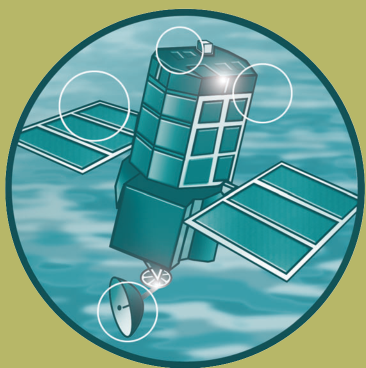
Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme

Modelling Extreme Rainfall Events

R&D Technical Report FD2210/TR

Environment Agency

defra
Department for Environment Food and Rural Affairs

Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme

Modelling Extreme Rainfall Events

R&D Technical Report FD2210/TR

Produced: May 2008

Author(s): Nigel Roberts

Statement of use
This report documents the outcome of Project FD2210
"Modelling Extreme Rainfall Events" concerned with assessing the ability of a storm-resolving Numerical Weather Prediction (NWP) model to predict extreme rainfall events. Findings will contribute to the development and understanding of new high-resolution NWP forecast systems and products for improved flood warning.

Dissemination status
Internal: Released Internally
External: Released to Public Domain.

Keywords:
High-resolution Numerical Weather Prediction (NWP) model, extreme rainfall, forecasting, predictability, case studies, warning

Research contractor:
Met Office, FitzRoy Road, Exeter, EX1 3PB
Project Initiator: Brian Golding Email: brian.golding@metoffice.gov.uk

Defra project officer:
Bob Hatton

Publishing organisation
Department for Environment, Food and Rural Affairs
Flood Management Division,
Ergon House,
Horseferry Road
London SW1P 2AL

Tel: 020 7238 3000 Fax: 020 7238 6187

www.defra.gov.uk/environ/fcd

Published by the Department for Environment, Food and Rural Affairs (Sept 2008)

Executive summary

This is the third and final report to be delivered from the Modelling Extreme Rainfall Events (MERE) project. The project began in November 2005 and finished in February 2008.

The objective was to investigate the ability of a storm scale configuration of the Met Office Numerical Weather Prediction (NWP) model (the Unified Model (UM)) to predict extreme rainfall events up to 18-24 hours ahead and to determine what it is about the meteorology of these situations that the model must capture in order to produce useful predictions for flood warning.

The project follows on from the Storm Scale Modelling project in which it was found that the UM with a grid spacing of ~1 km has the potential to deliver significantly improved forecasts of convective rainfall events.

Operational implementation of a 'storm-resolving' version of the UM is now considered to be of paramount importance for the future delivery of improved weather forecasts (including heavy rain) in the UK. The UM is currently run operationally with a grid-spacing of 4 km over the UK. A 1-1.5 km version is planned for 2009/10 when computer resources are enhanced.

The project was split into three stages:

Stage 1: To identify five extreme rainfall events and assess the ability of a storm-resolving NWP model to predict these events. The report from stage 1 was completed in March 2006.

Stage 2: To study two of the cases chosen in stage 1 in considerably more depth. The objectives were (1) to achieve the best possible high-resolution forecasts and in doing so gain an understanding of the meteorological mechanisms involved; (2) to determine what a storm-resolving model needs to get correct and where there are deficiencies in the model, and (3) to gain insight into the predictability of these types of events. The report from stage 2 was completed in March 2007.

Stage 3: To perform an in-depth analysis on another extreme case that differed meteorologically from those previously examined. To perform model sensitivity studies on both the new case and the two cases from stage 2. To investigate whether the storm-resolving model was capable of giving better forecasts of any new extreme events identified during the course of the project.

Key Findings

1. There are no differences between the physical and dynamical processes that lead to 'extreme' events compared to other heavy rainfall events. It is the coincidence and interaction of those processes that leads to extreme events and the difficulty in their prediction.
2. A storm-resolving NWP model is capable of providing useful forecasts of 'extreme' rainfall events. The use of a storm-resolving model has the potential to greatly improve on our current ability to predict such events provided that it is understood that the output must include information about forecast uncertainty (e.g. probabilities).

3. The accuracy of the forecasts may vary considerably from case to case and depends crucially on getting all the necessary meteorological components correct.
4. For many events is vital to represent accurately the larger-scale disturbances in the flow that may originate from outside the high-resolution domain; as well as any local effects.
5. A high-resolution grid (~1 km grid spacing) is absolutely essential to be able to represent the dynamics of more localised thunderstorms and many of the important local pre-cursors to the triggering of storms.
6. Some convective situations are inherently more predictable than others, and that predictability is strongly linked to the meteorology of the situation. A classification into three types of storm has been made. Knowledge of the likely type of storm can provide information about its predictability.
7. This work has lead to a greater understanding of both the meteorological processes that lead to extreme rainfall events and the strengths and weaknesses in the model in predicting them.

Recommendations

The development and operational implementation of a storm-resolving model (grid spacing ~ 1 - 2 km) should continue. It is likely to be the best route towards a significant improvement in our ability to provide warnings of severe convective storms.

There will always be uncertainty in forecasts of extreme rainfall events and the more precision we expect the more uncertainty there will be. It is essential that probabilistic outputs for users are developed and that users understand what they mean.

The model is capable of representing extreme storms, but getting the positioning correct presents a big challenge. Research and development into new data assimilation methods (for improving the initial state of forecasts) on the high resolution grid and the use of new types of observations in data assimilation is, and must continue to, play a vital role in the development of a storm-resolving model.

The uncertainty in a single 'deterministic' forecast is difficult to quantify and this can limit the usefulness of probabilistic outputs based on one forecast. We should now move towards developing a storm-resolving ensemble prediction system and ensemble-based probabilistic products. Initially, such a system would embed the storm-resolving model in a coarser-resolution ensemble to account for uncertainty in the larger-scale dynamics.

Further research is required to determine whether an estimate of the predictability can be deduced in advance from the meteorological conditions. Research into the use of 'time-lag' ensembles (combining the most recent forecast with older forecasts) as a possible means of determining predictability should also be undertaken.

High-resolution model output should be used in hydrological models for flood warning, particularly for orographically enhanced rainfall situations such as the Carlisle floods in 2005. For convective situations, a more probabilistic approach is necessary and needs to be developed.

There are still some outstanding issues to do with the way a storm-resolving model represents convective cells. Further work is needed in the areas of cloud microphysics and sub-grid-scale turbulence.

Contents

Figures

Section 2

Section 3

General regions of possible ascent and descent within the troposphere are depicted by the dashed circles. (b-f) pseudo Water Vapour imagery extracted from relative humidity fields at 300 to 600 hPa from the 12-km forecasts starting at 18 UTC 18th, 00, 03, 06 and 09 UTC 19th June as comparison against the imagery. The blue line is the same as the white line in (a). The dashed white lines mark the moist/dry air transition in the pseudo imagery.

Section 4

Figure 4-1. Rainfall accumulations over the period 13 to 18 UTC, 3rd August 2004 from (a) radar, (b) a 1 km forecast from 06 UTC with 76 vertical levels and (c) a 1 km forecast from 06 UTC with 38 vertical levels. The orange line in (b) and (c) enclosed the totals above 30 mm in (a) (over and to the west of London).

Figure 4-2. Rainfall accumulations over the period 13 to 18 UTC, 3rd August 2004 from (top left) radar, and a collection of 1 km forecasts from 06 UTC with different uses of the convection parametrization scheme (refer to text). The orange line in all but the top left panel encloses the totals above 30 mm from radar (over and to the west of London).

Figure 4-3. Graph of forecast skill using the Fractions skill Score (FSS) against spatial scale for the rainfall accumulation period shown in Figures 4-1 and 4-2 for (a) the 1 km forecast from 06 UTC compared with radar (red line), the 1 km forecast with weak convection scheme (blue line) compared with radar and the comparison between the two forecasts (green line). (b) the 1 km forecast from 06 UTC compared with radar (red line), the 1 km forecast with an unrestricted convection scheme (blue line) compared with radar and the comparison between the two forecasts (green line). The better the fit between forecast and radar or two forecasts the higher the value of FSS. The upper dashed line shows the minimum FSS value for a reasonable fit. The lower dashed line shows the FSS for a comparison with a randomly generated forecast.

Figure 4-4. Rainfall accumulations over the period 13 to 17 UTC 19th June 2005 from (a) radar and 1 km forecasts from 06 UTC (b) reference forecast, (c) forecast with 38 vertical levels instead of 76, (d) forecast with radiation on slopes included (see text), and (e) forecast with the convection scheme included as used in the UK 4 km model.

Figure 4-5. Graph of forecast skill using the Fractions skill Score (FSS) against spatial scale for the rainfall accumulation period shown in Figure 4-4 for (a) the 1 km forecast from 06 UTC with 76 levels compared with radar (red line), the 1 km forecast with 38 levels compared with radar (blue line) and the comparison between the two forecasts (green line). The better the fit between forecast and radar or two forecasts the higher the value of FSS. The upper dashed line shows the minimum FSS value for a reasonable fit. The lower dashed line shows the FSS for a comparison with a randomly generated forecast. The other dashed lines are described in the text.

Figure 4-6. Graph of forecast skill using the Fractions skill Score (FSS) against spatial scale for the rainfall accumulation period shown in Figure 4-4 for (a) the 1 km forecast from 06 UTC with 76 levels compared with radar (red line), the 1 km forecast with radiation on slopes (see text) with radar (blue line), the comparison between the two forecasts (green line), and the comparison between the forecasts in Figure 4-5 (dashed green line). The better the fit between forecast and radar or two forecasts the higher the value of FSS. The upper dashed line shows the minimum FSS value for a reasonable fit. The lower dashed line shows the FSS for a comparison with a randomly generated forecast. The other dashed lines are described in the text.

Section 5

Section 6

circle is described in the text. The smaller superimposed panel shows the area covered by the frames.

Figure 6-13. Shaded blobs show areas of rain exceeding 8 mm/hr every 10 minutes taken from the 1 km forecast starting at 06 UTC. Darker shades enclose rates >16mm/hr. Each time has been separated from the next by the length of the arrow labelled ‘Displacement imposed from frame to frame’. Further discussion in the text.

Figure 6-14. Diagnostics from the 1 km model forecast starting from 06 UTC for 18 UTC 4th July 2006, (a) Vertical velocity at 900 hPa (shaded) and 700 hPa (contoured, with blue=descent, red=ascent), (b) θ_w at 975hPa with vertical velocity at 700 hPa (as panel (a)), divergence at 900 hPa, (d) a composite picture.

Figure 6-15. Cross section along the line x2-y2 in Figure 6-14. Diagnostics from the 1 km model forecast starting from 06 UTC for 18 UTC 4th July 2006, (a) Relative humidity, (b) θ_w, (c) along section wind (shaded), vertical velocity (contoured every 2 m/s) and arrows superimposed to depict key flows, (d) a composite picture.

Section 7

Figure 7-1. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) with 76 vertical levels, and (b) with 38 vertical levels.

Figure 7-2. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar and the 38 and 76 level 1 km forecasts from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

Figure 7-3. Rainfall rates at 15 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) with 76 vertical levels, and (b) with 38 vertical levels.

Figure 7-4. Tephigrams at locations along the cross-section x-y in section 6, from the 1 km model forecasts with 76 and 38 vertical levels.

Figure 7-5. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain).

Figure 7-6. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar, the reference 1 km forecast (300x300km domain) and the 1.5 km forecast (450x450 km domain) from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

Figure 7-7. Rainfall rates at 15 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain).

Figure 7-8. Rainfall rates at 20 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain).

Figure 7-9. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (300x300 km domain).

threshold the 95th percentile value taken from all points in the domain (including zeros) (see appendix in section 4 for more information about the FSS). The dashed lines discriminate between a good agreement between forecasts (FSS>0.8), a reasonable agreement (0.5<FSS<0.8) and a poor agreement (FSS<0.5).

Section 8

Figure 8-1. Rainfall accumulation over the period 16 to 18 UTC 10th May 2006, from (a & b) radar, (c) 12 km forecast from 12 UTC, (d) 1 km forecast from 12 UTC.

Figure 8-2. Graph showing the scale at which a target or useful level of skill was reached by the 12 km forecast (blue line)and 1 km forecast (red line) for 2-hourly rainfall accumulations (lower is better). See appendix 1 for more information. The light shaded bars show the mean 2-hour rainfall accumulations (over small domain in Figure 8-1) and the grey bars the peak rainfall rates from radar (5x5km grid squares).

Figure 8-3. Rainfall accumulation over the period 15 to 17 UTC 13th August 2006, from (a & b) radar, (c) 12 km forecast from 12 UTC, (d) 1 km forecast from 12 UTC.

Figure 8-4. Graph showing the scale at which a target or useful level of skill was reached by the 12 km forecast (blue line)and 1 km forecast (red line) for 2-hourly rainfall accumulations (lower is better). See appendix 1 for more information. The light shaded bars show the mean 2-hour rainfall accumulations (over small domain in Figure 8-1) and the grey bars the peak rainfall rates from radar (5x5km grid squares).

Figure 8-5. (a) Rainfall accumulations from gauge measurements for the 24th and 25th June 2007 courtesy of the National Climate Information Centre (NCIC). (b) Rainfall accumulations from radar over the period 04 to 21 UTC 25th June 2007. The Ingham radar is marked with a star. The red boxes enclose the Hull and Sheffield urban areas.

Figure 8-6. Rainfall accumulations over the period 04 to 21 UTC 25th June 2007 from the North Atlantic European (NAE) 12 km operational forecasts starting from 00, 06, 12 and 18 UTC on the 24th and from 00 UTC on the 25th. The red boxes enclose the Hull and Sheffield urban areas.

Figure 8-7. Rainfall accumulations over the period 04 to 21 UTC 25th June 2007 from simulations starting at 15 UTC on the 24th using (a) the 1.5 km model and (b) the 4 km model. The red boxes enclose the Hull and Sheffield urban areas.

Figure 8-8. Rainfall accumulations over the period 04 to 21 UTC 25th June 2007 from simulations starting at 21 UTC on the 24th using (a) the 1.5 km model and (b) the 4 km model. The red boxes enclose the Hull and Sheffield urban areas.

Figure 8-9. Rainfall accumulations from gauge measurements for the 19th and 20th July 2007 courtesy of the National Climate Information Centre (NCIC).

Figure 8-10. Rainfall rates at 09 and 12 UTC 20th July 2007 from the operational UK 4 km forecast starting at 21 UTC on the 19th.

Figure 8-11. Rainfall accumulations over the period 00 to 12 UTC 20th July 2007 from (a) radar, and (b) a 1.5 km forecast starting at 00 UTC on the 20th. Courtesy of Humphrey Lean, Joint Centre for Mesoscale Meteorology (JCMM), Met Office.

Figure 8-12. Rainfall accumulations over the period 03 to 15 UTC 20th July 2007 from (a) radar, and (b) a 1.5 km forecast starting at 03 UTC on the 20th. Courtesy of Humphrey Lean, Joint Centre for Mesoscale Meteorology (JCMM), Met Office.

Figure 8-13. Rainfall accumulations over the period 06 to 18 UTC 20th July 2007 from (a) radar, and (b) a 1.5 km forecast starting at 06 UTC on the 20th. Courtesy of Humphrey Lean, Joint Centre for Mesoscale Meteorology (JCMM), Met Office.

Section 9

Figure 9-1. Probability that rainfall accumulations will exceed 50 mm in the period 13 to 19 UTC 3rd August 2004 from a 1 km forecast starting from 09 UTC. (a) Using a neighbourhood of 30 km, (b) Using a neighbourhood of 60 km. The crosses mark where 50 mm was exceeded from radar measurements.

Figure 9-2. Probability that rainfall accumulations will exceed 25mm in the period 17 to 19 UTC 4th July 2006 from a 1 km forecast starting from 06 UTC. (a) Using a neighbourhood of 30 km, (b) Using a neighbourhood of 60 km. The crosses mark where 25 mm was exceeded from radar measurements.

Figure 9-3. Probability that rainfall accumulations will exceed 75mm in the period 06 to 18 UTC 25th June 2007 from a 1 km forecast starting from 03 UTC. (a) Using a neighbourhood of 30 km, (b) Using a neighbourhood of 60 km. The black line encloses where rainfall exceeded 75 mm (from gauges) over the 24th and 25th.

Figure 9-4 Probability that rainfall accumulations will exceed 75mm, using a 60 km neighbourhood, in the period 06 to 15 UTC 20th July 2007 from 1.5 km forecasts starting from (a) 03 UTC, (b) 06 UTC, and (c) the two forecasts combined with equal weighting. Black contour encloses the probabilities >6% when the same processing was applied to radar data.

Figure 9-5. Diagnostics from the 1.5 km model forecast starting from 03 UTC 20th July 2007 for the period 06 to 15 UTC. (a) The maximum possible mean rainfall amount over a $24km^2$ area within each EA warning area given a forecast uncertainty of 60 km. (b) The probability of exceeding a mean accumulation of more than 50 mm over a $24km^2$ area within each EA warning area given a forecast uncertainty of 60 km.

Appendix

Figure A1. A schematic comparison between forecast and radar (see text)

Tables

Table 3-1. Note that the 1-km forecasts are labelled as starting at the same time as the 12-km forecasts. This is because they share the same analysis time even though the 1-km forecasts begin 1-hour later (see Figure 5).

Table 3-2. Note that the 1-km forecasts are labelled as starting at the same time as the 12-km forecasts. This is because they share the same analysis time even though the 1-km forecasts begin 1-hour later.

Appendices

Appendix 1: The generation of fractions and computation of the Fractions Skill Score.

1. Introduction

This is the third and final report to be delivered from the Modelling Extreme Rainfall Events (MERE) project. The project began in November 2005 and finished in February 2008. It was half funded by the Flood Forecasting and Warning theme of the joint Defra/Environment Agency Research Programme and half funded by the Met Office core research programme.

The objective was to investigate the ability of a storm scale configuration (~1 km grid spacing) of the Met Office Numerical Weather Prediction (NWP) model (the Unified Model (UM)) to predict extreme rainfall events up to 18-24 hours ahead and to determine what it is about the meteorology of these situations that the model must capture in order to produce useful predictions for flood warning. An extreme event is defined as having a return period of ~100 years at any particular location.

The project follows on from the Storm Scale Modelling project (Roberts 2005) in which it was found that the UM with a grid spacing of ~1km has the potential to deliver significantly improved forecasts of convective rainfall events.

To put this work in context; the operational implementation of a ‘storm-resolving’ version of the UM is now regarded as being of paramount importance in the future delivery of improved weather forecasts (including heavy rain) for the UK. The UM is currently run operationally with a grid-spacing of 4 km on a domain covering the UK. A 1-1.5 km version is planned for 2009/10 when computer resources are enhanced.

1.1 Objectives

The following objectives were set out in the project initiation document.

Objectives

1. Simulate a selection of extreme (~1:100yr) rainfall events using the convective scale NWP model

2. Identify the causes of shortcomings in each simulation and investigate possible improvements

3. Assess the ability of the model to reproduce features identified in the Extreme Event Recognition project (Hand *et al* 2004) as having contributed to the extreme nature of each storm

4. Assess the ability of the model to reproduce the observed precipitation in each case

5. Synthesise the results, draw conclusions on the predictability of extreme rainfall events using a convective scale model, and make recommendations on further work that could increase their predictability.

1.2 Stages

In order to meet the objectives above the project was split into three stages.

1.2.1 Stage 1

The tasks for stage 1 were to identify five extreme rainfall events and then assess the ability of a storm-resolving NWP model to predict those events. This was a first look at the capability of the model and was a route into more detailed investigation later on.

The five cases were:

1. Case A. 8th July 2004. Very high rainfall totals at the village of Wittering, Cambridgeshire from a localised thunderstorm that followed a large amount of frontal rain the previous night.

2. Case B. 3rd August 2004. Flash flooding in parts of northwest London as a result of intense thunderstorms during the afternoon.

3. Case C.16th August 2004. Very destructive flash flood at the village of Boscastle, North Cornwall from persistent localised thunderstorms during the afternoon.

4. Case D. 7-8th January 2005. Flooding in the city of Carlisle, following more than a day of persistent rain over the Lake District and northwest Pennines.

5. Case E.19th June 2005. Flash floods at the villages of Hawnby and Helmsley in North Yorkshire because of torrential rain from a very intense thunderstorm.

The findings from stage 1 were presented in the stage 1 report (Roberts 2006). A cut-down version of those results is presented in section 2 of this report.

1.2.2 Stage 2

In stage 2, two of the cases from stage 1 were to be examined in considerably more depth.
The objectives were

(1) To achieve the best possible high-resolution forecasts and in doing so gain an understanding of the meteorological mechanisms involved.
(2) To determine what a storm-resolving model needed to get correct in these situations and where there are deficiencies in the model.
(3) To gain insight into the predictability of these types of events.

The two cases used were

1. Case B. Flooding in northwest London following intense convection on the 3rd August 2004.
2. Case E. Flooding in the Hawnby area of North Yorkshire following severe convection on the 19th June 2005. (Case 5 above)

The findings from stage 2 were presented in the stage 2 report (Roberts 2007). A cut-down version of those results is presented in section 3 of this report.

1.2.3 Stage 3

In stage 3 another case was to be examined in detail and model sensitivity experiments were to be performed on the new case and the two cases from stage 2. The choice of the new case was important as it had to be different in nature from those already studied (but still extreme) to provide a new perspective on the model capability and the predictability of extreme events.

Case F: The case chosen was a flash flood in the village of Albrighton in the West Midlands on the 4th July 2006 that occurred because of a localised and very intense thunderstorm.

The findings from this case are presented in sections 5 to 7. The model sensitivity studies for the stage 2 cases are presented in section 4.

1.3 Further cases

During the course of the project there were some other extreme or high-profile cases. Some of them are also documented briefly in this report.

Case G: Large thunderstorm complex on the 10th May 2006 over Southern England that brought flooding to South Wales later in the evening.
Case H: Flash flooding from thunderstorms in Surry (west of London) 13th August 2006.
Case I: Severe flooding over northern England, particularly Hull and Sheffield, from a quasi-stationary frontal system on the 25th June 2007
Case J: Severe flooding over the West Midlands and Southern England from widespread heavy convective rain on the 20th July 2007

The events in the summer of 2007 were particularly notable and led to the independent review by Sir Michael Pitt 'Learning lessons from the 2007 floods' December 2007.

1.4 Other projects

This project does not stand alone. Also contained in section 10 is a description of the relationships between recent and ongoing research projects directed at understanding and predicting flood-producing rainfall events over the UK and Europe. The section focuses on the use of high-resolution Numerical Weather Prediction (NWP) models and projects funded from four main sources (a) The Met Office PWS customer group research programme, (b) Department for Food and Rural Affairs (Defra) in cooperation with the Environment Agency (EA) (c) Natural Environmental Research Council (NERC) and (d) European / Global funding bodies.

1.5 The model

All the model simulations were performed using the Met Office UM. The Met Office's Unified Model (UM) solves non-hydrostatic, deep-atmosphere dynamics using a semi-implicit, semi-Lagrangian numerical scheme (Cullen *et al.* 1997, Davies *et al* 2005). The model includes a comprehensive set of parametrizations, including surface exchange (Essery, *et al.* 2001), boundary layer (Lock *et al.* 2000), mixed phase cloud microphysics (Wilson and Ballard 1999 + enhancements) and convection (Gregory and Rowntree 1990 + enhancements). The model runs on a rotated latitude/longitude horizontal grid with Arakawa C staggering, and, a terrain-following hybrid-height vertical coordinate with Charney-Philips staggering. Soil moisture fields are generated offline from observations using the UM surface exchange scheme (Smith et al 2006). Data assimilation in the operational forecasts used 3DVAR (Lorenc et al 2000) or from 2006 4DVAR (Rawlins et al 2007) and latent heat nudging (Jones and Macpherson 1997).

2. Case studies in stage 1

This section gives a brief overview of the cases that were studied in stage 1 of the project and forecasts of those events using a model grid spacing of 12, 4 and 1km. Cases B and E were studied further in stage 2 and are only described briefly in this section. They are described more in section 3. Refer to the stage 1 report (Roberts 2006) for a more comprehensive account.

Forecasts were compared with radar data that had been processed and quality controlled through the Met Office Nimrod system (Golding 1998, Harrison et al 2000).

2.1 Case A 8th July 2004 Wittering flood

During the 7th July 2004 an area of low pressure moved northwards from Biscay towards the English Channel (Figure 2-1). On the northern flank of the low a slow-moving front produced a significant amount of rain over parts of the East Midlands. At Wittering (52.62N, 0.48W) 52.6mm of rain was recorded in the 12 hour period between 21UTC on the 7th and 09UTC on the 8th. After the front had moved through, heavy showers and thunderstorms developed. The heaviest storms were scattered over the same area where the front had brought the most rain the previous night. A further 51.0mm was recorded at Wittering over 2 hours from 13UTC to 15UTC on the 8th from one of these downpours. In total 107.4mm was recorded over the 19 hour period up to 15UTC on the 8th. Using the Flood Estimation Handbook (FEH) method this gave a return period of ~150 years. The reason for the high totals at some locations was the slow moving nature of the storms.

