Forecasting Research Technical Report 407, Met Office: Exeter; 30.
Roberts NM. 2003b. Precipitation diagnostics for a high resolution forecasting system. Met Office Forecasting Research Technical Report 423, Met Office: Exeter.
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, Lean HW. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.* **136**: 78–96.
Sandery PA, Brassington GB, Craig A, Pugh T. 2010. Impacts of ocean–atmosphere coupling on tropical cyclone intensity change and ocean prediction in the Australian region. *Mon. Weather Rev.* **138**: 2074–2091.
Scaife AA, Arribas A, Blockley E, Brookshaw A, Clark RT, Dunstone N, et al. 2014a. Skilful predictions of European and North American Winters. *Geophys. Res. Lett.* **41**: 2514–2519.
Scaife AA, Athanassiadou M, Andrews MB, Arribas A, Baldwin MP, Dunstone N, et al. 2014b. Predictability of the quasi-biennial oscillation and its northern winter teleconnection on seasonal to decadal timescales. *Geophys. Res. Lett.* **41**: 1752–1758.
Scaife AA, Copsey D, Gordon C, Harris C, Hinton T, Keeley SJ, et al. 2011. Improved Atlantic blocking in a climate model. *Geophys. Res. Lett.* **38**: L23703.

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.

Shapiro M, Shukla J, Brunet G, Nobre C, Beland M, Dole R, et al. 2010. An Earth-System prediction initiative for the 21st Century. *Bull. Am. Meteorol. Soc.* **91**: 1377–1388.

Shelly A, Xavier P, Copsey D, Johns T, Rodriguez JM, Milton S, *et al.* 2014. Coupled versus uncoupled hindcast simulations of the Madden-Julian Oscillation in the Year of Tropical Convection. *Geophys. Res. Lett.* **41**: 5670–5677.

Sibley A, Cox D, Titley H. 2015. Coastal flooding in England and Wales from Atlantic and North Sea storms during the 2013/2014 winter. *Weather*, DOI: 10.1002/wea.2471 (in press).

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, DOI: 10.1002/qj.2298.

Smith EA, Asrar G, Furuhama Y, Ginati A, Mugnai A, Nakamura K, *et al.* 2007. International global precipitation measurement (GPM) program and mission: an overview. In *Measuring Precipitation from Space*. The Netherlands: Springer; 611–653.

Smith GC, Roy F, Brasnett B. 2013. Evaluation of an operational ice-ocean analysis and forecasting system for the Gulf of St Lawrence. *Q. J. R. Meteorol. Soc.* **139**: 419–433.

Sun J, Xue M, Wilson JW, Zawadzki I, Ballard SP, Onvlee-Hooimeyer J, *et al.* 2014. Use of NWP for nowcasting convective precipitation. *Bull. Am. Meteorol. Soc.* **95**: 409–426.

Tang Y, Lean HW, Bornemann J. 2012. The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorol. Appl.* **20**: 417–426.

Tennant W, Beare S. 2014. New schemes to perturb sea-surface temperature and soil moisture content in MOGREPS. *Q. J. R. Meteorol. Soc.* **140**: 1150–1160.

Theis SE, Hense A, Damrath U. 2005. Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. *Meteorol. Appl.* **12**: 257–268.

UK Government. 2013. National Risk Register for Civil Emergencies. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/211867/NationalRiskRegister2013_amended.pdf (accessed 25 November 2014)

UK Government. 2014. UK floods 2014: government response and recovery. https://www.gov.uk/government/news/uk-floods-2014-government-response (accessed 25 November 2014)

Warner JC, Armstrong B, He R, Zambon JB. 2010. Development of a coupled ocean–atmosphere-wave-sediment transport (COAWST) modelling system. *Ocean Model.* **35**: 230–244.

Wilson JW, Crook NA, Mueller CK, Sun J, Dixon M. 1998. Nowcasting thunderstorms: a status report. *Bull. Am. Meteorol. Soc.* **79**: 2079–2099.

Wood N, Staniforth A, White A, Allen T, Diamantakis M, Gross M, *et al.* 2013. An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Q. J. R. Meteorol. Soc.* **140**: 1505–1520, DOI: 10.1002/qj.2235.