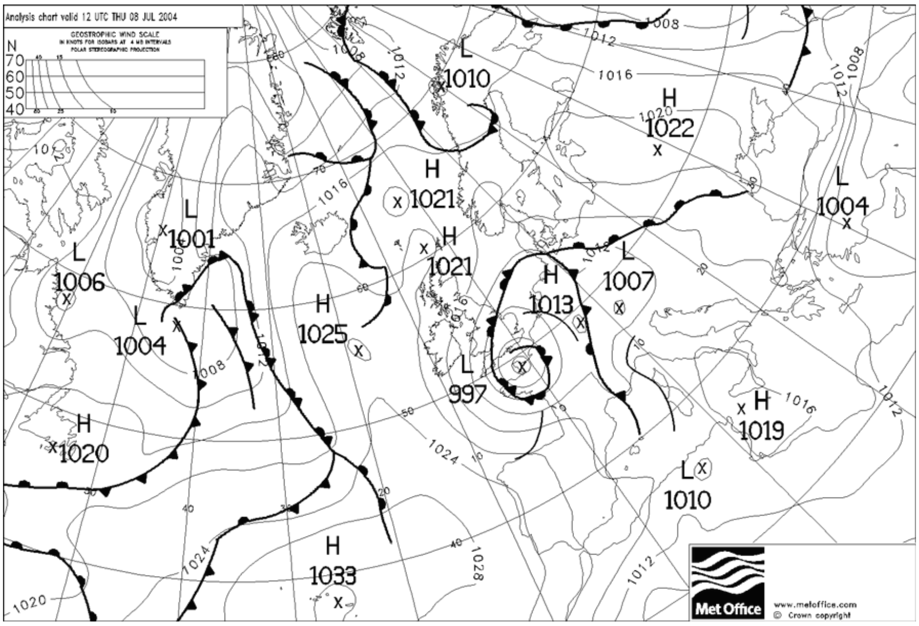


Figure 2-1. Synoptic analysis for 12 UTC 8th July 2004.

The frontal period of rain was well forecast in the 12 km mesoscale model and in the 4 km and 1 km simulations. It was the later thunderstorms that presented problems for the mesoscale model.

Figure 2-2 (right) shows rainfall accumulations from radar averaged to the same 12km grid as the operational mesoscale model. When averaged to 12km, the radar accumulations are not large (~13mm in the Wittering area) because the highest totals were concentrated over small areas. This highlights a problem with trying to predict convective storms using a 12km model. Even if the model is predicting the correct average amount of rain, it can not give an indication that there may be larger quantities on small scales.

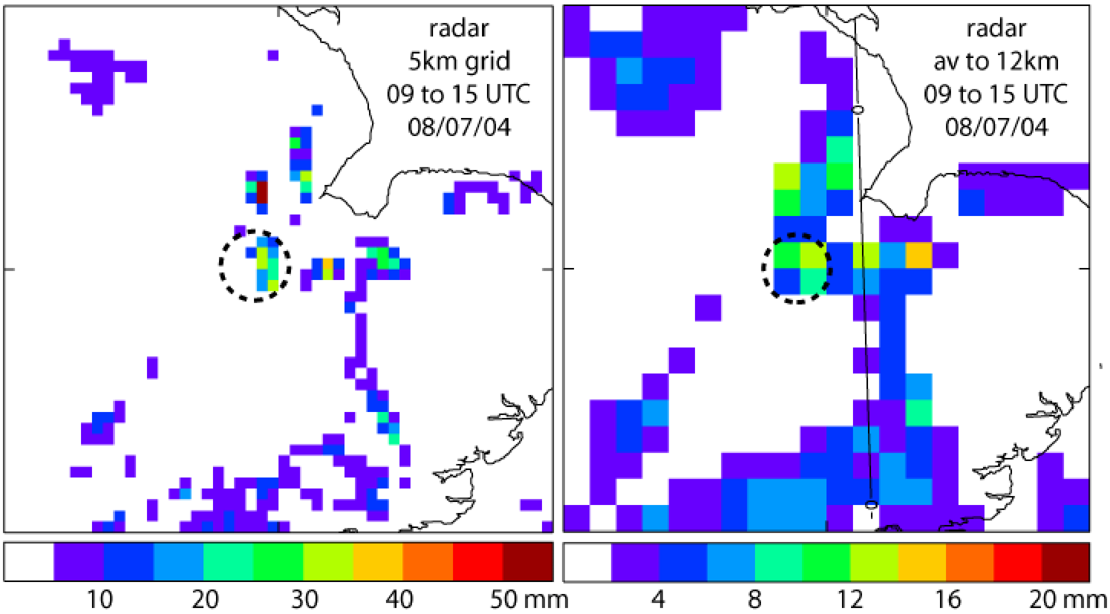


Figure 2-2. Rainfall accumulations from radar over the 6-hour period 09 UTC to 15 UTC 08/07/04 on a 5km grid (left) and averaged to 12km (right). Dashed circles are 25km diameter centred on Wittering (to within a few km)

A comparison of 12, 4 and 1 km forecasts from 06 UTC on the 8th is shown in Figure 2-3. The 12 km model did not produce enough rain. The rain that was predicted was too light and fell over too wide an area. In contrast to the 12 km forecast, the 4 and 1 km forecasts produced totals that were too high compared to radar.

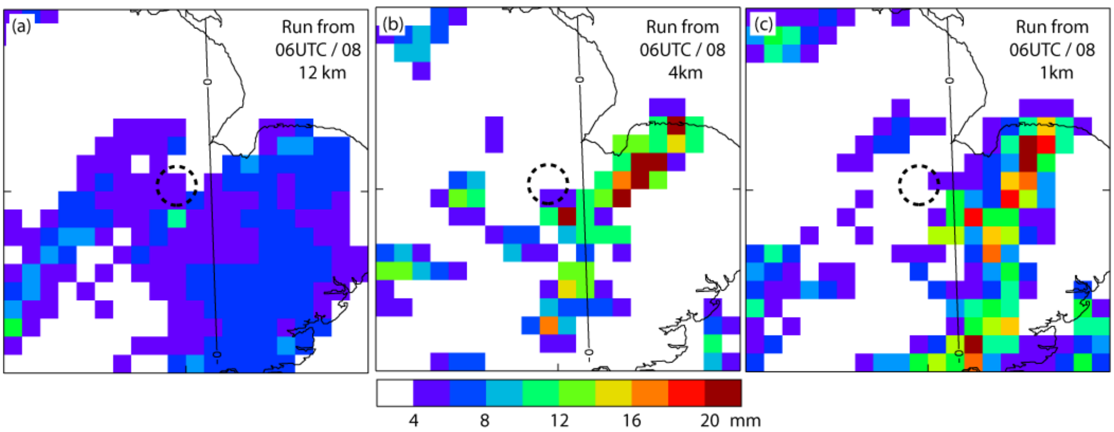


Figure 2-3. Rainfall accumulations over the 6-hour period 09 UTC 08/07/04 to 15 UTC 08/07/04, projected on to the 12 km grid, from forecasts starting at 06 UTC. (a) the operational 12 km mesoscale model, (b) the 4 km model and (c) the 1 km model. Dashed circles are 25km diameter centred on Wittering.

On this occasion the increase in resolution to 4 km and 1 km made little difference to the spatial distribution of the rain. However, the higher-resolution forecasts did give a better indication that high rainfall totals were a possibility somewhere in the locality. At the forecast resolution (rather than averaged to 12 km) the 4km forecast produced more than 50-60mm of rain over some grid squares and the 1km forecast more than 40-50mm in places. The 4km model probably had peak values that were too high. The 1km forecast, on the other hand, was more realistic but may have slightly under-predicted the highest totals. The reason why the 1km model had too much rain when averaged to 12km is because the showers were too extensive.

In this instance, the precise locations of these localised storms were very difficult to predict. The higher-resolution forecasts could provide a better warning that high totals were possible in the area.

This case was not considered for further investigation in stage 2, because it was thought unlikely that the spatial accuracy of the forecasts could be improved without the use of high-resolution data assimilation, which is outside the scope of this project.

2.2 Case B 3rd August 2004 London floods

This case was chosen for further investigation in stage 2, and will only be very briefly described here. Model results will be shown in section 3.

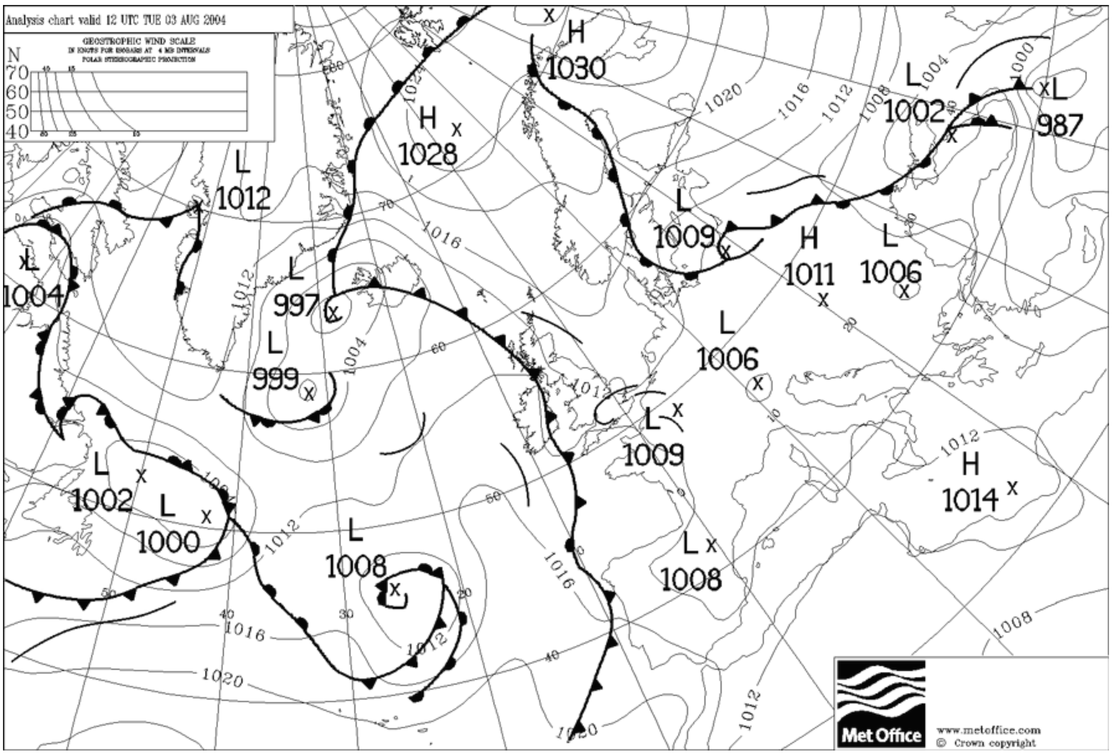


Figure 2-4. Synoptic chart for 12 UTC 3rd August 2004

The synoptic pattern is shown in Figure 2-4. Most of the UK was under the influence of hot humid air from the continent. During the middle of the afternoon heavy thunderstorms developed across southern England. These storms organised into a band and progressed north into the Midlands.

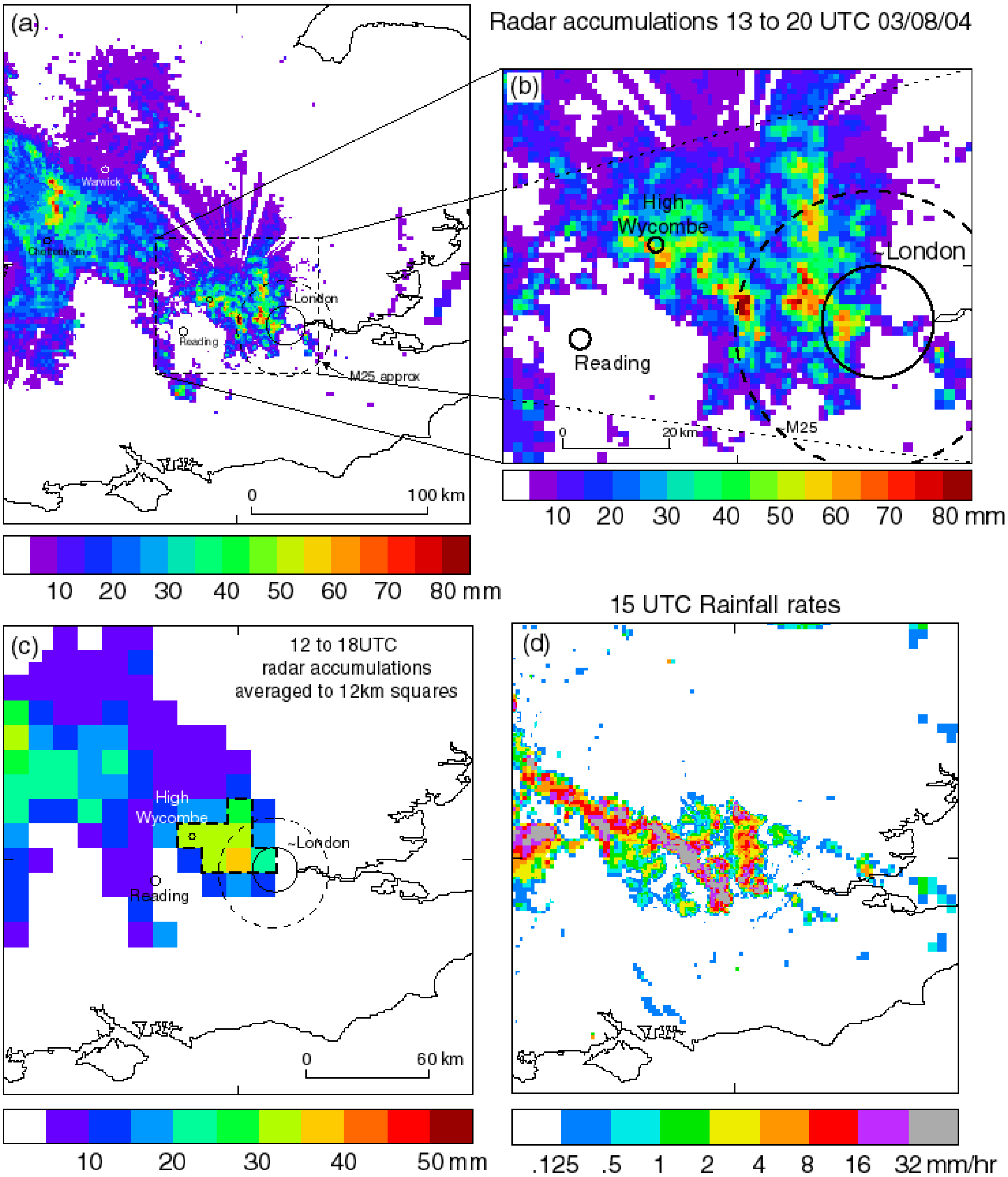


Figure 2-5. Radar accumulations for the period 13 to 20 UTC 3rd August 2004 over Southeast England and a smaller sub area (1km resolution over sub-area). (c) Radar accumulations over the period 12 to 18 UTC averaged to the operational mesoscale model 12km grid. The dashed line encloses accumulations > 20mm. (d) Rainfall rates from radar at 15 UTC on a 1km grid over the area of interest.

Rainfall accumulations from radar are displayed in Figure 2-5. More than 80mm of rain was measured in parts of northwest London. This was enough to lead to flash flooding and major disruption to commuter traffic. At High Wycombe 42.4mm was collected by a rain gauge in 38 minutes, with peak rainfall rates in excess of 200mm/hour. The value of 42.4mm in an hour gives a return period of ~64 years using the Flood Estimation Handbook (FEH) method. Given that High Wycombe was not in the area of highest accumulations observed by radar, it is likely that more extreme point values would have been obtained in parts of West London if gauge measurements had been obtained.

2.3 Case C 16th August 2004 Boscastle

Thunderstorms developed along the north coast of Cornwall (SW England) during the late morning and into the afternoon of 16th August 2004. During the afternoon a severe flash flood in the village of Boscastle inundated around 60 homes and washed 30 cars into the harbour.

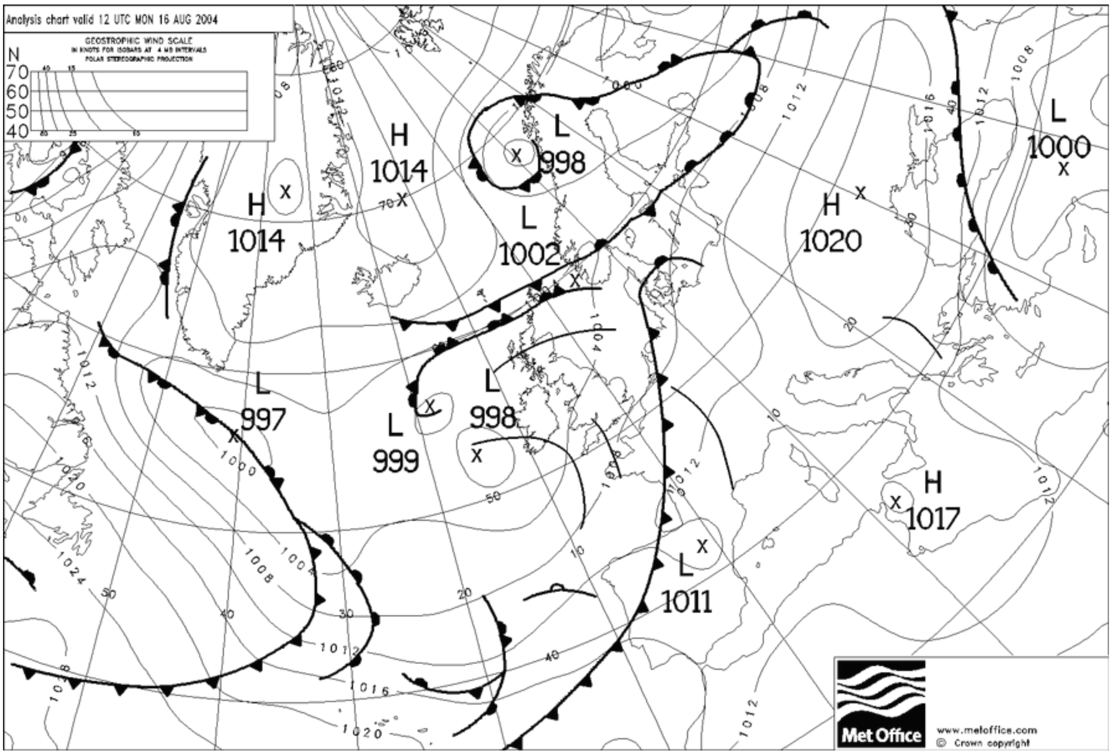


Figure 2-6. Synoptic chart for 12 UTC 16th August 2004

Figure 2-6 shows the synoptic pattern at midday (GMT). The southwest of England was under the influence of a south-westerly flow and between two troughs (marked as black lines). The storms appeared to have developed along a line of local convergence that may have been due to the land/sea contrast in the vicinity of the north coast of Cornwall. The individual storm cells kept re-generating at the same location to the southwest of Boscastle before tracking northeast and producing large amounts of rain over the small river catchment just inland from Boscastle before moving off to allow the next cell to repeat the process. It was the succession of small but torrential storms propagating over the same location that produced the very high rainfall totals.

The highest recorded daily rainfall amount was 200.4mm at Otterham (Figure 2-7). Radar estimates at 5km resolution gave more than 70mm over a 6-hour period. For a much more comprehensive overview of the meteorology, refer to Golding et al (2005).

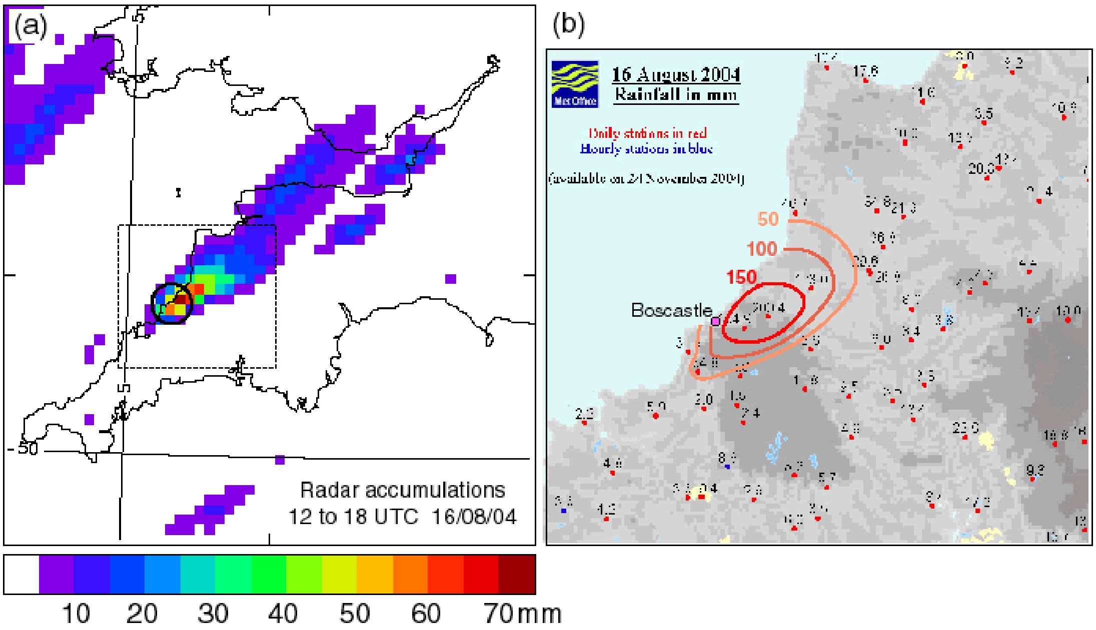


Figure 2-7. Rainfall accumulations for 16th August 2004. (a) From radar for the period 12 to 18 UTC (b) From rain gauges for the whole day with hand drawn contours for 50, 100 and 150mm. Gauge measurements courtesy of the National Climate Information Centre (NCIC) and the Environment Agency (EA). Circle in (a) is 20km diameter centred at Boscastle.

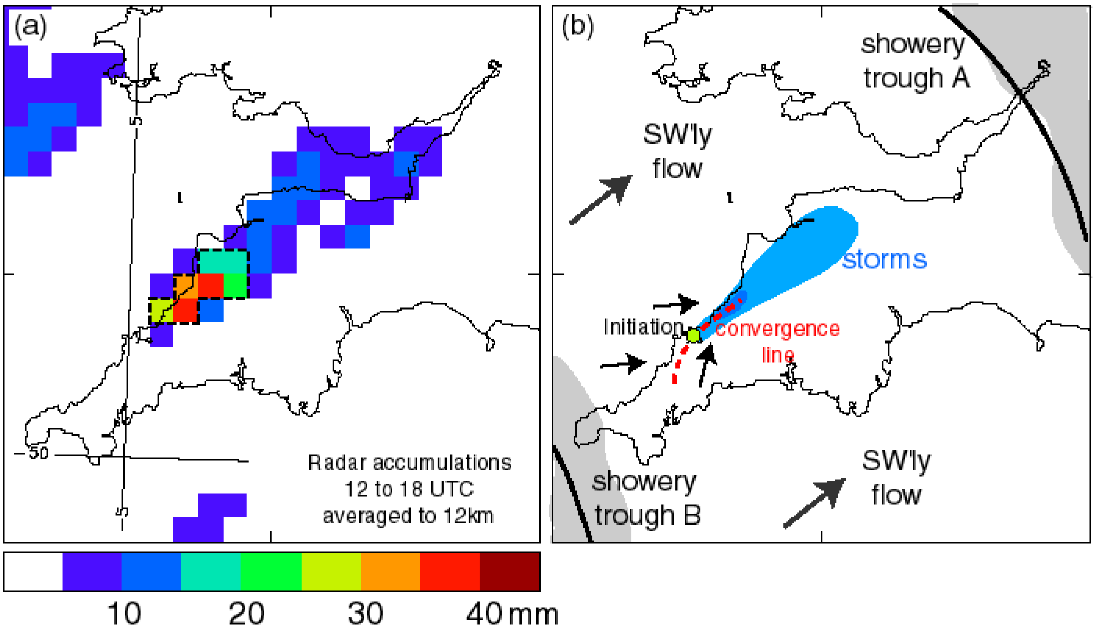


Figure 2-8. Rainfall accumulations from radar over the period 12 to 18 UTC 16th August 2004 averaged to the operational mesoscale model 12km grid. Dashed line encloses accumulations > 15mm. (b) Schematic of the meteorological situation during the afternoon of 16th August 2004.

Figure 2-9 shows the rainfall accumulations (projected on to the 12 km grid) over the period 12 to 18 UTC from 12, 4 and 1 km forecasts starting from 00 UTC. The accumulations can be directly compared with those from radar in Figure 2-8(a). The 12km forecast was extremely poor. Nowhere over the southwest of England did the predicted rainfall exceed 8mm. Other 12 km forecasts were equally poor.

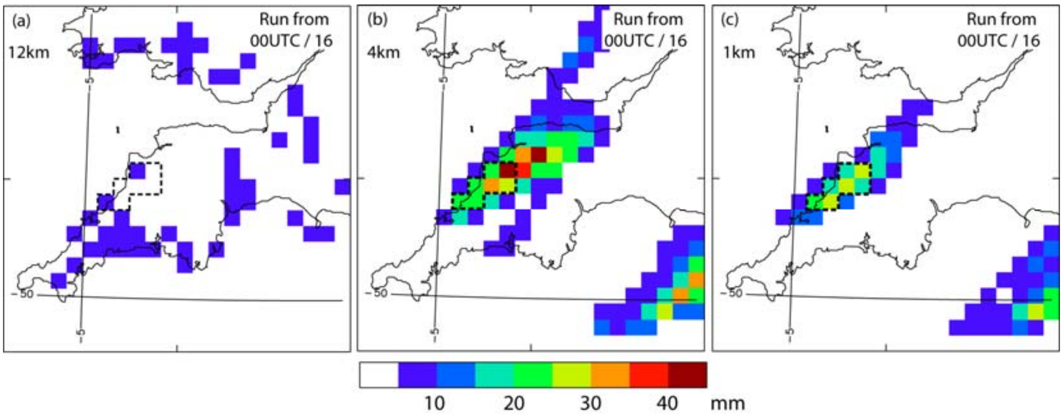


Figure 2-9. Rainfall accumulations over the period 12 to 18 UTC 16th August 2004 projected on to a 12 km grid from forecasts starting from 00 UTC. (a) 12 km forecast, (b) 4 km forecast, (c) 1 km forecast. The dashed line encloses the area of highest accumulations from radar (Figure 2-8).

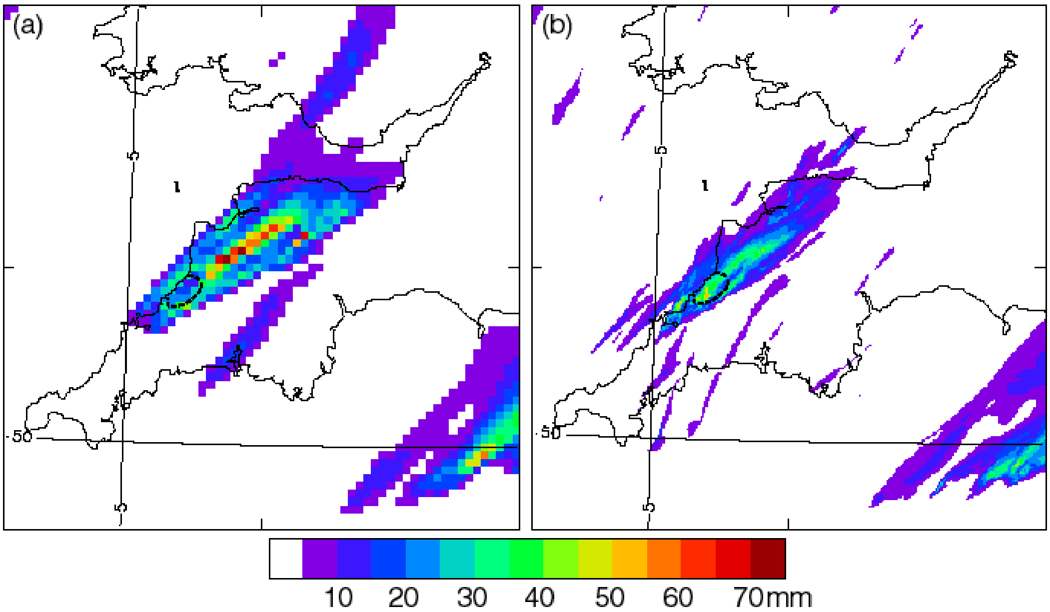


Figure 2-10. Rainfall accumulations over the period 12 to 18 UTC 16th August 2004 from (a) 4km forecast starting at 01 UTC and (b) 1km forecast starting at 01 UTC. Dashed line encloses accumulations > 50mm from radar (5km resolution), or 100mm from gauges (Figure 2-7).

In contrast to the 12km forecast, both the higher-resolution simulations predicted high totals (Figures 2-9(b & c) and 2-10). The 4km model produced amounts of rain that were comparable to those from radar, but the rain was displaced ~20-30 km to the northeast. The 1km forecast did not generate sufficient rain, but the location of the highest totals was almost perfect.

The high-resolution models were able to give a an indication that very high rainfall totals were possible locally on that day in the Boscastle area, and with suitable post-processing could have been used to warn of a chance of flooding, even without being able to reproduce the exact nature of the event.

This case was not chosen for further investigation in stage 2 because other investigations of this event were already underway.

2.4 Case D 7-8th January 2005 Carlisle

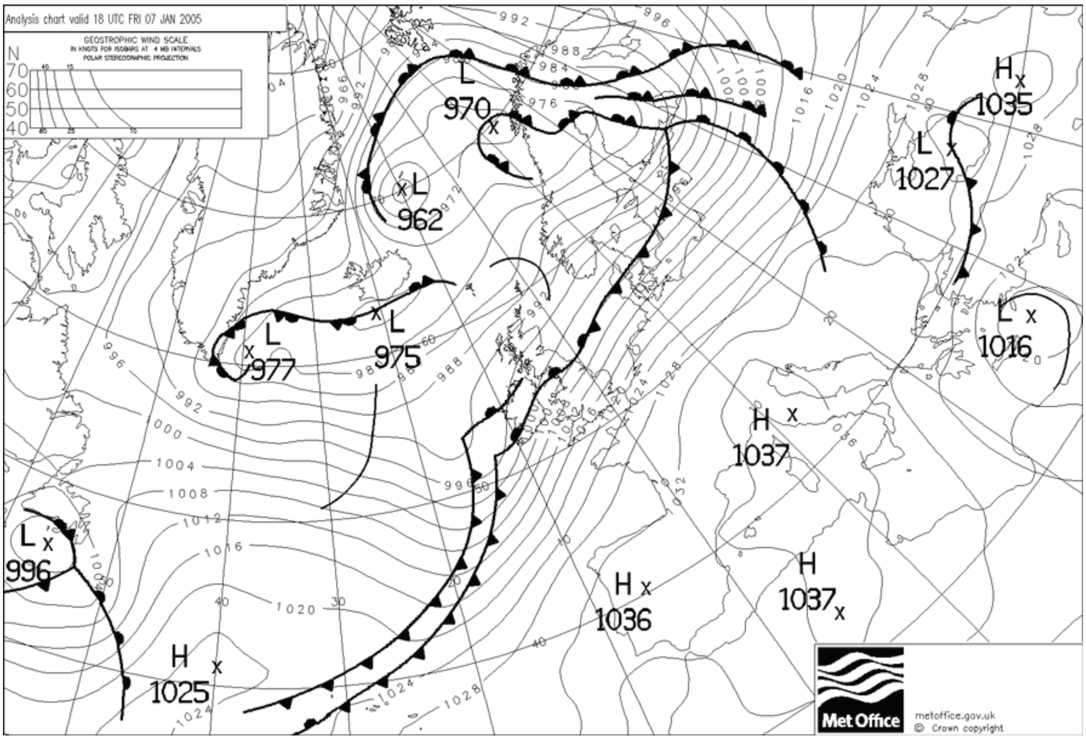


Figure 2-11. Synoptic situation for 18UTC 7th January 2005.

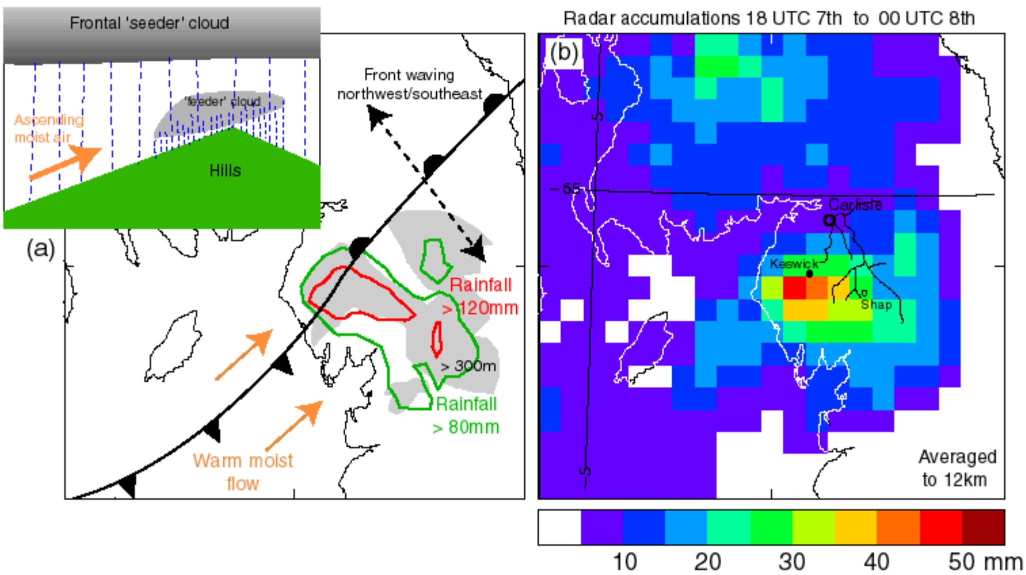


Figure 2-12. (a) Schematics of the meteorology of the period of orographically enhanced rainfall and the ‘seeder feeder’ mechanism. (b) Rainfall accumulations from radar over the 6-hour period 18 UTC 07/01/05 to 00 UTC 08/01/05 averaged to the operational mesoscale model 12km grid. The black lines show significant rivers that feed into Carlisle.

During the 7th and through into the early hours of the 8th a quasi-stationary waving frontal system brought a period of almost continuous rain to much of Cumbria (Figure 2-11). Very large quantities of rain were measured over the mountains of the Lake

District because of enhancement by the so-called ‘seeder-feeder’ effect (Figure 2-12(a)). The ‘feeder’ cloud is formed over the hills as warm moist low-level flow is forced to ascend, then as rain from the frontal ‘seeder’ cloud falls through the ‘feeder’ cloud it washes out additional droplets to produce a much higher rainfall rate than would have otherwise occurred from the frontal cloud alone.

In the 36 hours from 00UTC 7th until 12UTC 8th more than 160mm of rain was estimated by radar over some 5x5km squares and more than 100mm over a very wide area (~2700km^2). Point measurements from rain gauges over the 24-hour period from 09UTC 7th to 09UTC 8th gave 119.8mm at Shap and 95.0mm at Keswick.

Periods of 3 or 6 hours from several forecasts (starting from 18 UTC 6th, 00, 06, 12 and 18 UTC 7th) have been pieced together to make up the 24-hour accumulations from 00 UTC 7th to 00 UTC 8th shown in Figure 2-13. These are compared with accumulations from rain gauges and radar over the same period. It was noticeable that the radar measurements were significantly lower than the gauge measurements over the areas where most of the rain fell. The accuracy of the radar measurements suffered in this situation because the beam was at a higher elevation than the seeder-feeder process. Even though an adjustment was made to allow for this, the radar measurements are regarded as less reliable than those from gauges. The adjustment is made on the basis of long-term average conditions and is not responsive to all of the specific conditions at a particular time. For that reason the gauge totals are shown in preference in Figure 2-13(a).

In comparison to the gauge totals, the 12 km model clearly produced too little rain and the highest totals fell too close to the south and west coasts rather than over the hills. In contrast, the 4 km model gave a much more accurate distribution of the rainfall. The rainfall totals were also much closer to those observed, although the highest amounts were probably somewhat too low. The 1 km model produced, probably, the most accurate distribution of the highest totals, although it appears to have generated too much rain in general.

These results support the view that a 12 km model is incapable of getting the correct distribution of the rainfall in a seeder-feeder situation over Cumbria because of an inadequate representation of the orography (Figure 2-14) even if the general meteorological situation is largely correct. It has been shown that forecasts from the Met Office Unified Model with a grid spacing of 1 or 4 km were capable of producing more accurate predictions of the rainfall than the 12 km model operational at the time. The 4 km and 1 km models did produce somewhat too much rain, but this can be partly attributed to an over-prediction of the frontal (feeder) rain in all the forecasts (whatever resolution) on this occasion, which would have been amplified more by the seeder-feeder process in the higher-resolution models.

Typically, we should expect that the higher the resolution of the model the more accurate the catchment-scale precipitation forecasts will be for this type of situation.

As a result of this work high-resolution rainfall forecasts from this case have been fed into the Probability Distributed Model (PDM) hydrological model developed at the Centre for Ecology and Hydrology (Moore, 2007). Results are encouraging and are expected to be published.

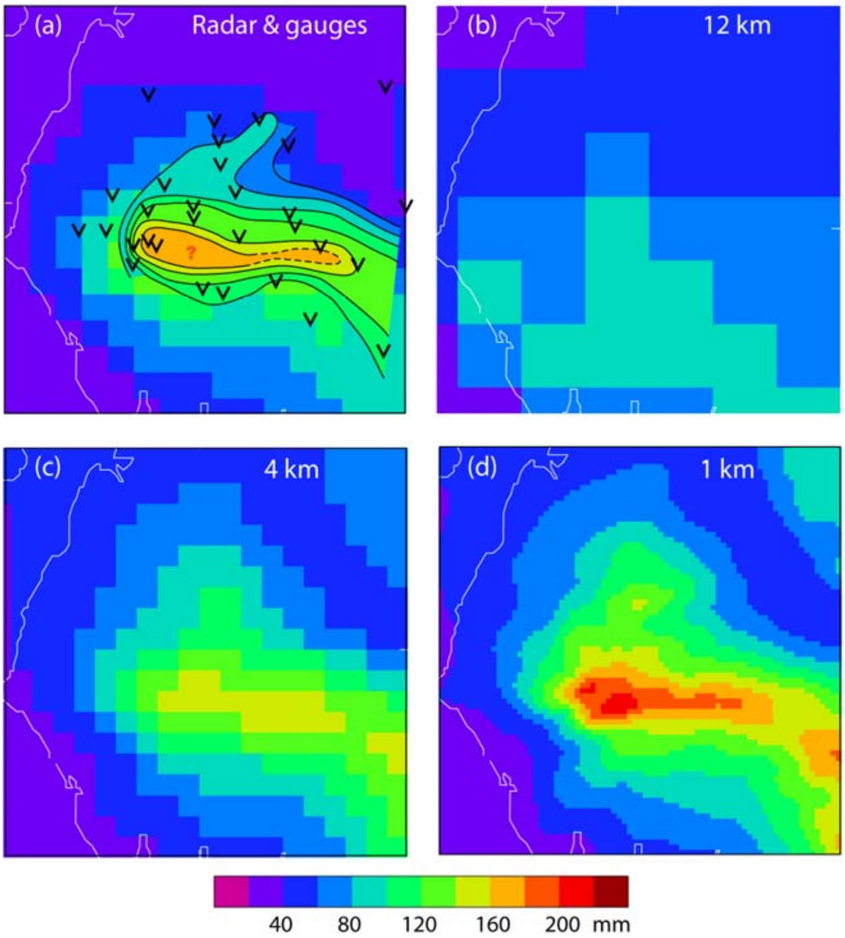


Figure 2-13. Rainfall accumulations for the period 00 UTC 7th to 00 UTC 8th January 2005. (a) Hand analysis from gauges (locations at bottom of 'V') contoured with shading on top of radar (block colours). (b, c, d) From a sequence of forecasts (see text) at 12, 4 and 1km.

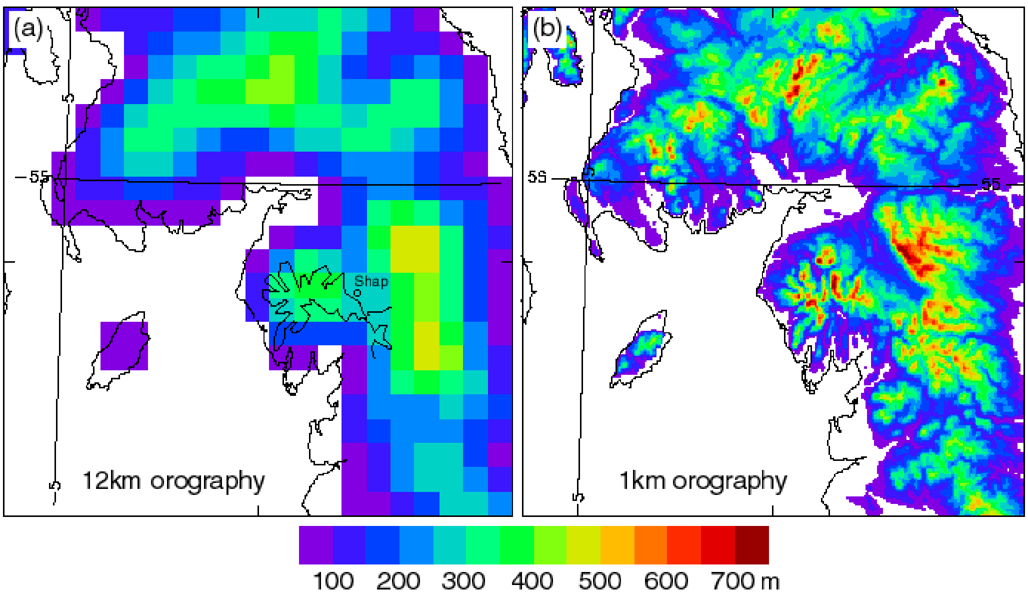


Figure 2-14. Model orography for (a) operational 12km mesoscale model and (b) 1km grid-length model. The black line in (a) gives a rough guide to a 300m contour in (b).

2.5 Case E 19th June 2005 North Yorkshire

This case was chosen for further study in stage 2. It will only be briefly described here. High-resolution model simulations are shown in section 3.

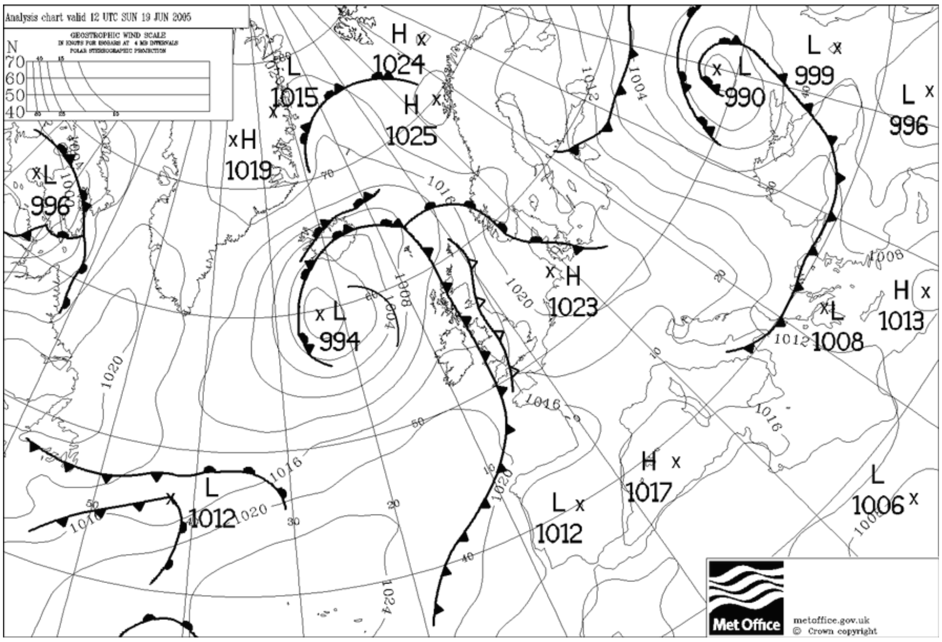


Figure 2-15. Synoptic chart for 12 UTC 19th June 2005.

On the 19th June 2005 most of the UK was under the influence of hot humid air that had been advected northwards from the continent. The synoptic chart for 12 UTC is shown in Figure 2-15. Thunderstorms developed during the afternoon in northern England. The most intense rainfall was concentrated in a small region to the northwest of the village of Hawnby in North Yorkshire. Figure 2-16(a) shows rainfall totals in excess of 90mm in the 6-hour period from 12 to 18 UTC; with most of that rain falling between 16 and 17 UTC. At Hawnby itself the radar measured around 70mm. The rain gauge at Hawnby (information provided by the Environment Agency) recorded 69.4mm in three hours, of which 59.8mm fell in one hour and 50.6mm in just half an hour. All these totals have a return period of more than 200 years (using the Flood Estimation Handbook method). The result of all that rain was flash flooding at Hawnby, the larger nearby village of Helmsley and the surrounding area.

Rainfall accumulations over the period 12 to 18 UTC 19/06/05 from four operational mesoscale model forecasts are shown in Figure 2-17. The pictures can be directly compared with radar accumulations over the same period and on the same grid (Figure 2-16(b)). None of the forecasts managed to predict the high rainfall totals although they did all produce some convective rain. The forecast that started at 00 UTC did the best by generating totals over 25mm, but that is not even close to the 70mm observed by radar on the same grid, and the highest totals were wrongly positioned by more than 50km.

The failure of the 12 km mesoscale model forecasts sets the scene for the investigation presented in section 3.

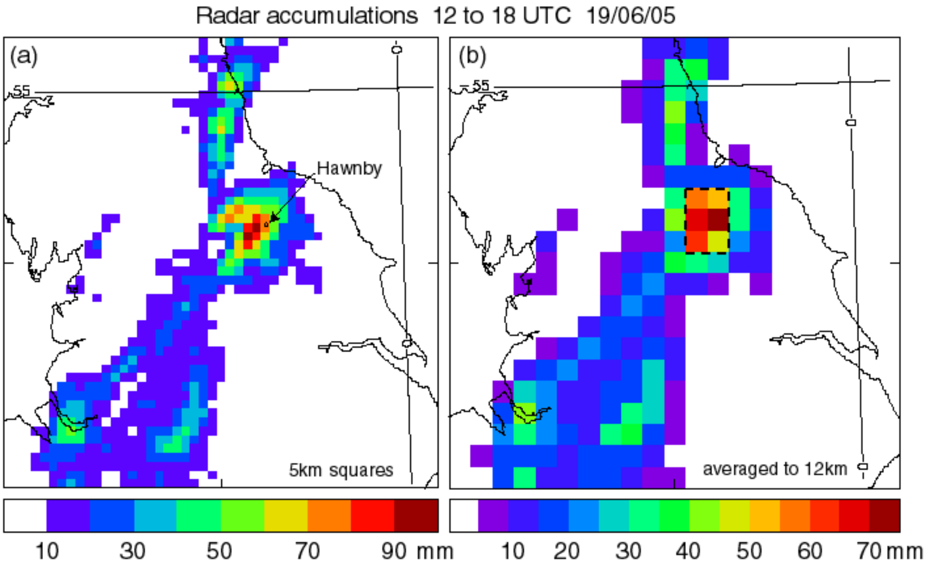


Figure 2-16. Rainfall accumulations from radar over the period 12 to 18 UTC 19th June 2005, (a) on 5km grid and (b) averaged to operational mesoscale model 12km grid. Dashed line in (b) encloses accumulations >45mm.

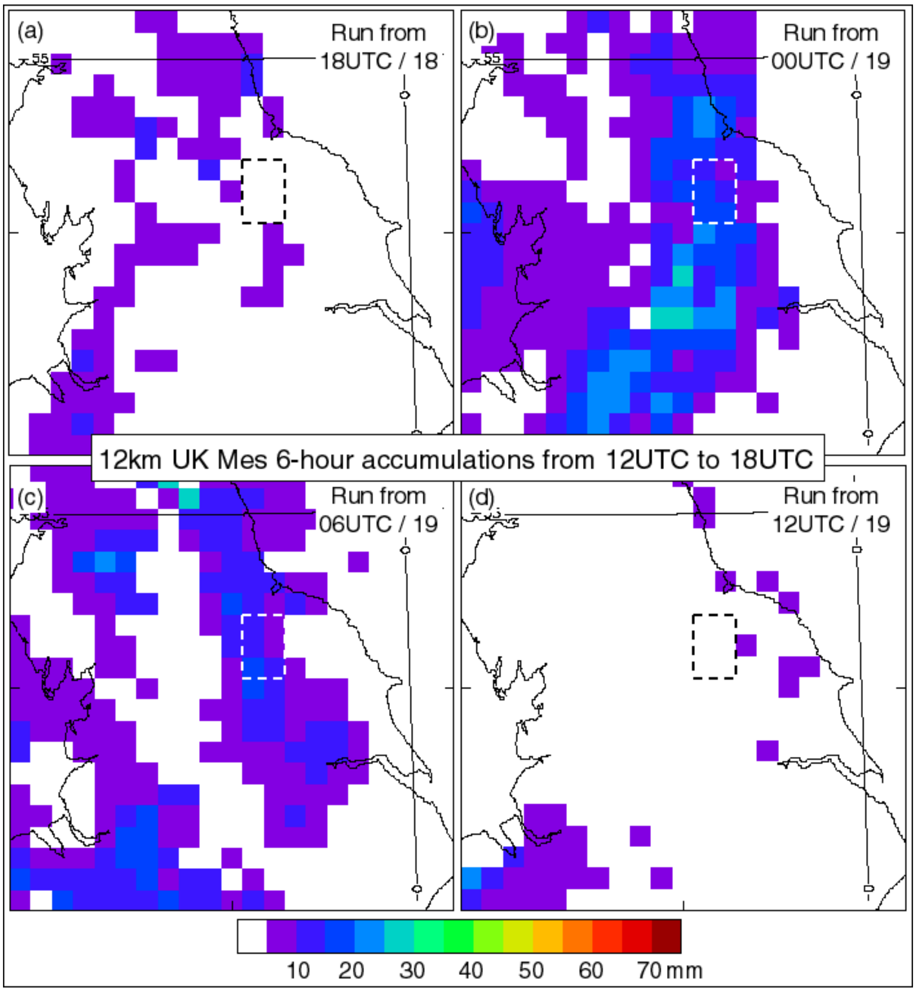


Figure 2-17. Operational mesoscale model forecasts of rainfall accumulations over the 6-hour period 12 to 18 UTC 19/06/05. Forecasts started from 18 UTC 18/06/05 and 00, 06, 12 UTC 19/06/05. Dashed line encloses accumulations > 45mm from radar (Figure 2-16(b)).

3. Further examination of cases B and E

This section gives a shortened version of what was presented in the stage-2 report. It is based on a detailed examination of 1 km simulations for Cases B and E.

3.1 Case B. 3rd August 2004

The event has already been described in section 2. Figure 3-1 shows the rainfall accumulations from radar over the afternoon & evening of the 3rd August 2004.

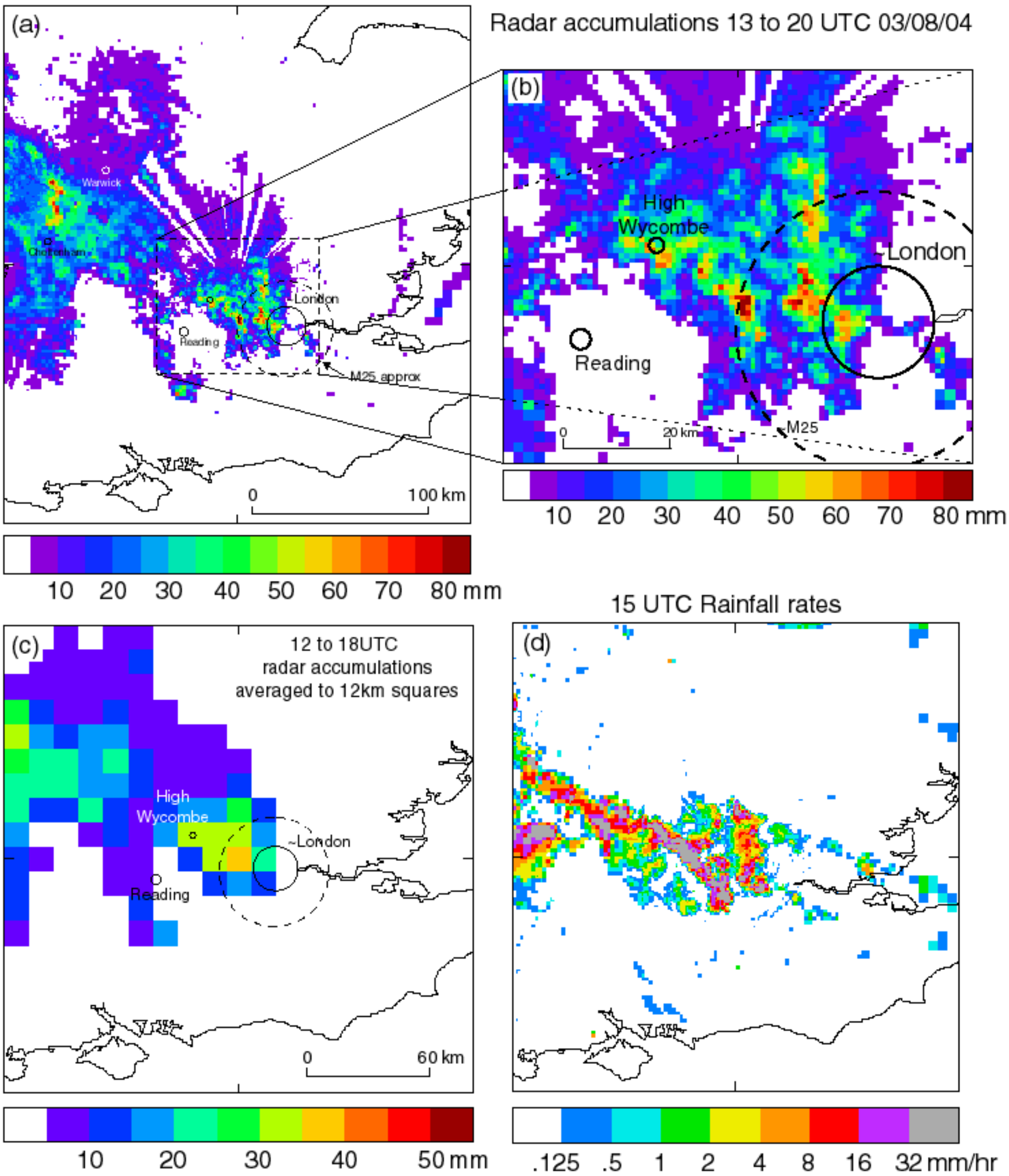


Figure 3-1. (a & b) Radar accumulations for the period 13 to 20 UTC 3rd August 2004 over Southeast England and a smaller sub area (1km resolution over sub-area). (c) Radar accumulations over the period 12 to 18 UTC averaged to the operational mesoscale model 12km grid. (d) Rainfall rates from radar at 15 UTC on a 1km grid over the area of interest. The circles in (a), (b) and (c) give a guide to the boundaries of inner and outer London.

3.1.1 The forecasts

The Unified Model was run with a grid spacing of 12, 4 and 1 km from several start times; 18 and 21 UTC on the 2nd and 00, 03, 06, 09 and 12 UTC on the 3rd August 2004. The forecasts are labelled in table 3-1. The focus of this investigation was on 1 km forecasts, not because the 4 km forecasts are unimportant, but because the objective was to examine whether an accurate 1 km forecast could be achieved and put that in the context of the 12 km forecasts.

Start time	12km	1km
18 UTC 02/08/04	LN12km18-02	LN1km18-02
21 UTC 02/08/04	LN12km21-02	LN1km21-02
00 UTC 03/08/04	LN12km00-03	LN1km00-03
03 UTC 03/08/04	LN12km03-03	LN1km03-03
06 UTC 03/08/04	LN12km06-03	LN1km06-03
09 UTC 03/08/04	LN12km09-03	LN1km09-03
12 UTC 03/08/04	LN12km12-03	LN1km12-03

Table 3-1. Note that the 1-km forecasts are labelled as starting at the same time as the 12-km forecasts. This is because they share the same analysis time even though the 1-km forecasts begin 1-hour later (see Figure 5).

12 km

Rainfall accumulations over the period from 13 to 18 UTC from radar and from several 12-km forecasts are shown in Figure 3-2. All the accumulations were projected on to the same 5km grid.

The forecast from 18 UTC on the 2nd (LN12km18-02) was clearly the least accurate in terms of rainfall location (panel (b)). The other forecasts were more accurate spatially because they were able to produce a northwest-southeast oriented band of rain >5mm. There appears to have been a sharp change towards an improved rainfall distribution between LN12km18-02 and the later forecasts. The later forecasts, although a better fit spatially, did not however necessarily produce more accurate rainfall totals. LN12km00-03 and LN12km09-03 produced far too little rain. LN12km03-03 and LN12km06-03 both produced more realistic totals, although still not enough. The forecast from 06 UTC (LN12km06-03) was the best 12-km forecast.

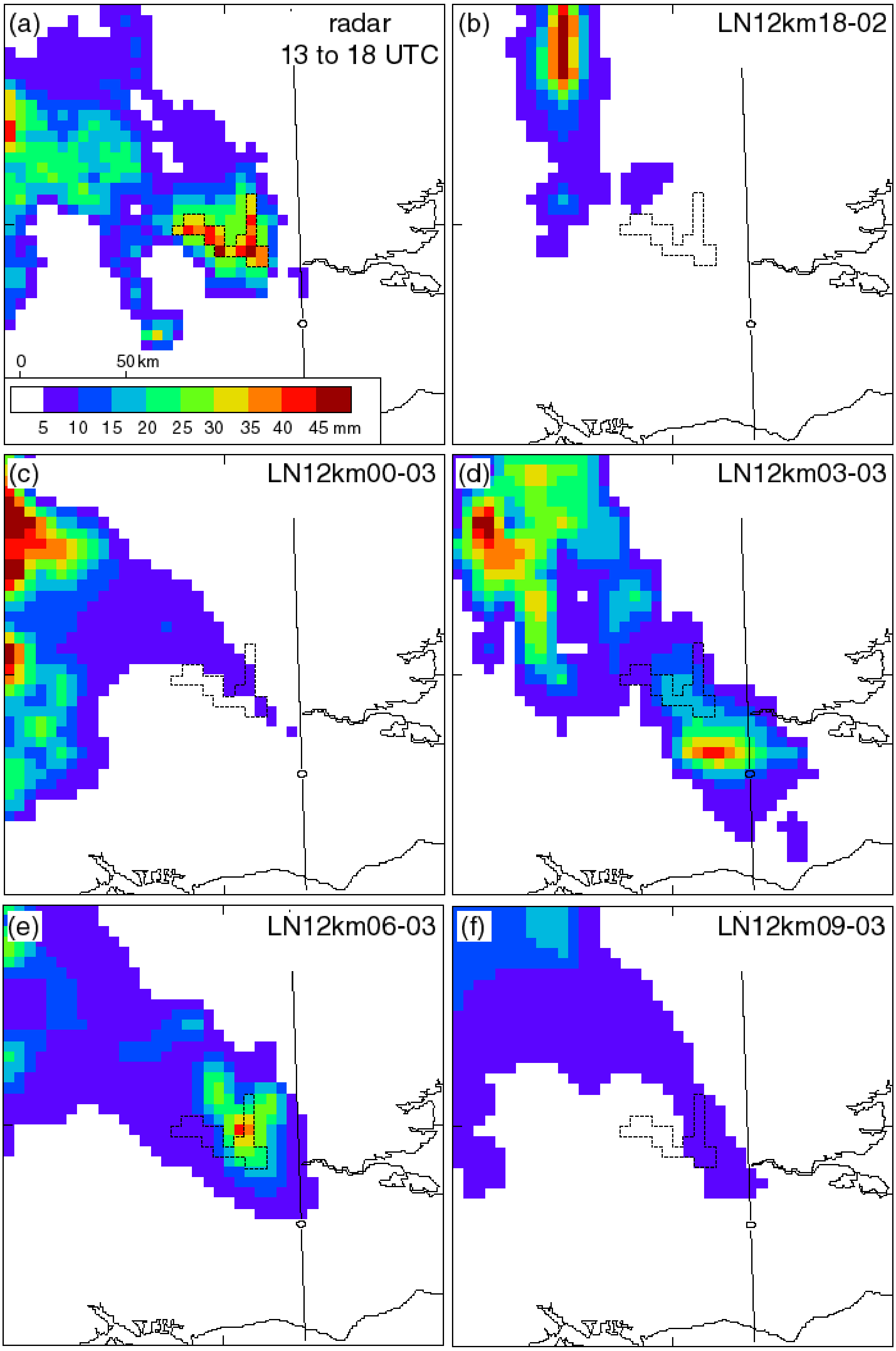


Figure 3-2. (a) Rainfall accumulations for the period 13 to 18 UTC 3rd August 2004 projected on to a 5-km grid from (a) radar, (b-f) 12-km model forecasts starting from 18 UTC 2nd, 00 UTC 3rd, 03 UTC 3rd, 06 UTC 3rd, 09 UTC 3rd.

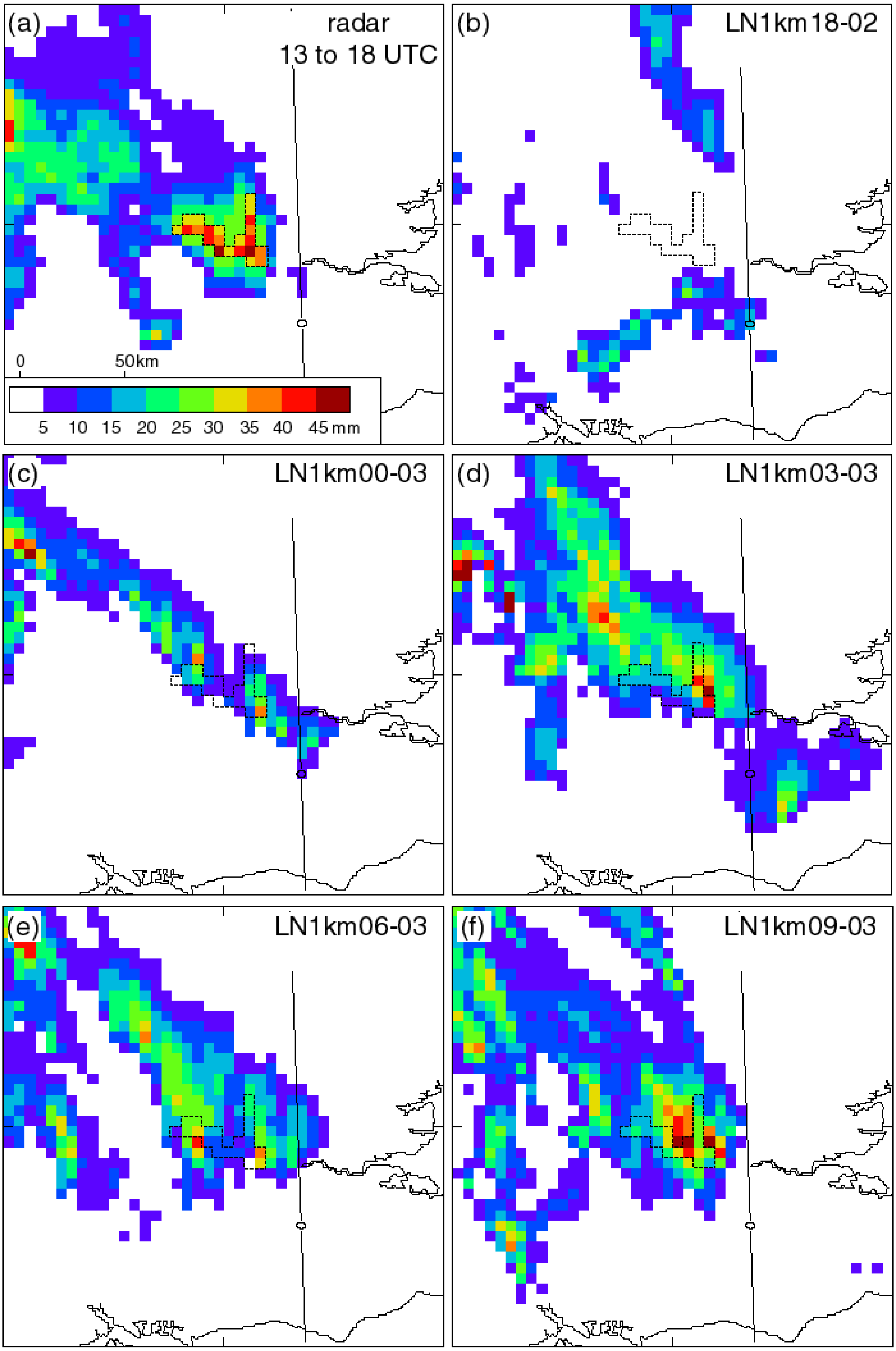


Figure 3-3. (a) Rainfall accumulations for the period 13 to 18 UTC 3rd August 2004 projected on to a 5-km grid from (a) radar, (b-f) 1-km model forecasts starting from 18 UTC 2nd, 00 UTC 3rd, 03 UTC 3rd, 06 UTC 3rd, 09 UTC 3rd.

1 km

Rainfall accumulations over the period from 13 to 18 UTC from radar and from several 1-km forecasts are shown in Figure 3-3 (equivalent to the 12-km forecasts in Figure 3-2). All the accumulations were projected on to the same 5km grid. As seen in the 12-km forecasts, the forecast from 18 UTC on the 2nd (LN1km18-02) was the least accurate in terms of rainfall location. The other forecasts were more accurate spatially because, like their 12-km counterparts, they were able to produce a northwest-southeast oriented band of rainfall. The improvement between LN1km18-02 and the other later forecasts is similar to what was seen in the 12-km model and suggests that a change in the representation of the larger-scale flow was responsible.

The latest 1-km forecasts (shorter lead times) were the most accurate spatially, and, unlike their 12-km counterparts, they were able to produce rainfall totals very close to those observed by radar. Both the forecasts from 06 and 09 UTC (LN1km06-03 and LN1km09-03) were extremely good forecasts, especially the one form 09 UTC (LN1km09-03), which gave an excellent prediction of high rainfall totals in the London area.

3.1.2 Reasons for the differences in forecast accuracy

The larger-scale flow

Each of the 1-km forecasts was initiated from 12-km model fields (1 hour into the 12-km forecast) and the same 12-km forecast provided the 1-km model with information at the boundaries (via a short section of the 4 km model) throughout the forecast period. For this reason, the larger-scale flow pattern within the 1-km forecasts (e.g. low pressure areas, frontal zones) was largely controlled by the 12-km model. Therefore, if we wish to examine whether there are any larger-scale errors present in the 1-km forecasts, we are justified in first looking at the behaviour of the 12-km forecasts.

The Water Vapour image from the geostationary Meteosat satellite for 12 UTC, (2-3 hours before the storms triggered), is shown in Figure 3-4(a). Two features of interest have been highlighted in the picture. Firstly, the oval-shaped blue area 'D', enclosed by a black arrow. This is where there was a local lowering of the tropopause (boundary between tropospheric air and drier stratospheric air) and where there was rotation at upper levels. Secondly, the orange/red areas bounded by a red line. This is where an area of cloud has formed from ascent in association with the upper-level rotation and was subsequently drawn around in the rotation.

Temperature and humidity fields from NWP model forecasts have be used to generate pseudo WV imagery for comparison with the satellite WV imagery. This is also shown in Figure 3-4 for several of the 12-km forecasts. The forecast from 18UTC on the 2nd (LN12km18-02,) does not match the WV image particularly well. The dry (blue) region is too stretched and too far south and the hooked area of cloud too far east. The forecast that initiated 6 hours later at 00 UTC (LN12km00-03) showed a noticeable improvement in the position of both the dry region and the hook of cloud.

The improvement in the upper-tropospheric flow-pattern between the forecast from 18 UTC on the 2nd (LN12km00-03) compared to 00 UTC on the 3rd (LN12km18-02) is matched by the improvement in the spatial distribution of the rain shown earlier in Figure 3-2. Since the larger-scale evolution of the 1-km forecasts is controlled to a large extent by their equivalent 12-km forecasts there should also be an improvement in the 1-km forecast from 00UTC on the 3rd (LN1km18-02) compared to the forecast from 18UTC on the 2nd (LN1km00-03). Again, we saw in Figure 3-3 that this is the case. In fact, it is more noticeable than that seen at 12km.

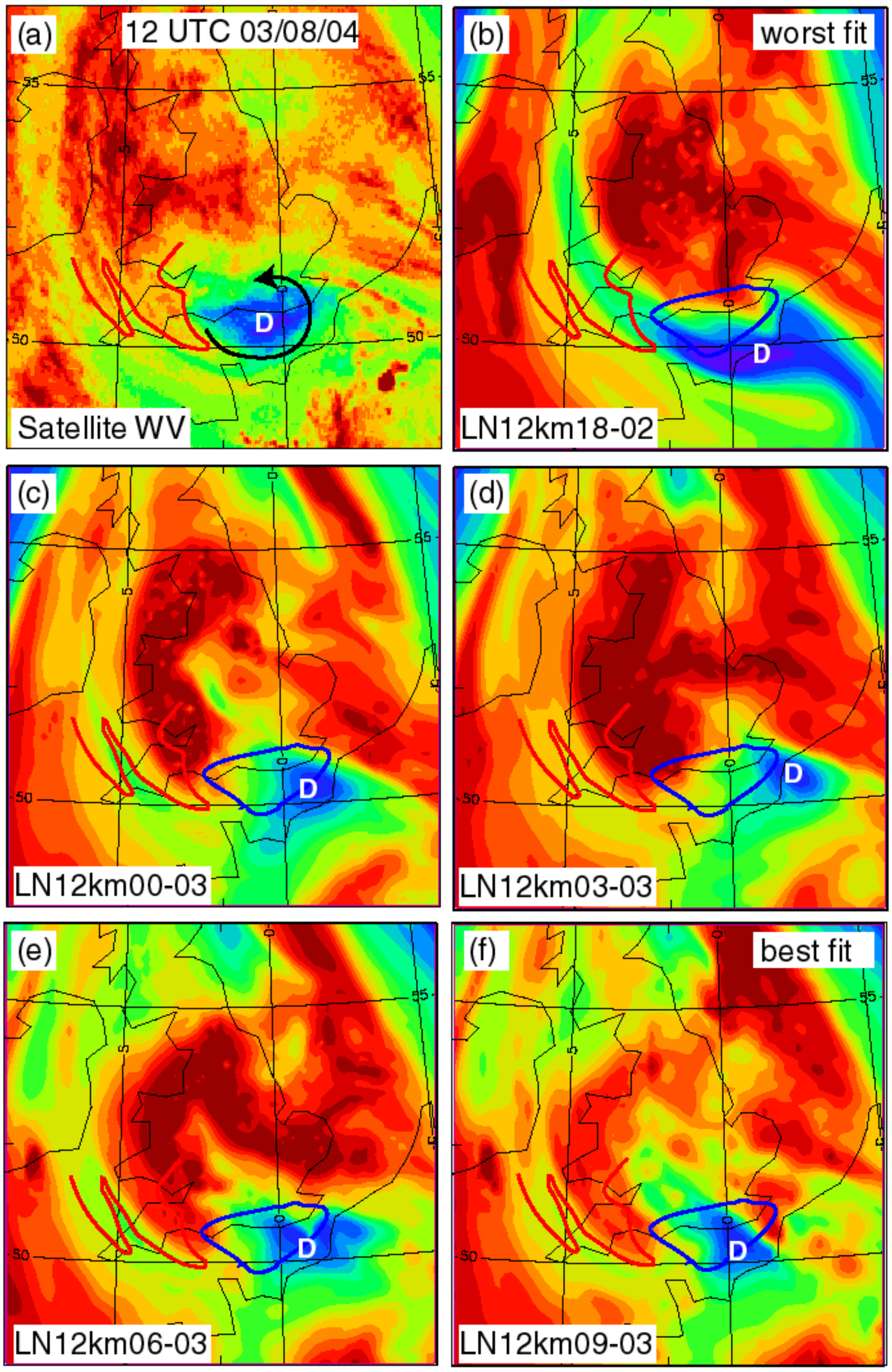


Figure 3-4. (a) Meteosat Water Vapour image for 12 UTC 03/08/04. Areas with warmer brightness temperatures (dry regions) are blue, areas with cooler brightness temperatures (moist/cloudy regions) are orange/red. 'D' marks the centre of a small dry region which signifies a small region of rotation at upper levels (7-10km). The red lines mark the edge of the moist/cloudy region wrapping round the dry area. (b-f) pseudo Water Vapour imagery extracted from relative humidity fields at 300 to 600 hPa from the 12-km forecasts starting at 18 UTC 2nd, 00, 03, 06 and 09 UTC 3rd August as comparison against the imagery. The blue line markes the dry region in the satellite image. The red line is the same as in (a).

There is not much difference in the fit to WV imagery at 12UTC on the 3rd between the forecasts starting at 00UTC on the 3rd and subsequent forecasts starting from 03, 06 and 09 UTC on the 3rd. The 12-km forecast starting from 09UTC on the 3rd has the best fit and the equivalent 1 km forecast is the most accurate.

It is essential for the larger-scale dynamics to be well represented to allow the possibility of a good forecast because the larger-scales control the envelope region within which convection can develop. In this case, when the fit between the WV imagery and the model pseudo WV imagery was poor, the rainfall distribution was also poor. When the fit to the imagery was better, the forecasts produced a better rainfall evolution.

The importance of the local dynamics

It has been shown that the position of an upper-level vortex played an important role in the ability of the models to predict the storms. The 1-km forecasts from 00, 03, 06 and 09 UTC were all in reasonable agreement with the imagery and produced reasonable forecasts. However there were also more subtle differences in skill between those forecasts, and it was the forecast from 09UTC (LN1km09-03) in particular that gave the most accurate local predictions of the rainfall amounts.

Figure 3-5 shows the observed and forecast convergence lines at 13 UTC before convection in the areas of interest had triggered and superimposed on that the observed and forecast rainfall 30 minutes later. The convergence lines show where the wind from one direction was meeting the wind from another direction near the ground. It is striking that both the forecast and observed storms triggered along their respective convergence lines. It appears that in this case, if the convergence lines are correctly represented, and given that everything else is broadly correct, the storms will initiate in the correct locations. So part of the key to getting the detail in the rainfall correct is to get the local dynamics accurately represented.

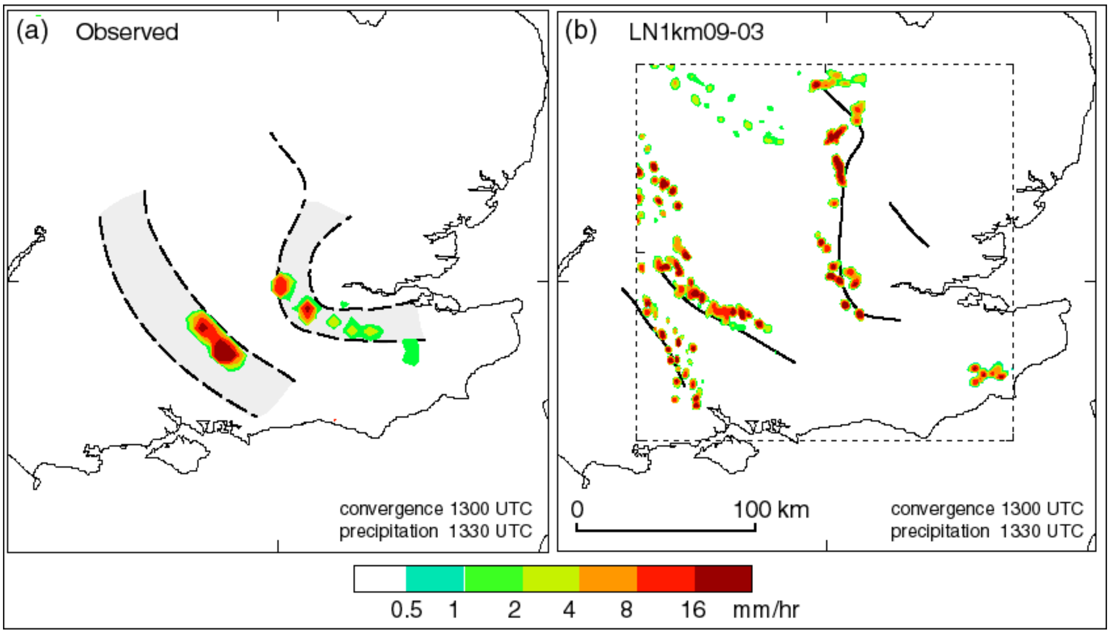


Figure 3-5. (a) Grey shading shows the observed bands of low-level convergence at 13 UTC 3rd August 2004. Shading is the new rainfall 30 minutes later from radar. (b) Black lines are lines of low-level convergence at 13 UTC in the 1-km forecast from 09 UTC. Shading is the new rainfall in the same forecast 30 minutes later (within the dashed box).

In the report from stage 2 (Roberts 2007) it was shown that the convergence lines prior to storm development in the 1 km forecast from 06 UTC (LN1km06-03) were not as accurately positioned as in the forecast from 09 UTC (LN1km09-03) and this is part of the reason why that forecast did not produce as accurate a forecast.

It was also shown that the convergence line in the London area occurred at the boundary of the cold outflow from some previous showers. The 1 km forecast from 09 UTC was able to generate the convergence line because it captured the earlier showers.

The 12-km forecasts from 06 and 09 UTC had no convergence lines in the London area because a 12-km model is incapable of generating such a feature with such a coarse grid and had insufficient resolution to resolve the previous showers. This demonstrates that there are important meteorological pre-cursers to convection that a 12-km model is simply unable to represent adequately.

3.1.3 Summary for case B

The 1-km model was able to produce a very good forecast of the convective storms that lead to flooding in parts of northwest London. The ability of the model to predict the storms was dependent on getting two components of the flow correct in the lead up to the storms.

Component 1: The upper-level vortex and associated dynamics needed to be correct. This component involves scales larger than the scales of the storms themselves. If the larger-scale flow is not sufficiently correct the forecast will be poor whatever the model resolution.

Component 2: The local dynamical pre-cursors also had to be correct for precision in the forecast. The best 1-km forecast (from 09 UTC, LN1km09-03) was more accurate than its predecessor (from 06 UTC, LN1km06-03) because it had a low-level convergence line in approximately the correct location just prior to the time the storms triggered, whereas the previous forecast did not.

The 1-km model significantly out-performed the 12-km model for all but one of the equivalent forecasts. The 1-km model was more consistent from forecast to forecast than the 12-km model because of a more accurate representation of the storms by the model dynamics. The variability of the 12-km forecasts came from a flip-flop between how much the convection was ‘resolved’ on the grid or represented by the convection parametrization scheme.

3.2 Case E. 19th June 2005: Flooding in North Yorkshire

The event has already been described in section 2. Figure 3-6 shows the synoptic situation a couple of hours before the storms developed. Figure 3-7(a) shows the high rainfall accumulations, measured by radar, that lead to the severe flooding.

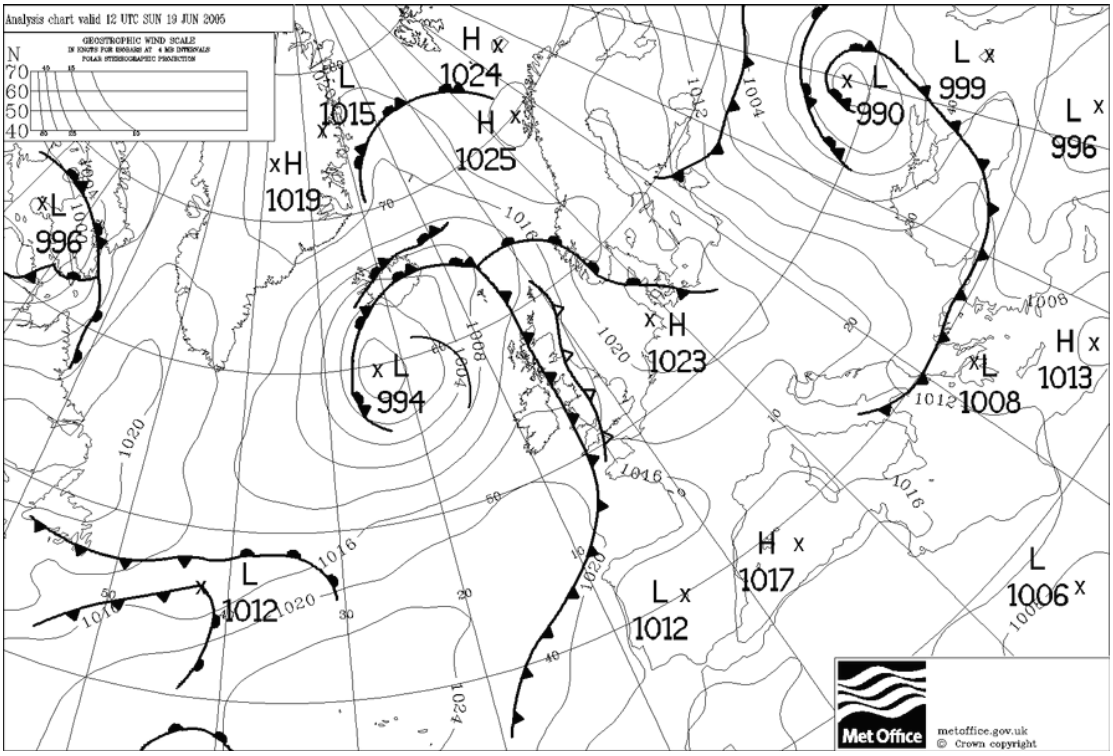


Figure 3-6. Synoptic chart for 12 UTC 19th June 2005.

Forecasts have been run from start times of 18 and 21 UTC on the 18th and 00, 03, 06, 09 and 12 UTC on the 19th June 2005. Each of the forecasts has been given a label shown in Table 3-2.

Start time	12km	1km
12 UTC 18/06/05	YK12km12-18	YK1km12-18
18 UTC 18/06/05	YK12km18-18	YK1km18-18
21 UTC 18/06/05	YK12km21-18	YK1km18-18
00 UTC 19/06/05	YK12km00-19	YK1km18-19
03 UTC 19/06/05	YK12km03-19	YK1km18-19
06 UTC 19/06/05	YK12km06-19	YK1km18-19
09 UTC 19/06/05	YK12km09-19	YK1km18-19
12 UTC 19/06/05	YK12km12-19	YK1km18-19

Table 3-2. Note that the 1-km forecasts are labelled as starting at the same time as the 12-km forecasts. This is because they share the same analysis time even though the 1-km forecasts begin 1-hour later.

3.2.1 The accuracy of the forecasts

12 km

Rainfall accumulations over the period 14 to 17 UTC from the 12-km forecasts initiating at 18 UTC on the 18th and 00, 03, 06 and 09 UTC on the 19th are displayed in figure 3-7. None of these forecasts was able to produce anything even vaguely resembling the high rainfall totals observed by radar. Two other forecasts (not shown) from 12 UTC on the 18th and 12 UTC on the 19th were no better. The 12-km model was unable to predict this storm.

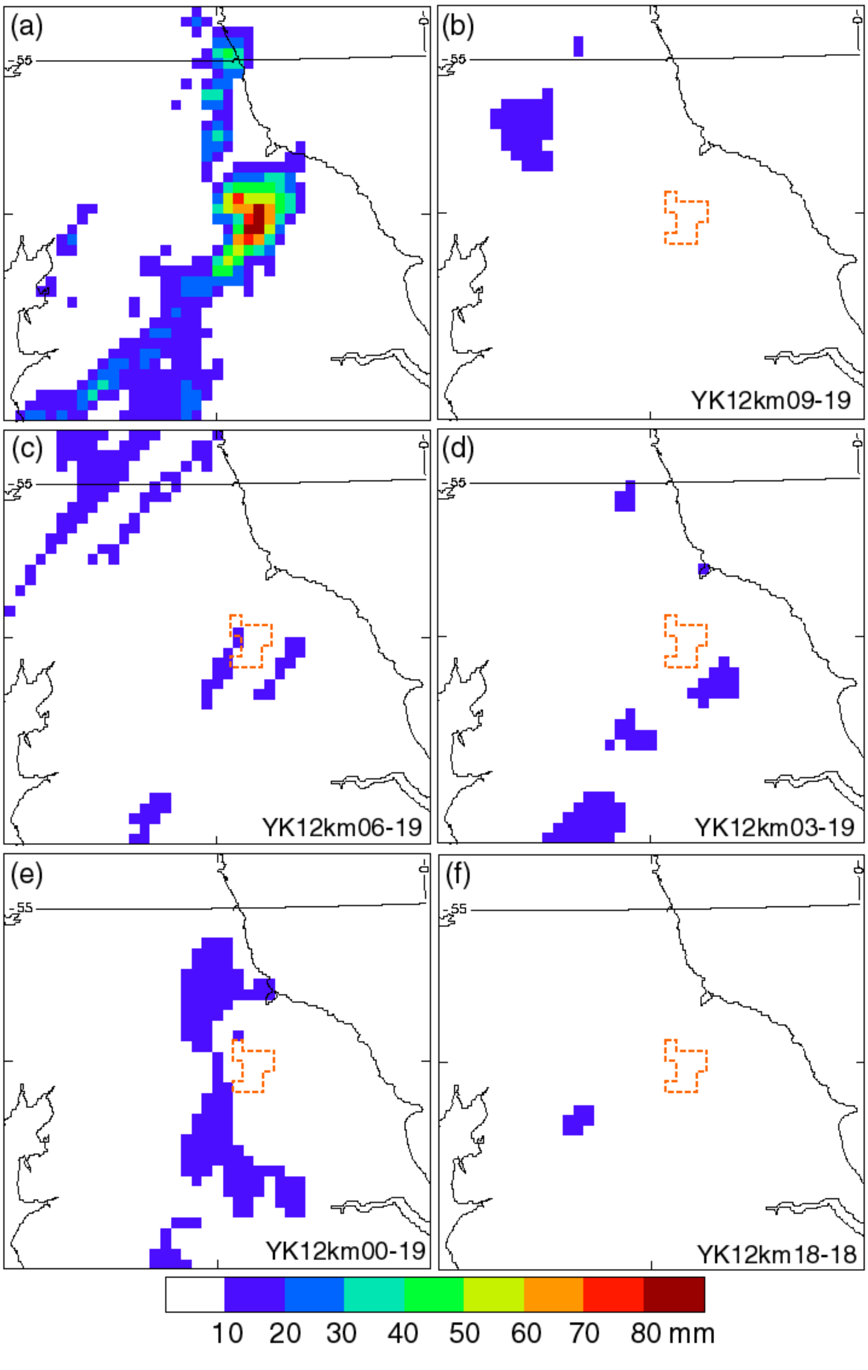


Figure 3-7. (a) Rainfall accumulations for the period 14 to 17 UTC 19th June 2005 projected on to a 5-km grid from (a) radar, (b-f) 12-km

model forecasts starting from 09, 06, 03, 00 UTC, 19th and 18 UTC 18th.

1 km

Forecasts run at 1 km from the same times are shown in Figure 3-8. The 1-km model was also unable to predict the high rainfall totals. The two other forecasts (not shown) from 12 UTC on the 18th and 12 UTC on the 19th were equally poor. Note, though, that there were still differences between the 12 km and 1 km forecasts. Although the 1-km model was unable to produce the high accumulations seen in the radar, it did manage to produce some local accumulation of 30-40mm in some of the forecasts, which do not show up when averaging to a 5-km grid, and it also generated the rain in southwest-northeast oriented bands as were observed.

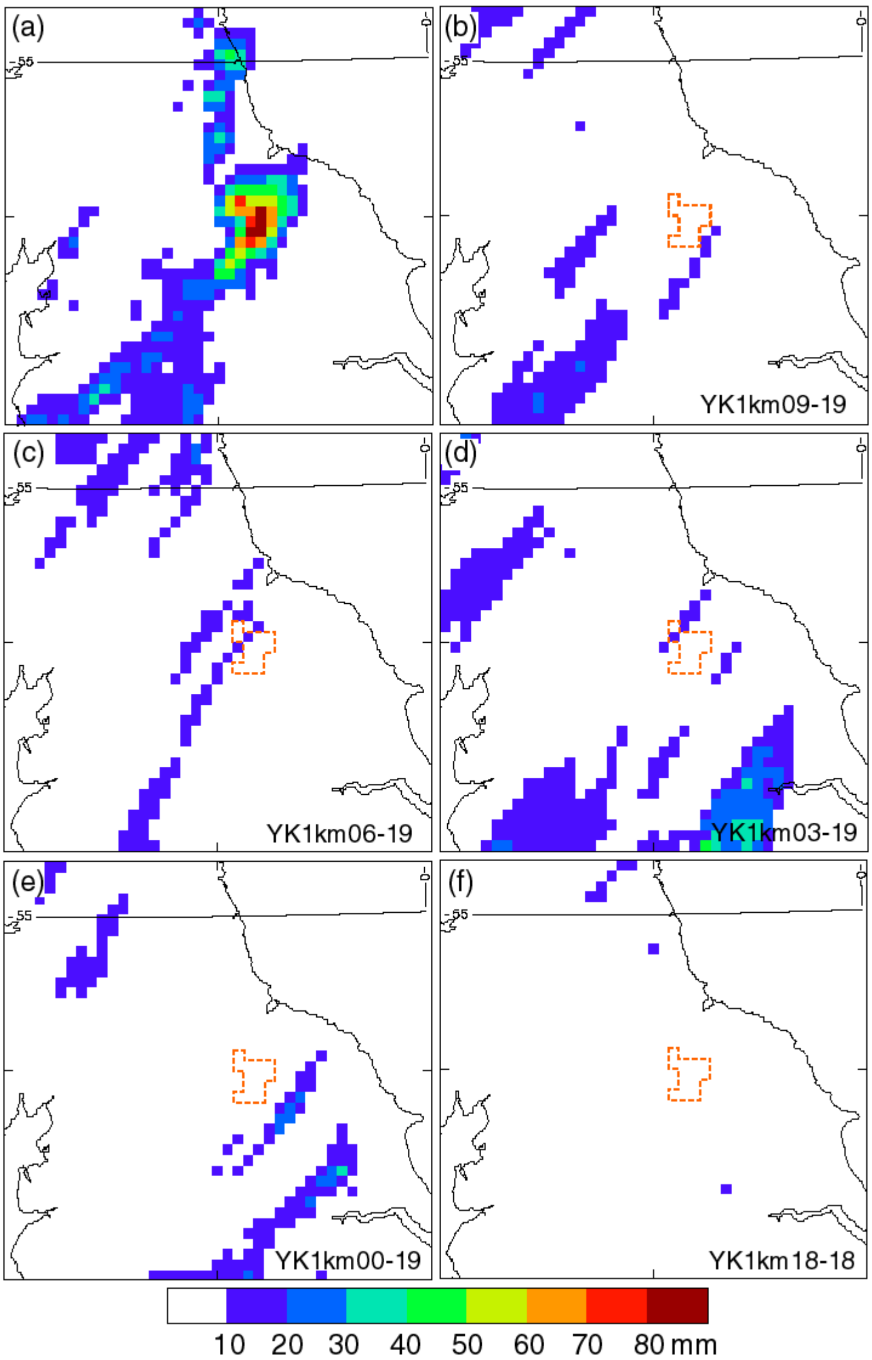


Figure 3-8. (a) Rainfall accumulations for the period 14 to 17 UTC 19th June 2005 projected on to a 5-km grid from (a) radar, (b-f) 1-km model forecasts starting from 09, 06, 03, 00 UTC, 19th and 18 UTC 18th.

In summary, neither the 12-km or 1-km models were able to generate the storm of interest in any of the forecasts. The 1-km model did a better job at predicting the possibility of localised storms and the banded nature of the storm area, but it still failed to produce a forecast that resembled the actual event.

3.2.2 Reason for the poor forecasts – the larger-scale flow

The immediate question that needs answering is why the forecasts were unable to reproduce the observed storm. The fact that improved resolution did not help much suggests that there may have been a problem with the representation of the larger-scale flow.

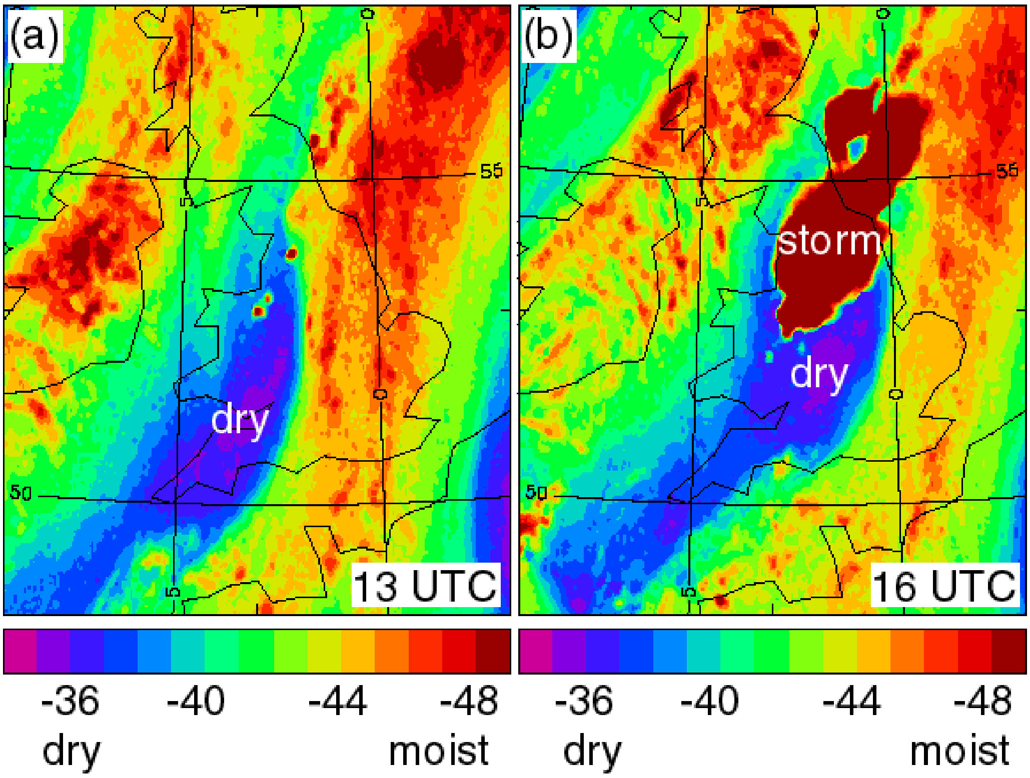


Figure 3-9. Meteosat Water vapour imagery for 13 and 16 UTC 19th June 2005. The blue colours show high brightness temperatures and therefore dry upper-tropospheric air. The /browns/reds/oranges show moist/cloudy upper-tropospheric air. The storm can be seen at 16 UTC growing from smaller clouds at 13 UTC.

Again, as for the previous case, we turn to Meteosat Water vapour (WV) imagery to give a view of what was happening with the larger-scale dynamics. We can see from Figure 3-9 that the storms developed within the northern part of a dry zone in the WV imagery and subsequently moved with that dry area. The sharp boundary between the dry region (blue) and the moist/cloudy area (orange/red) is marked by a white line in Figure 3-10(a) and indicates the position of the jet stream; the strong core of winds at an altitude of around 10 km. This boundary (the white line) had twisted into an S shape because of rotation at upper-levels. It is common for thunderstorms to develop ahead of areas of rotation at upper-levels (Browning and Roberts 1994, Roberts 2000)

Figure 3-10 shows how well pseudo WV images for 12 UTC from the 12-km model compared with the actual imagery. Pseudo WV imagery from the 1-km model is not shown because the domain is too small to cover the area required for a meaningful comparison. However, the 1-km forecast were very similar to their equivalent 12-km forecasts over the 1-km domain, so what we infer for the 12-km forecasts is also true for the 1-km forecasts.

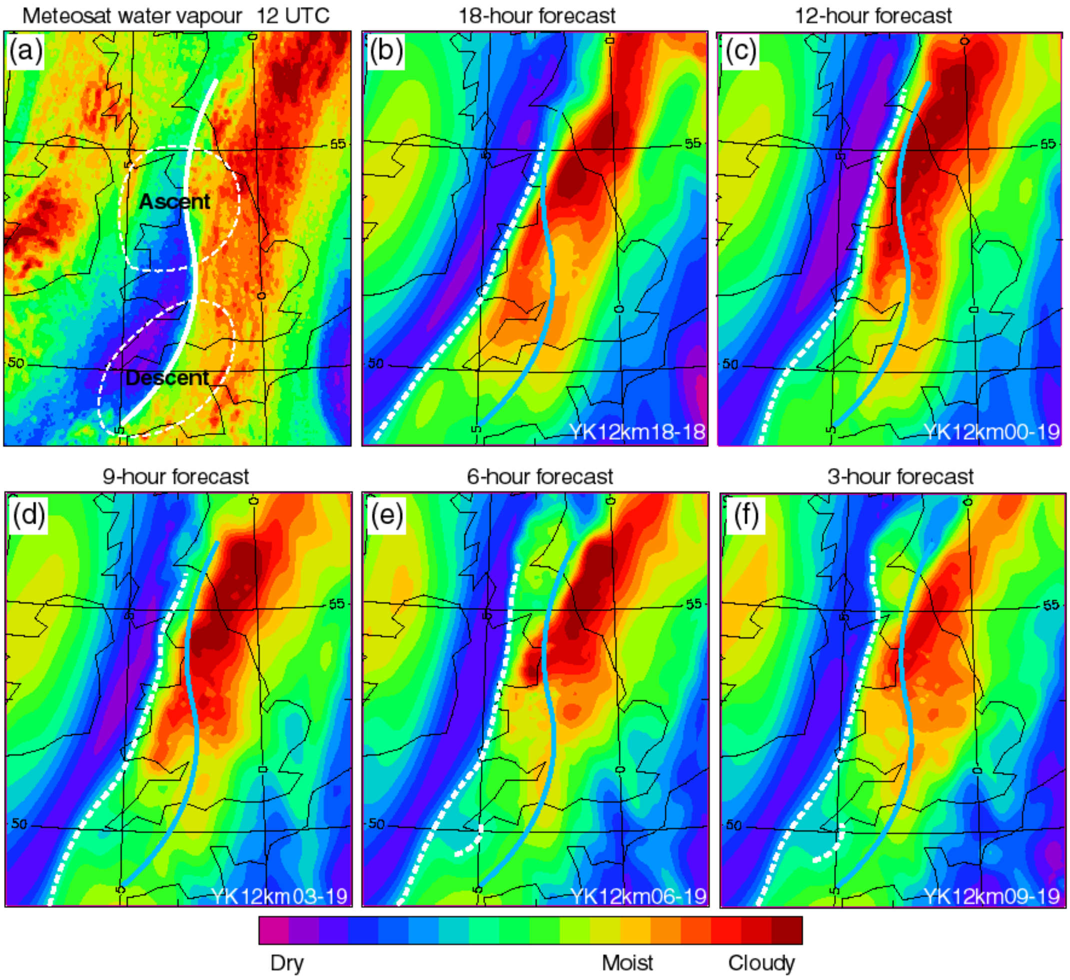


Figure 3-10. (a) Meteosat Water Vapour image for 12 UTC 19/06/05. Areas with warmer brightness temperatures (dry regions) are blue, areas with cooler brightness temperatures (moist/cloudy regions) are orange/red. The white line marks the sharp transition between the drier and moister air (typically the axis of the jet stream). General regions of possible ascent and descent within the troposphere are depicted by the dashed circles. (b-f) pseudo Water Vapour imagery extracted from relative humidity fields at 300 to 600 hPa from the 12-km forecasts starting at 18 UTC 18th, 00, 03, 06 and 09 UTC 19th June as comparison against the imagery. The blue line is the same as the white line in (a). The dashed white lines mark the moist/dry air transition in the pseudo imagery.

It is immediately evident that none of the 12-km forecasts had the degree of waving (rotation) that was observed. The boundary between the dry and moist air is much straighter. Clearly, the larger-scale upper-tropospheric dynamics was wrong in all the forecasts.

The report from stage 2 showed that the 1 km forecasts produced essentially the correct surface pressure pattern and essentially the correct temperatures, humidities and winds near the ground. They had the cold front correctly positioned and even had a deflection of the flow around the North York Moors that was observed.

The reason all the forecasts were poor was down to the error in the rotation of the upper-level flow.

3.2.3 Summary

Both the 1-km or 12-km models failed to reproduce the location and quantity of rainfall that was observed. Some of the 1-km model forecasts did manage to generate heavy showers in the right general area, but the showers were too small and transitory to produce very high rainfall accumulations. As for the previous case, the accuracy of the forecasts depended on getting both the larger-scale and local-scale components of the flow correct.

On this occasion the error in the larger-scale dynamics lead to the poor forecasts despite the local meteorology being well represented. The message from this case and the previous one is that both the larger-scale and local dynamics must be sufficiently correct for a good forecast. The larger scales should not be ignored even when the model resolution is very high.

4. Sensitivity experiments for cases B and E

This section shows brief results from some of the sensitivity experiments performed for cases B and E. These were part of stage 3 and were not presented in the stage 2 report. The purpose of this section is firstly to show the impact of scientifically defensible changes to the model grid or formulation in the 1 km model forecasts, and secondly to show the benefit of using the Fractions Skill Score verification approach to identify differences between forecasts over a range of scales.

4.1 Case B: 3rd August 2004

4.1.1 Change from 76 to 38 vertical levels

The purpose of this experiment is to see whether a poorer vertical resolution has a significant impact on the forecast. Increased vertical resolution has been shown to give a better representation of slanted frontal circulations in a high-resolution model (Lean and Clark 2003), however for convective situations it is not so clear whether it matters so much. In this case the representation of the low-level convergence lines and the 'lid' that initially inhibited convection would be modified by a reduced number of vertical levels.

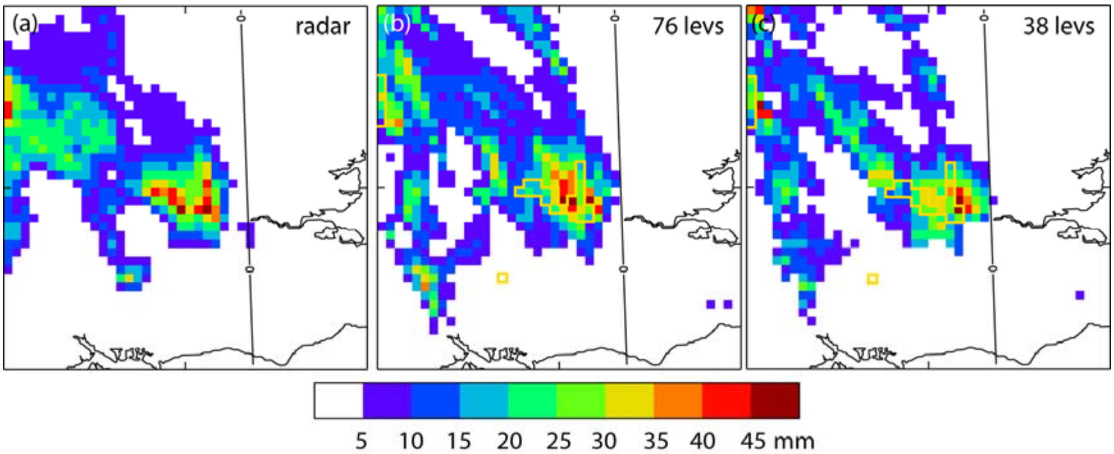


Figure 4-1. Rainfall accumulations over the period 13 to 18 UTC, 3rd August 2004 from (a) radar, (b) a 1 km forecast from 06 UTC with 76 vertical levels and (c) a 1 km forecast from 06 UTC with 38 vertical levels. The orange line in (b) and (c) enclosed the totals above 30 mm in (a) (over and to the west of London).

The rainfall accumulations in Figure 4-1 show the impact of changing from 76 to 38 vertical levels in the 1 km forecast from 09 UTC (the best forecast). The general pattern of rainfall did not change. The detail did though. The pixels with the highest accumulations moved a little and the number of pixels exceeding 5 mm reduced. From a subjective viewpoint, the reduction in the number of vertical levels made the forecast slightly worse in the London area, but the change was no greater than the expected uncertainty in the detail of the rainfall amounts.

No general conclusions about the impact of vertical resolution can be drawn from this case in isolation. The reduction in the number of vertical levels has changed the (inherently uncertain) detail in the forecasts. It has not changed the essence of the forecast. This is not surprising since the larger-scale upper-level flow played such an important role in this case and that was largely unchanged. The run with 38 levels still had a convergence line prior to the formation of the storms (just like the 76-level run) and is the reason for the high totals being focussed in the London area in both forecasts. 76 levels is still considered a better option in general.

4.1.2 The use of the convection scheme

The operational UK 4 km model uses the convection parametrization scheme (Gregory and Rowntree 1990 +enhancements) to represent the effect of showers that can not be resolved on the grid. To make sure that the convection scheme does not operate too strongly and inhibit the model from representing the larger showers on the grid, it has been restricted according to Roberts (2003).

In a 1 km model the convection scheme is usually switched off because it is thought that the grid spacing is small enough to resolve the showers sufficiently well. However, even on a 1 (or 1.5) km grid there will be clouds that can not be resolved and there will also be mixing of the air within the clouds that is also not resolved.

In these experiments the convection scheme was used in the 1 km forecast from 09 UTC (larger domain) with differing restrictions:

Experiment 1: No convection scheme (Labelled 'No conv scheme')

Experiment 2: Strongly restricted convection scheme - same as in UK 4 km model (Labelled: 'Very weak conv scheme')

Experiment 3: Less strongly restricted convection scheme (Labelled: 'Weak conv scheme')

Experiment 4: Only slightly restricted convection scheme (Labelled: 'Strong conv scheme'

Experiment 5: Convection scheme operating without the restriction applied. This is the same as used in the old 12 km mesoscale model (Labelled: 'Strong conv scheme (30 min). The '30 min' refers to a 'CAPE closure timescale' of 30 minutes – see Roberts 2003.

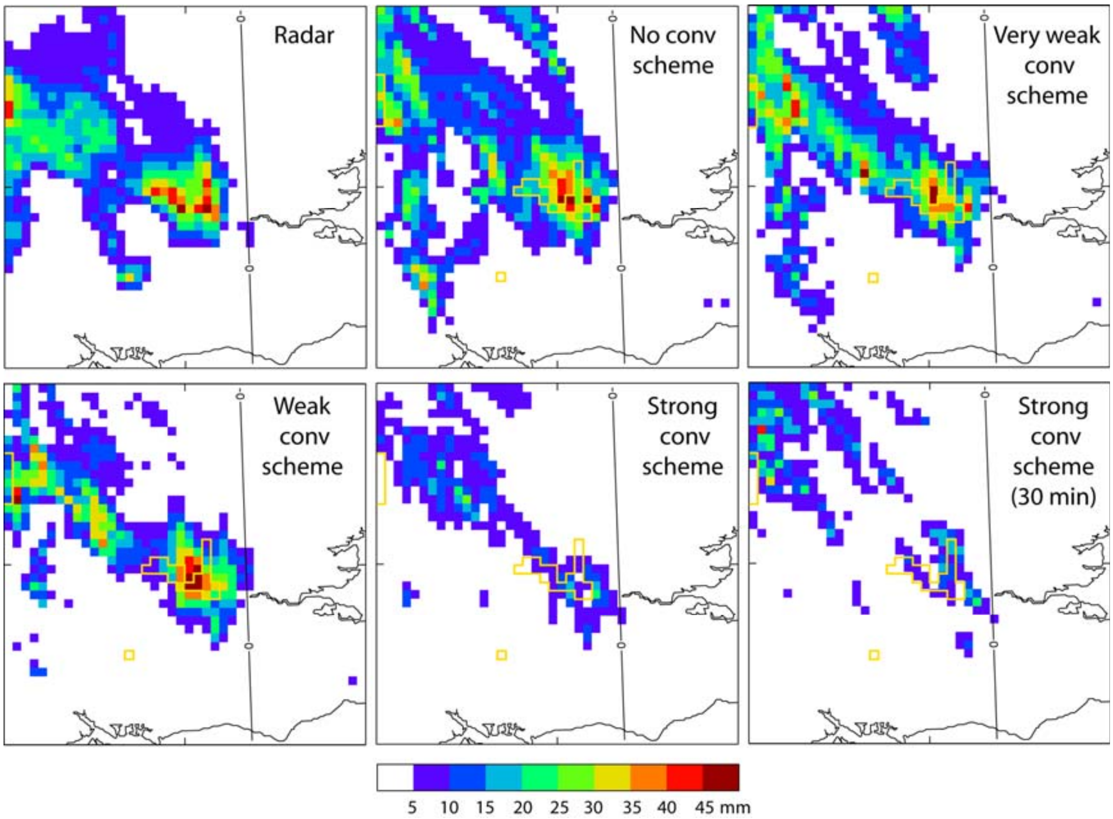


Figure 4-2. Rainfall accumulations over the period 13 to 18 UTC, 3rd August 2004 from (top left) radar, and a collection of 1 km forecasts from 06 UTC with different uses of the convection parametrization scheme (refer to text). The orange line in all but the top left panel encloses the totals above 30 mm from radar (over and to the west of London).

Figure 4-2 shows the differences between the forecasts in terms of rainfall amounts. When the convection scheme was used in the same way as in the 4 km model (top right panel) there was a decrease in the number of pixels with totals > 5 mm compared to the run with no convection scheme, but the high totals were maintained. When the convection scheme was restricted a little less (bottom left) there was a further decrease in the number of pixels > 5 mm, but again the high totals were maintained (particularly in the London area). So, when the convection scheme was strongly restricted, the overall rainfall pattern did not change very much and high rainfall totals were still been produced. There were variations on small scales (a few grid squares); in the particular locations of the highest totals, but the areas with high rainfall did not change. A benefit of using the restricted convection scheme was a reduction in the number of small 'spotty' showers that initiated too widely and too early in the run with no convection scheme.

In contrast, when the convection scheme was not restricted much or at all (bottom centre and bottom right), the amount of rain was dramatically reduced, particularly the high totals in the London area. In those experiments the convection scheme acted to remove much of the convective instability before the model could generate showers on the grid. The convection scheme tended to produce rainfall totals over too wide an area that were too small to show up in Figure 4-2.

The results indicate that the use of a restricted convection scheme is not harmful in this case if it is not allowed to dominate the solution and may even be beneficial for representing convective clouds and turbulence that can not be resolved. It's probably not the best way of doing this however; other sub-grid-scale turbulence scheme are likely to work better since the convection scheme is not designed for such a fine grid, but it does give insight into the model response to varying amounts of parametrised convection.

Quantitative verification

The Fractions Skill Score (FSS) measure (Roberts & Lean 2008) has been used to compare the forecast rainfall accumulations with accumulations measured by radar. The method is described in appendix 1. The use of the FSS to compare with radar gives a quantitative assessment of forecast skill at different spatial scales. Forecast skill should typically improve with spatial scale in the sense that it should be easier to predict whether rain will fall somewhere within a large area than a smaller one (e.g. a county rather than a town). The FSS can also been used to compare a forecast with other forecasts. The comparison of one forecast with another provides information about the scales over which the forecasts agree best and therefore it also shows the scales that have been affected most by a particular change to the model.

The FSS can vary between 0 and 1. The closer to 1 the better the agreement between the forecast and radar (or another forecast). To get an understanding of what constitutes a good score or a bad score; it was found in an idealised setup (very similar to the one presented by Roberts & Lean (2008)) that the scale at which the FSS= ~0.5 is the scale at which a forecast is correct a third of the time (i.e. in the areas where rain is either forecast or observed the forecast is correct over one third of the domain at that scale). In terms of categorical verification statistics it is the same as saying the Critical Success Index CSI=0.33. The same setup showed that the scale at which the FSS=0.8 is the scale at which a forecast is correct two thirds of the time (two thirds of the domain is correct at that scale, CSI=0.67). Therefore, it is reasonable to construct as a working guideline that for a given spatial scale a value of FSS < 0.5 means that the forecast is in poor agreement with whatever it is being compared with (radar or another forecast) at that scale, for 0.5 < FSS < 0.8 the forecast is in reasonable agreement, and for FSS > 0.8 the forecast is in good agreement (but still with some differences).

Figure 4-3 shows how forecast skill (using the FSS) varied with spatial scale for the top 2% of accumulations shown in Figure 4-2. Panel (a) shows scores for the 'weak conv scheme' (bottom left in Figure 4-2) and panel (b) shows scores for the 'Strong conv scheme (30 min)' (bottom right in Figure 4-2).

Starting with Figure 4-3(a). At smaller scales (up to ~20 km) the forecast with no convection scheme (red line) and the forecast with a 'weak convection scheme' (blue line) had almost exactly the same skill when compared with radar. The forecast with the weak convection scheme was a little more skilful at the very smallest scales, but both had poor skill at those scales. At scales larger than ~20 km both forecasts had high skill (FSS>0.8), and the forecast with no convection scheme was more skilful.

The green line shows the FSS comparison between the two forecasts (rather than compared to radar). At the smallest scales (< 15 km) they were in reasonable agreement and at scales larger than 15 km they had a good agreement (FSS>0.8). This says that although the two forecasts differed most at the smallest scales, the differences had little impact on the skill (when compared to radar) at those scales. In other words, the fine-scale detail was behaving like noise in both forecasts. However, at larger scales, despite the forecasts being very similar, the differences that there were did contribute to some differences in skill (the reference forecast with no convection scheme being more skilful).

In contrast, the green line in Figure 4-3(b) shows that the forecast with the unrestricted convection scheme was in poor agreement with the reference forecast (FSS < 0.5) at scales as large as ~75km and did not have a good agreement (FSS <0.8) until ~140km. It was also considerably less skilful compared to radar at all scales (blue line well below red line), with values of FSS less than 0.5 even at a scale of 70 km. This reflects the differences seen in Figure 4-2 that the changes in the forecast brought about by introducing the standard unrestricted convection scheme were large and detrimental. In this case, and probably in general, it is not a good idea to use a convection parametrization scheme in this way in a ~1 km model.

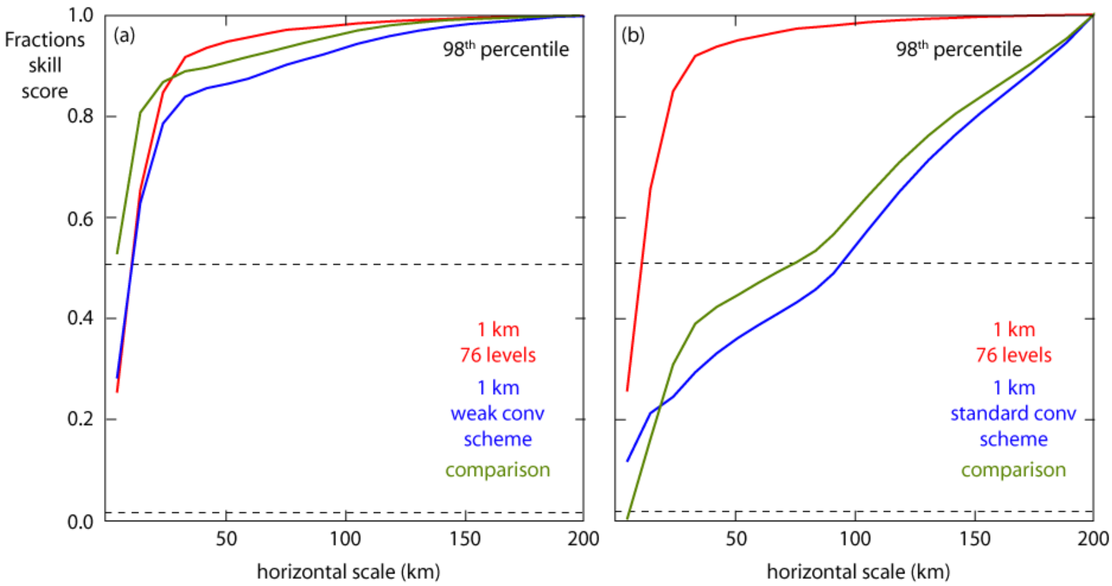


Figure 4-3. Graph of forecast skill using the Fractions skill Score (FSS) against spatial scale for the rainfall accumulation period shown in Figures 4-1 and 4-2 for (a) the 1 km forecast from 06 UTC compared with radar (red line), the 1 km forecast with weak convection scheme (blue line) compared with radar and the comparison between the two forecasts (green line). (b) the 1 km forecast from 06 UTC compared with radar (red line), the 1 km forecast with an unrestricted convection scheme (blue line) compared with radar and the comparison between the two forecasts (green line). The better the fit between forecast and radar or two forecasts the higher the value of FSS. The upper dashed line shows the minimum FSS value for a reasonable fit. The lower dashed line shows the FSS for a comparison with a randomly generated forecast.

4.2 Case E: 19th June 2005

Three sensitivity experiments are discussed for this case:

1. Change from 76 to 38 vertical levels
2. Change to use of 'radiation on slopes' (see below)
3. Inclusion of the convection scheme as used at 4 km

Radiation on slopes means that the slope of the ground is taken into account when the sun is shining, so that a slope facing the sun will be warmed more than a slope facing away. This change has gone into the latest versions of the model because it is more realistic. It wasn't an issue for older coarser resolution models in which the slopes in the orography dataset were always very shallow. The purpose of this experiment is to see what sort of difference it makes to convective rain in a hilly area because it appears that, in this case, the North York Moors played a role.

The three experiments are only briefly analysed because these types of changes were not going to correct the major problem with the forecasts of this event that was due to the error in the larger-scale dynamics. Nevertheless, the results from the objective verification are interesting and show what can be deduced by using the FSS to investigate the impact of model changes.

The rainfall accumulations in Figure 4-4 show that the three changes to the model did make some differences to the details in the rainfall pattern, but the overall signal remained essentially the same in all three. The forecast with 38 levels appears to be the most different from the reference 76-level forecast. The run with the radiation on slopes included was the most similar to the reference forecast.

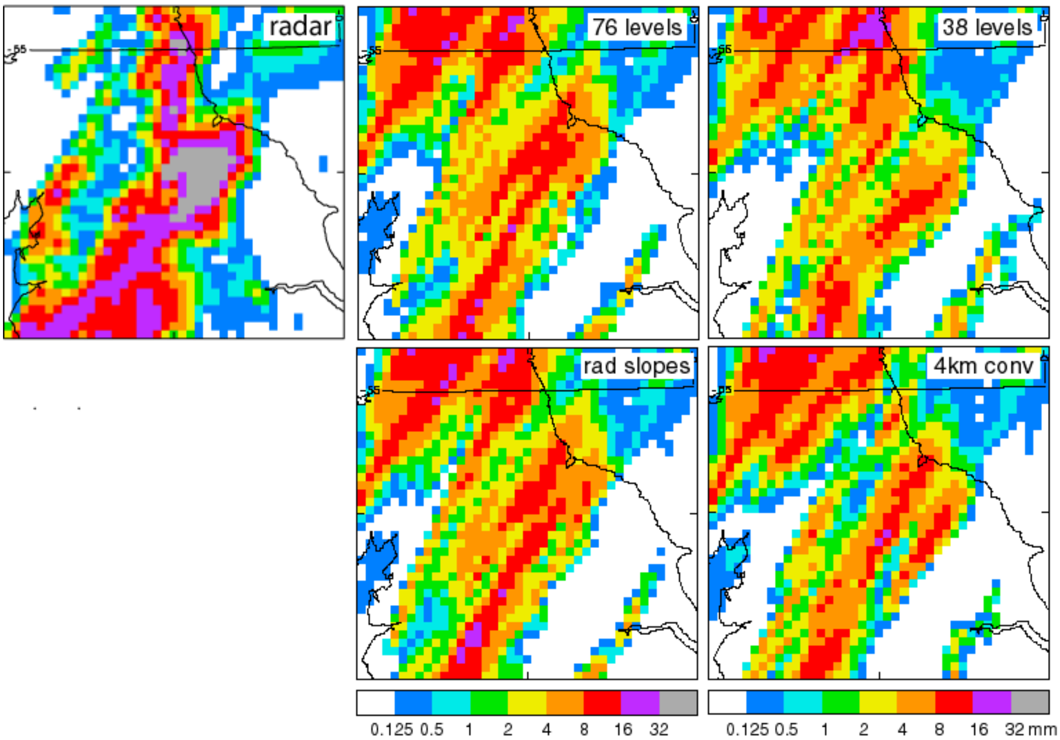


Figure 4-4. Rainfall accumulations over the period 13 to 17 UTC 19th June 2005 from (a) radar and 1 km forecasts from 06 UTC (b) reference forecast, (c) forecast with 38 vertical levels instead of 76, (d) forecast with radiation on slopes included (see text), and (e) forecast with the convection scheme included as used in the UK 4 km model.

The scale-selective verification can shed more light on the impact of each of the changes. Figures 4-5 shows the variation in skill with spatial scale for the experiment to half the number of vertical levels. It is the change that seemed to have had the biggest effect. Figure 4-6 shows the same type of graph for the radiation on slopes, which appears to have had the least impact of the three.

The blue and red lines in Figure 4-5 show that for the highest 2% of rainfall totals (98^{th} percentile threshold), both the reference 76-level forecast or the 38-level forecast had poor skill at most scales (FSS < 0.5 up to ~120 km) . They were both equally poor and neither was more accurate than a purely random (but unbiased) forecast at scales up to ~50 km. This was expected because of the error in the larger-scale dynamics discussed in section 3. What's interesting is that the green line, showing the comparison between the two forecasts, reveals that they were in poor agreement at the smallest scales (FSS < 0.5 at scales < ~10 km) and then had only reasonable agreement (0.5 < FSS < 0.8) up to around 60 km yet these differences do not translate into a change of forecast skill against radar (red and blue lines are nearly the same). This shows that the impact of the change to the vertical resolution is on scales shorter than the main gross error. It also reveals that the 38-level forecast is in effect behaving like a random rearrangement of the 76-level forecast over most scales (particularly the smaller scales), which neither added or removed forecast skill on this occasion.

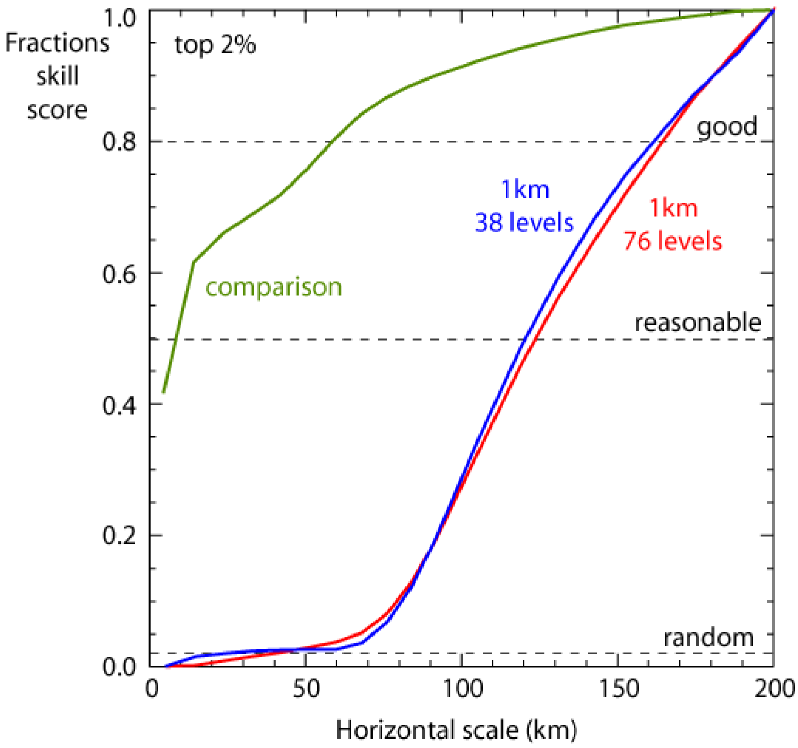


Figure 4-5. Graph of forecast skill using the Fractions skill Score (FSS) against spatial scale for the rainfall accumulation period shown in Figure 4-4 for (a) the 1 km forecast from 06 UTC with 76 levels compared with radar (red line), the 1 km forecast with 38 levels compared with radar (blue line) and the comparison between the two forecasts (green line). The better the fit between forecast and radar or two forecasts the higher the value of FSS. The upper dashed line shows the minimum FSS value for a reasonable fit. The lower dashed line shows the FSS for a comparison with a randomly generated forecast. The other dashed lines are described in the text.

Now consider the radiation on slopes experiment. The continuous green line in Figure 4-6 indicates that the run with radiation on slopes was in good agreement with the reference forecast at scales greater than ~25 km. They only really differed to the extent that they could be labelled as having 'reasonable' rather than 'good' agreement at scales less than ~20 km (FSS < ~0.8). In other words, there wasn't very much difference between the forecasts at most scales. Contrast this with the comparison between the reference and 38-level forecasts (taken from Figure 4-5 and shown by the dashed line in Figure 4-6), which showed greater differences (lower values of FSS) at all scales less than ~120 km. Yet, despite the differences between the reference forecast and the forecast with radiation on slopes being less, those differences nevertheless translated into an improved forecast when radiation on slopes was included (blue line above red line at all scales). Recall from Figure 4-5, that 38-level forecast, although more different to the reference forecast, had no impact on the forecast skill.

It appears that the differences the inclusion of radiation on slopes made to the forecast (even though not large) had a positive impact on skill at all scales. It shows that (for this case anyway) the inclusion of radiation on slopes was beneficial; that the changes to the rainfall were linked to improvements in skill rather than more random in nature. This is even despite the forecast being so poor to start with and there being no correction of the problem with the upper-level jet stream. It suggests that variations in solar heating over hills played a part in this case and may do in other situations with convection in hilly areas.

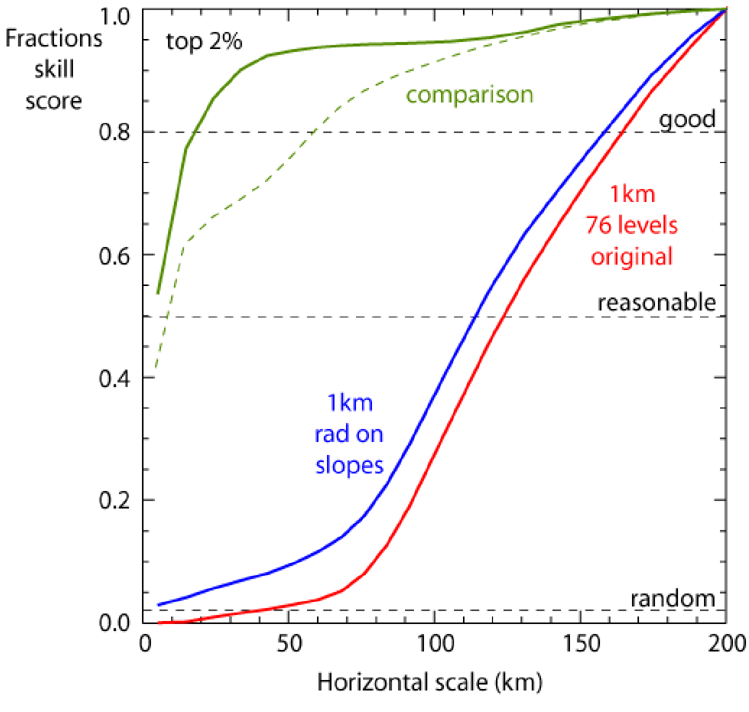


Figure 4-6. Graph of forecast skill using the Fractions skill Score (FSS) against spatial scale for the rainfall accumulation period shown in Figure 4-4 for (a) the 1 km forecast from 06 UTC with 76 levels compared with radar (red line), the 1 km forecast with radiation on slopes (see text) with radar (blue line), the comparison between the two forecasts (green line), and the comparison between the forecasts in Figure 4-5 (dashed green line). The better the fit between forecast and radar or two forecasts the higher the value of FSS. The upper dashed line shows the minimum FSS value for a reasonable fit. The lower dashed line shows the FSS for a comparison with a randomly generated forecast. The other dashed lines are described in the text.

4.3 Final comments

The results from the sensitivity experiments have shown the benefit of using the FSS approach to compare forecasts. The method provides a means of diagnosing the scales that are impacted most by a change to the model formulation and whether skill is added at those scales.

The approach could be used in future to compare members from a high-resolution ensemble to determine the predictability of particular rainfall features within an ensemble.

5. Detailed study 3: The Albrighton flood, 4th July 2006.

Part 1. Operational forecasts and high-resolution simulations

This is the first of three sections about this case. It was chosen to be studied in stage 3 of the project for two reasons. (1) The meteorology differed significantly from the events studied in stage 2 (section 3); the storm was much more localised and did not develop in association with larger-scale rotation in the upper troposphere. (2) The 1 km model had some success at representing the storms, which meant that an investigation of the meteorology and model behaviour would be meaningful.

In this section the event will be described and then a variety of operational and post-event forecasts will be shown. The operational forecasts had a horizontal grid spacing of either 12 or 4 km. The post-event forecasts were run with a grid spacing of 4 or 1 km.

In the next section the meteorology of the storm and the 1 km model forecasts will be investigated more closely.

In the third section the impact of sensitivity studies on the 1 km (or 1.5 km) forecasts will be examined.

5.1 What happened

An intense localised thunderstorm caused flooding of homes in the Albrighton area on the evening of the 4th July 2006. Albrighton is situated just south of the M54 midway between Wolverhampton and Telford in the West Midlands. The storm began at around 17.30 UTC (6.30 pm local time) and was most intense over the area of interest between 18.10 and 18.40 UTC and then again between 18.50 and 19.15 UTC. An Environment Agency (EA) report about the event stated that there were no official raingauges within the small catchment of Albrighton Brook (~5km^2) where the flooding occurred. The nearest EA gauge at the nearby village of Cosford recorded 42mm. A private raingauge within the catchment recorded 90mm in ~2 hours. The limited number of gauges means that radar information is needed to complete the picture. Again, from the EA report a radar pixel (Nimrod quality controlled) within the catchment showed a rainfall total of 171mm.

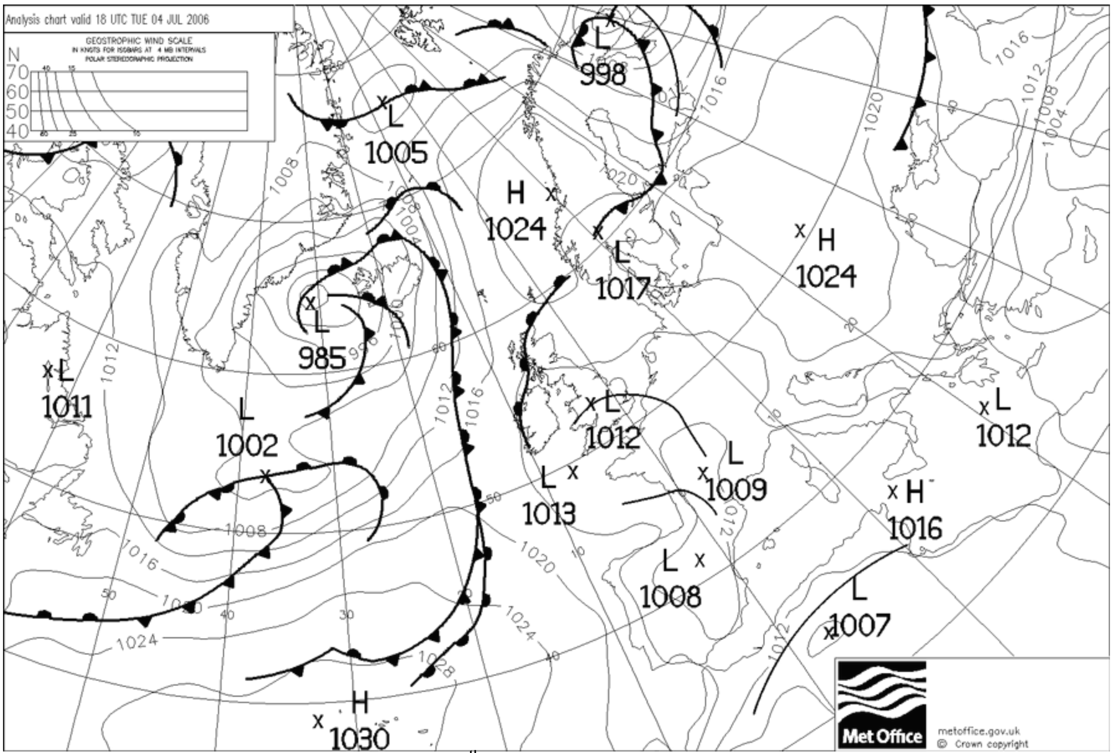


Figure 5-1. Synoptic chart for 18 UTC 4th July 2006.

The storms developed in a hot and humid air that had recently arrived from the near continent. The winds were light and mostly from the south or southeast.

Hourly rainfall accumulations for the period of interest from radar on a 5 km grid are shown in Figure 5-2. The 4-hour accumulations over the period 17 to 21 UTC (Figure 5-3) reached a maximum of nearly 50 mm over a 5 km square. This is lower than the high gauge totals and fine-scale radar totals because the very high rainfall amounts were very localised. A sequence of rainfall rate snapshots is shown in Figure 5-4.

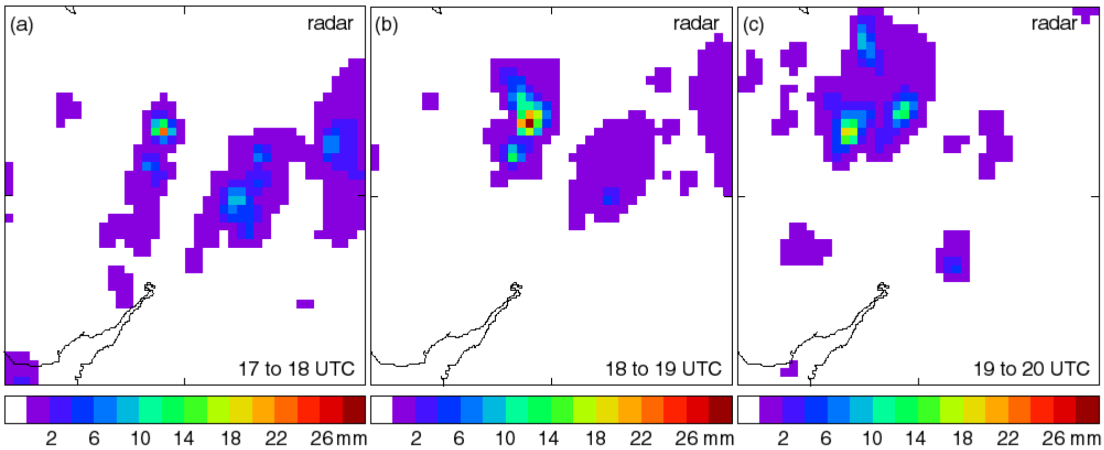


Figure 5-2. Hourly accumulations from radar, 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on a 5-km grid.

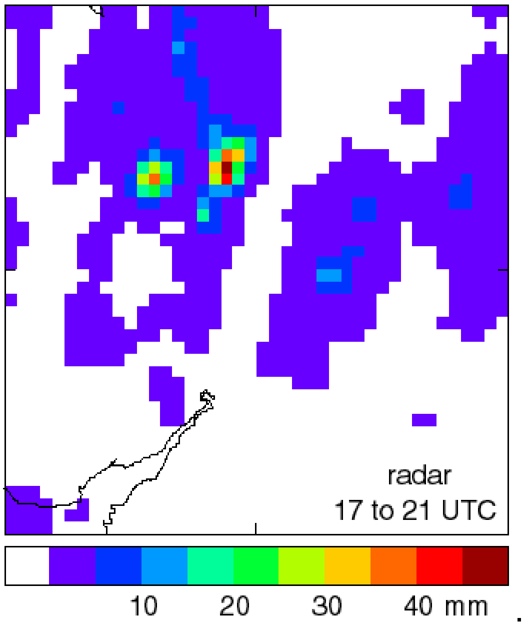


Figure 5-3. Rainfall accumulations from radar over the period 17 to 21 UTC on a 5-km grid.

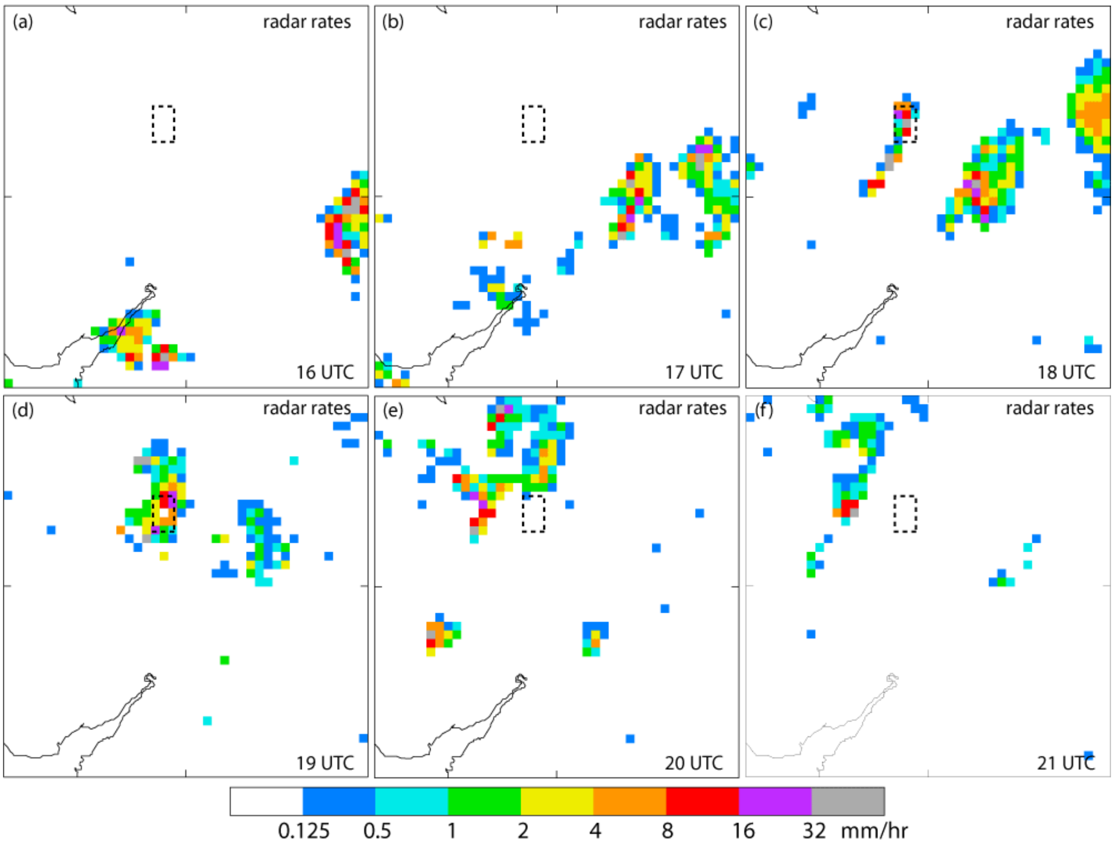


Figure 5-4. Rainfall rates from radar at 16, 17, 18, 18, 20 and 21 UTC on a 5-km grid. The dashed rectangle encloses the area of highest rainfall totals over the period 17 to 21 UTC.

5.2 The forecast configurations

This is the sort of very intense localised storm that we would expect the current operational models with a grid spacing of 12 km to fail to predict because a storm that small can not be resolved. The UK model with a grid spacing of 4 km should be better, but even that resolution is not theoretically fine enough.

Output from a variety of forecasts will now be presented. Firstly the operational forecasts at the time and then re-runs of high-resolution simulations. The forecasts presented are:

1. The then operational mesoscale model forecasts (12 km grid spacing) from 06 and 12 UTC.
2. The operational North Atlantic and European (NAE) model forecasts (12 km grid spacing) from 06 and 12 UTC.
3. The now operational UK model (4 km grid spacing) forecasts from 09 and 15 UTC.
4. Re-runs using the UK 4-km model from 06 and 12 UTC.
5. Re-runs using a regional 300x300 km 1-km model from 06 and 12 UTC.

The model domains are shown in Figure 5-5.

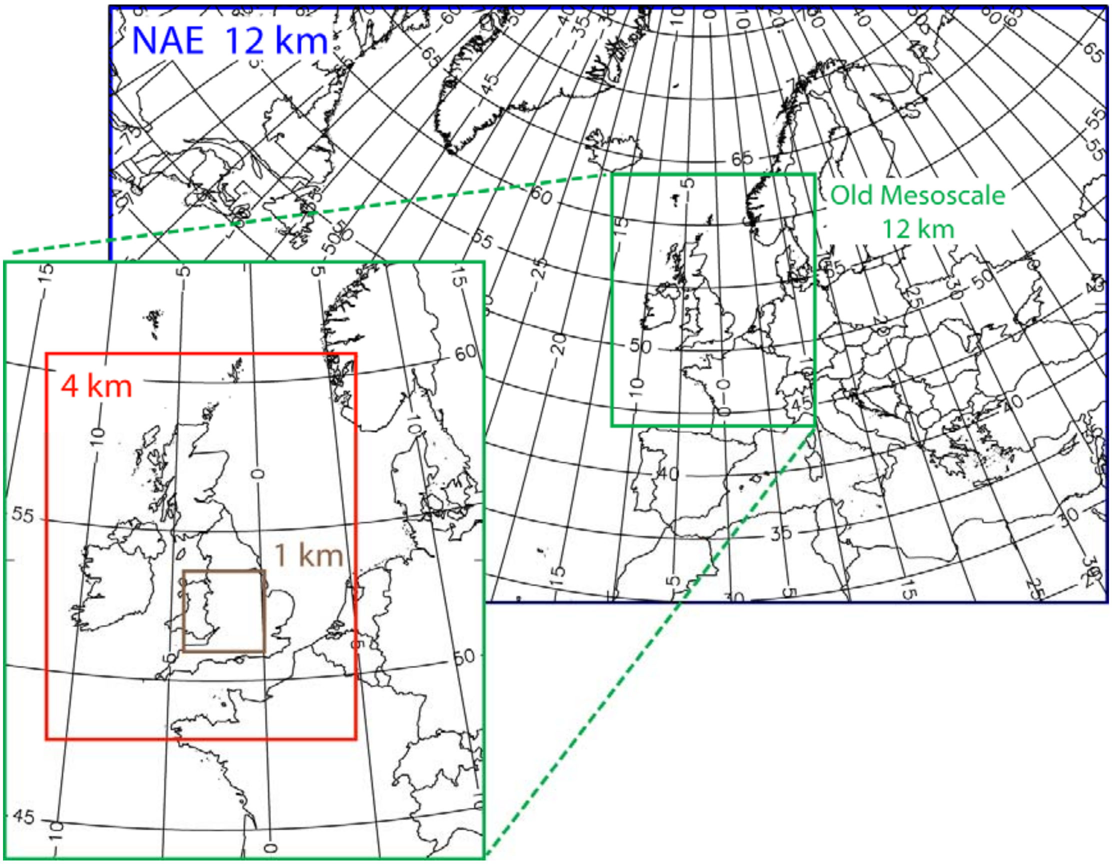


Figure 5-5. Domains for the operational North Atlantic European (NAE) 12 km model, the old mesoscale model (12 km), the operational 4 km model and the 1 km model set up for this study.

5.3 The operational forecasts

5.3.1 The Mesoscale model forecasts

The mesoscale model was the highest-resolution operational model in the Met Office for several years. It was operational at the time of this event, but has now been superseded the North Atlantic European (NAE) model (12 km grid spacing). It had a grid spacing of 12 km and covered an area that included the UK, North Sea and much of France (Figure 5-5). The forecasts shown started from 06 and 12 UTC and included data assimilation to make the starting point (analysis) as close a fit to the available observations as possible. Information at the boundaries came from the operational global model.

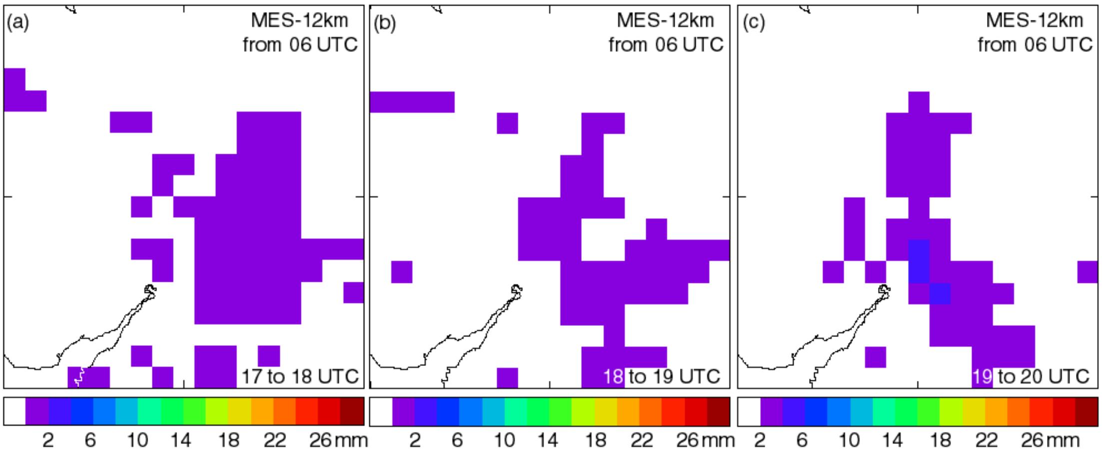


Figure 5-6. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from the mesoscale model forecast starting at 06 UTC.

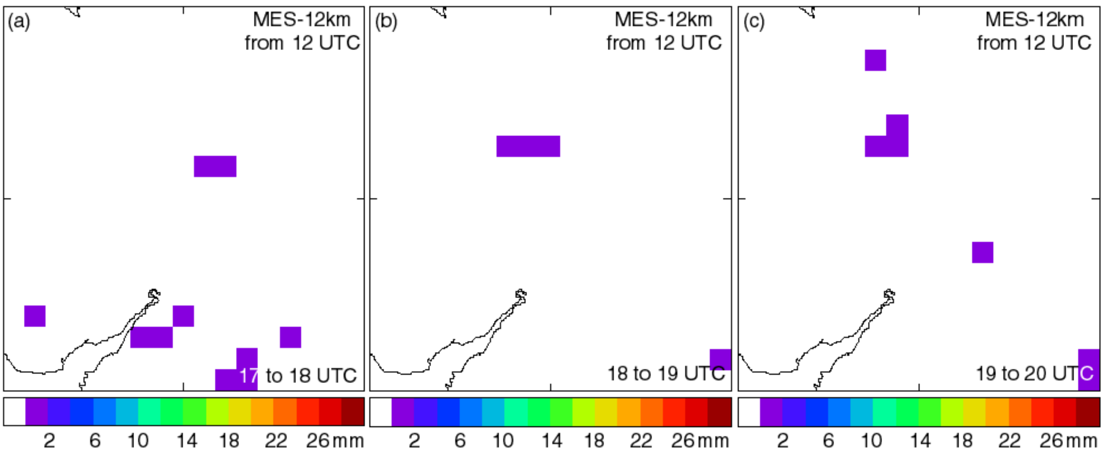


Figure 5-7. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from the mesoscale model forecast starting at 12 UTC.

Figures 5-6 and 5-7 show that neither the 06 or 12 UTC mesoscale model forecasts were able to produce more than a few millimetres of rain. Both forecasts failed to produce the storm (or anything even close) because the model has insufficient resolution. The rainfall that was generated came from the convection parametrization scheme (Gregory and Rowntree 1990) that is designed to represent the effect of unresolved showers, but by its nature, will tend to give a bland field of widespread low-precipitation amounts because it is representing the average effects of convection rather than individual storms.

5.3.2 North Atlantic European (NAE) model

The North Atlantic European (NAE) has a grid spacing of 12 km and covers much of the North Atlantic and Europe. The forecasts shown started from 06 and 12 UTC and included data assimilation to make the starting point as close to observations as possible. Information at the boundaries came from the operational global model.

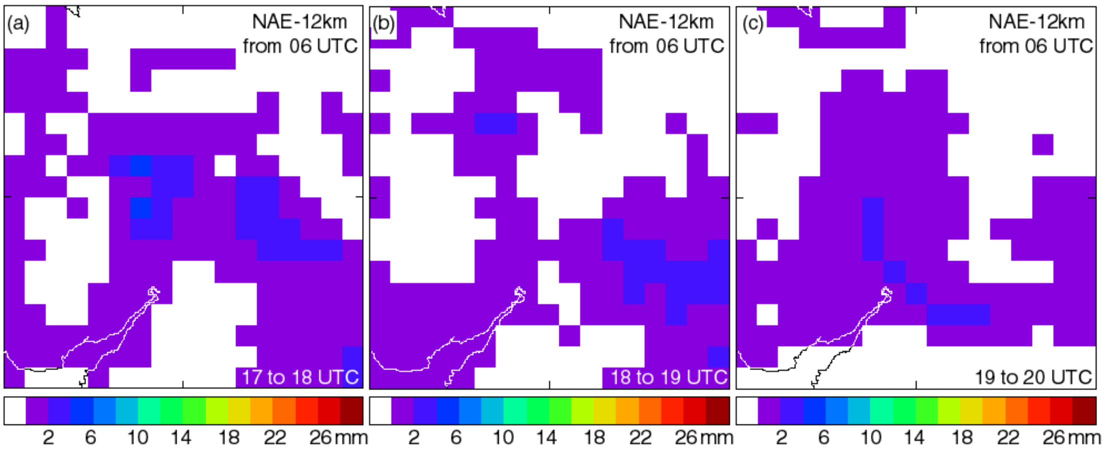


Figure 5-8. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from the NAE model forecast starting at 06 UTC.

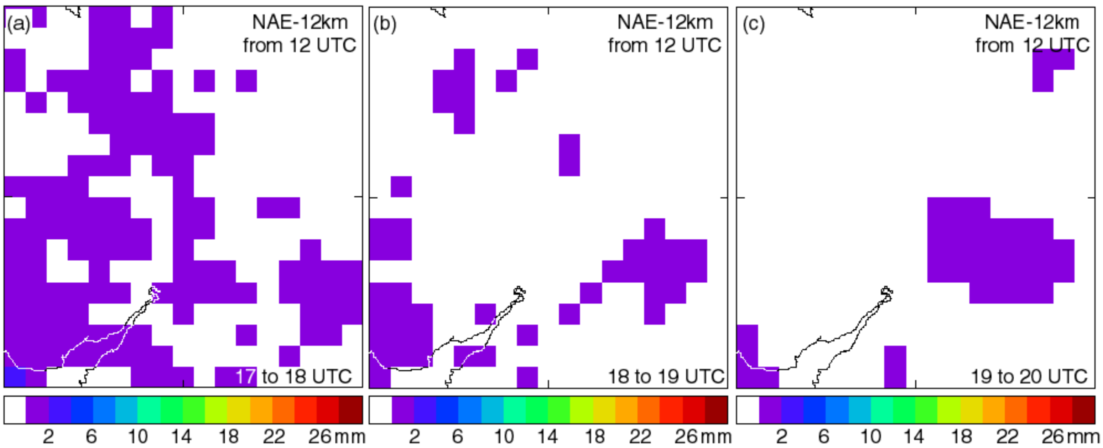


Figure 5-9. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from the NAE model forecast starting at 12 UTC.

Figure 5-8 and 5-9 show that, as for the mesoscale model, neither the 06 or 12 UTC NAE model forecasts were able to produce more than a few millimetres of rain. The reason is the same – the precipitation came from the convection parametrization scheme rather than being resolved because the 12-km grid spacing is not sufficient to represent this type of storm.

5.3.3 UK 4-km model

The UK 4-km model has replaced the mesoscale model as the finest-resolution model run operationally at the Met Office. It has a grid spacing of 4 km and covers the UK. The forecasts shown started from 09 and 15 UTC and included data assimilation to make the starting point as close to observations as possible. Information at the boundaries came from the NAE model.

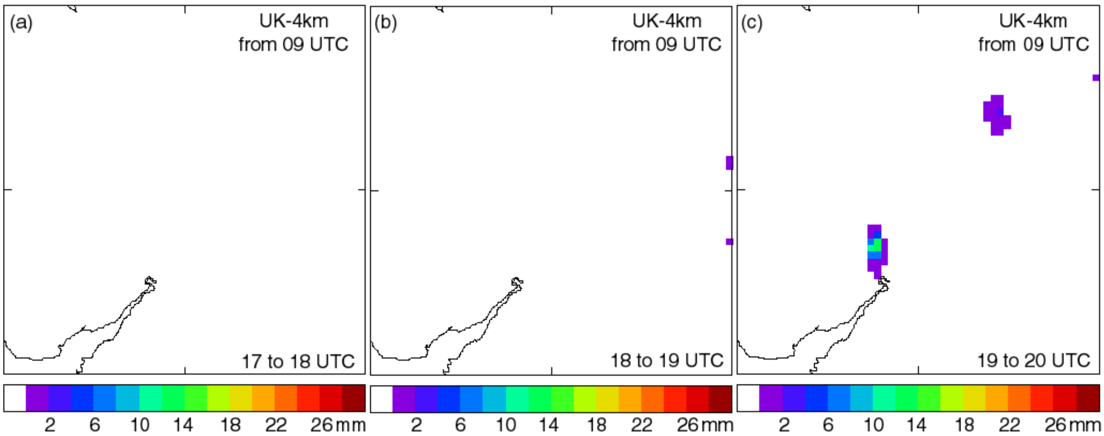


Figure 5-10. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from the UK 4-km model forecast starting at 09 UTC.

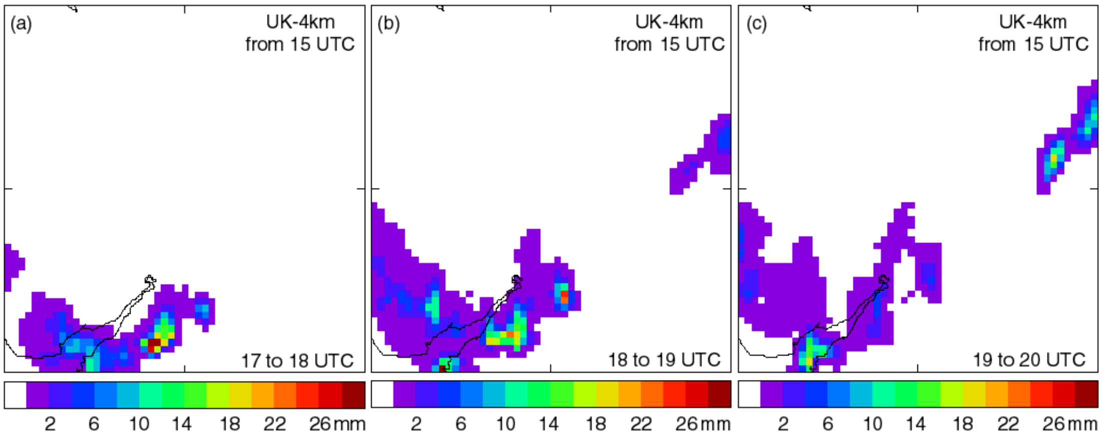


Figure 5-11. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from the UK 4-km model forecast starting at 15 UTC.

The two UK4-km model forecasts were very different. The first, from 09 UTC, (Figure 5-10) produced essentially no rainfall anywhere in the area shown at the time when the storm occurred (17 to 19 UTC). It did, however develop a storm to the south after 19 UTC and some significant rain at later times (not shown). The model was not able to develop resolved convection quickly enough. This is a typical deficiency of a 4-km model when it is trying to represent convection that is small and not strongly forced by the larger-scale dynamics. The 4-km grid is not fine enough for this type of situation. The 4-km model does have a convection parametrization scheme operating, but it is restricted so that it can only produce very small amounts of rain (not seen on these pictures). Experience from previous experiments has shown that if the convection scheme had been made more active it would probably not have helped because the forecast would have just tended towards the behaviour of the 12-km forecasts.

The second forecast, from 15 UTC, (Figure 5-11) produced realistic and intense storms at the right time, but they were in the wrong place (by >80km). The model was probably able to generate storms to start with because they existed at the start of the forecasts and put in by the data assimilation system. However, the model was unable to produce new storms further north. This is likely to have been a result of insufficient resolution again and perhaps an environment that was less likely to develop convection. All the models we have seen so far had less convection in the later forecasts.

5.4 The post-event forecasts

5.4.1 Post-event forecasts – 4 km model

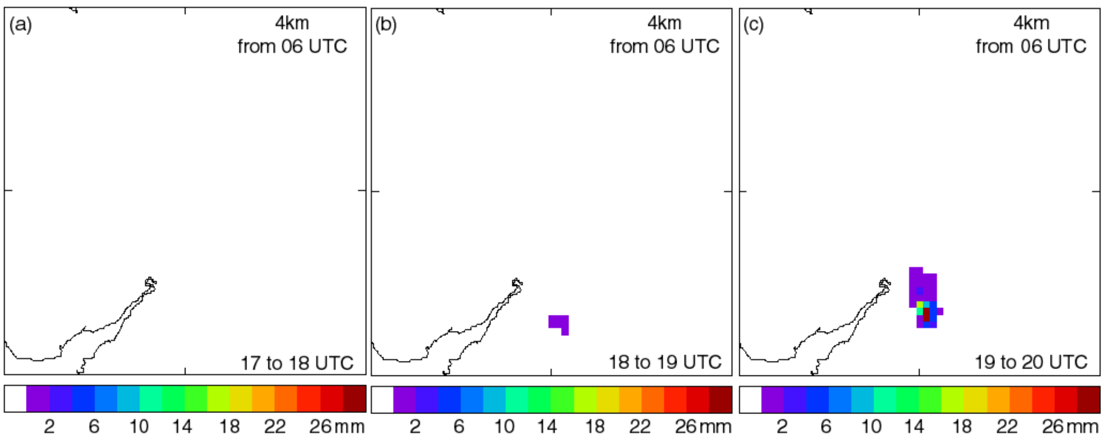


Figure 5-12. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from a run of the UK 4-km model forecast starting at 06 UTC.

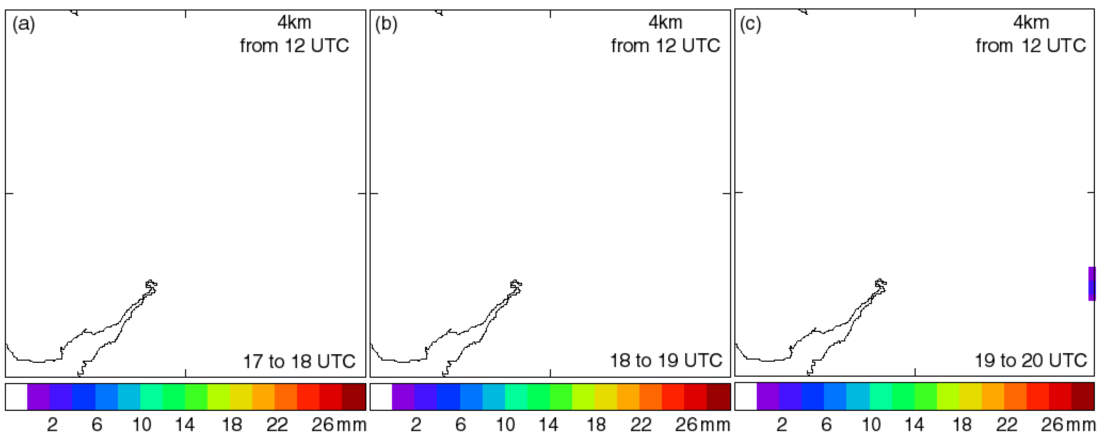


Figure 5-13. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from a run of the UK 4-km model forecast starting at 12 UTC.

The post-event runs have used the same 4-km domain as the operational UK 4km model. The difference is that the forecasts started off as 12-km mesoscale model forecasts from 06 and 12 UTC and were interpolated to 4-km an hour later at 07 and 13 UTC respectively. This means that there was no data assimilation on the 4-km grid (although there was at 12 km during the operational cycle). The 4-km model used information from the 12-km mesoscale model forecasts to provide information at the boundaries.

The post-event 4-km forecast from 06 UTC was very similar to the operational UK 4-km forecast from 09 UTC. There were no storms between 17 and 19 UTC, but then one developed too far south in the subsequent hour (although not in the same place). It is not surprising that those two forecasts should be similar as they start only 3 hours apart and have the same resolution-dependent behaviour.
The post-event 4-km forecast from 12 UTC produced no storms through the period shown (or shortly after). It appears that the conditions were less favourable for convection to break out in this forecast compared to the 06 UTC run. This trend is consistent with there being less rain in the 12 UTC mesoscale forecast compared to that at 06 UTC. The storms in the 15 UTC UK 4-km forecast are not seen in this 4-km forecast because at 12 UTC there were no storms in reality for the data assimilation system (operating at 12-km only) to latch on to.

5.4.2 Post-event forecasts - 1 km model

The 1-km model domain used was 300x300 km and covered a little more area than shown in Figures 5-14 to 5-16. The forecasts started off as 12-km mesoscale model forecasts from 06 and 12 UTC and were interpolated to the 1-km grid an hour later at 07 and 13 UTC respectively. This means that there was no data assimilation on the 1-km grid (although there was at 12 km during the operational procedure). The 1-km forecasts used the 4-km forecasts (see above) to provide information at the boundaries.

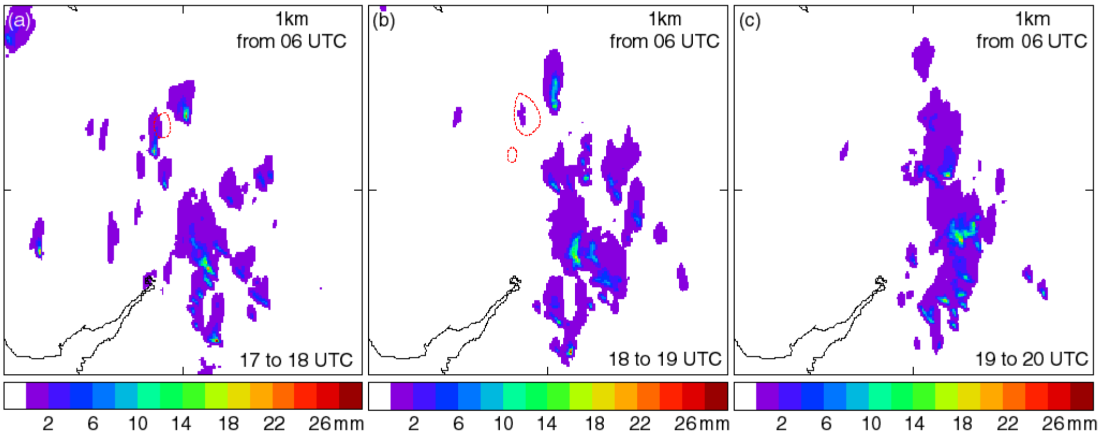


Figure 5-14. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from a run of the 1-km model forecast starting at 06 UTC.

The 1-km forecast from 06 UTC (Figures 5-14 and 5-15) generated storms in the domain. It was able to produce two storms close to where the observed storm occurred. This is a significant improvement on the other forecasts which either produced a broad area of light rain (12km) or failed to produce any storms or delayed and wrongly positioned them (4km). However, the rainfall totals were too low and the storms were too small. The 1-km forecast had a peak pixel of ~26 mm compared to ~47 mm in the radar on the 5-km grid. The radar on a 1-km grid recorded more than 100 mm. Therefore, the forecast generated rainfall totals of somewhere between 4 and 8 times less then observed. This sort of under-prediction of the peak amount and storm extent has been seen in the 1-km model on other occasions even though the domain average rainfall may be reasonable or even too high. The showers broke out too widely and were too small (the ‘measles’ effect) – even when taking into account that the showers observed by radar on the 5-km grid appear larger than they really were. The showers in the 1-km forecast did not persist for long enough or grow large enough to produce the large rainfall totals required. They were intense though (up to ~100mm/hour in the cell cores), which means that the model was able to produce some significant totals very locally.

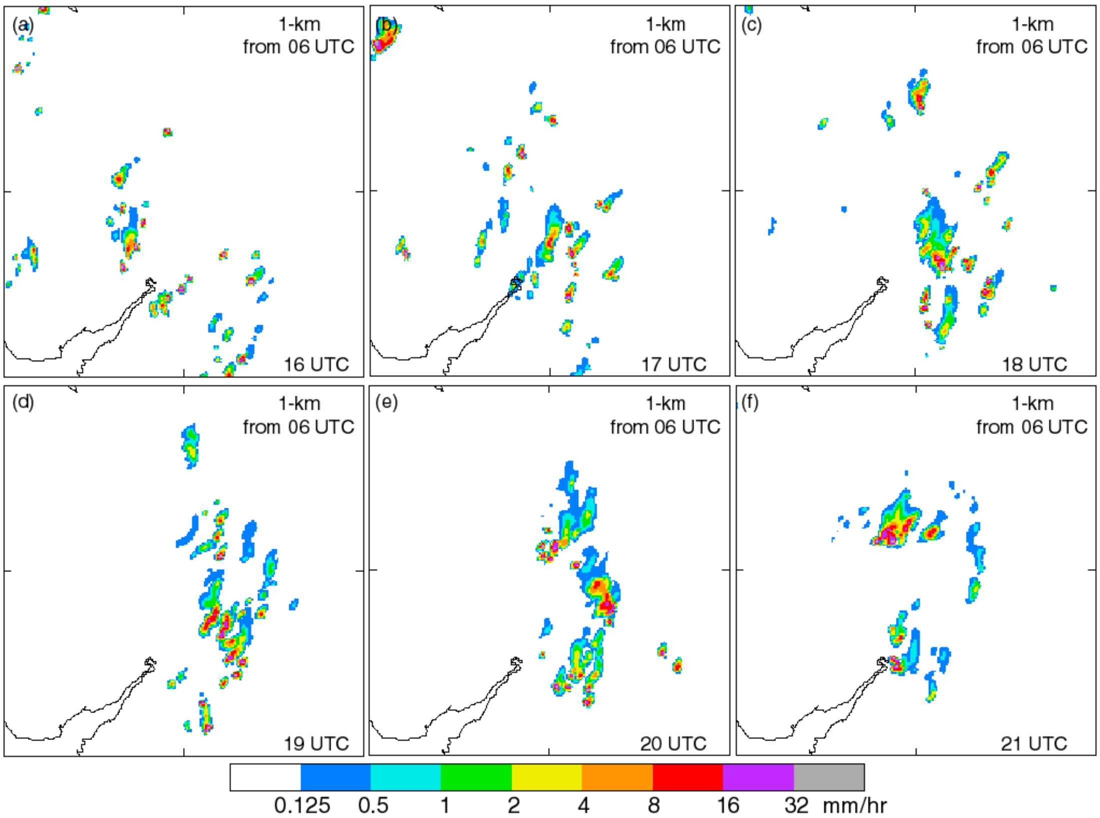


Figure 5-15. Rainfall rates at 16, 17, 18, 19, 20, 21 UTC on 4th July 2006 from a run of the 1-km model forecast starting at 06 UTC.

The 1-km forecast from 12 UTC (Figure 5-16) produced far less rain than the 06 UTC forecast. This is in line with the other resolutions and may be to do with a less convective environment in the model in the later forecasts. It did not produce significant rain in the area of interest.

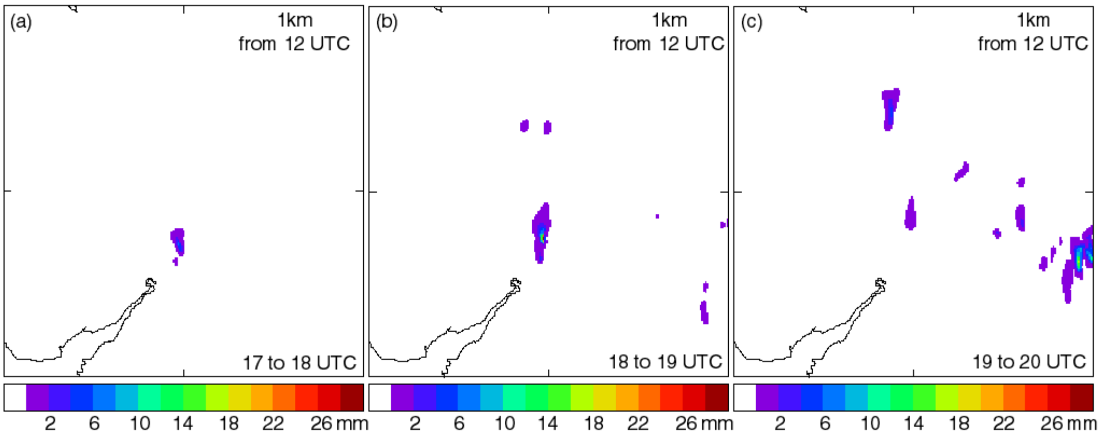


Figure 5-16. Hourly rainfall accumulations over the periods 17 to 18 UTC, 18 to 19 UTC and 19 to 20 UTC on 4th July 2006 from a run of the 1-km model forecast starting at 12 UTC.

5.5 Overview of the forecasts

- The 12-km forecasts, both NAE and mesoscale models, were unable to produce a storm because they have insufficient resolution to represent this type of situation and have to rely on a convection parametrization scheme to represent the effect of the unresolved showers.

- The 4-km forecasts did produce storms with high rainfall totals, but they developed far too late and in the wrong place. The 4-km grid is also too coarse for this type of event.

- The 1-km model forecast from 06 UTC was able to produce very heavy showers and gave high totals close to where the storm occurred at about the right time. However, the showers tended to be too small and transient to produce the very high rainfall totals observed. Nevertheless the 1 km model gave by far the best forecast of all the resolutions and is investigated further in section 6.

6. Detailed study 3: The Albrighton flood, 4th July 2006.

Part 2. The meteorology and 1 km forecasts.

We now focus on the 1 km forecasts from 06 UTC and 12 UTC. This section uses the output from those two forecasts as well as observations to explain why the storms developed, why the 06 UTC forecast generated more showers and why large rainfall totals were possible.

6.1 The meteorological conditions before the storms

The Albrighton flood came about because of torrential rain from a thunderstorm. However, unlike the other two thunderstorm situations studied in stage 2 (London floods 030804 and Hawnby flood 190605), there was no obvious larger-scale dynamical mechanism involved. Water vapour imagery (not shown) indicated that there was a dry plume of air in the upper troposphere, which is indicative of warm and dry sub-tropical air aloft. There is no indication the sort of upper-level vortex associated with a lowering of the tropopause that was seen in the other cases.

At the surface the winds were very light from a mainly south-easterly direction and the pressure field fairly slack. Figure 6-1 shows that there was a shallow area of low pressure to the southeast of where the storms broke out but the pressure gradients were very weak. The air at low levels was however very warm and moist with wet-bulb potential temperatures (θ_w) near the surface exceeding 18 or 19^{o}C over a wide area (Figure 6-2). This is high enough to allow severe storms if the atmosphere is unstable to convection.

6.2 The differences between the 06 and 12 UTC forecasts

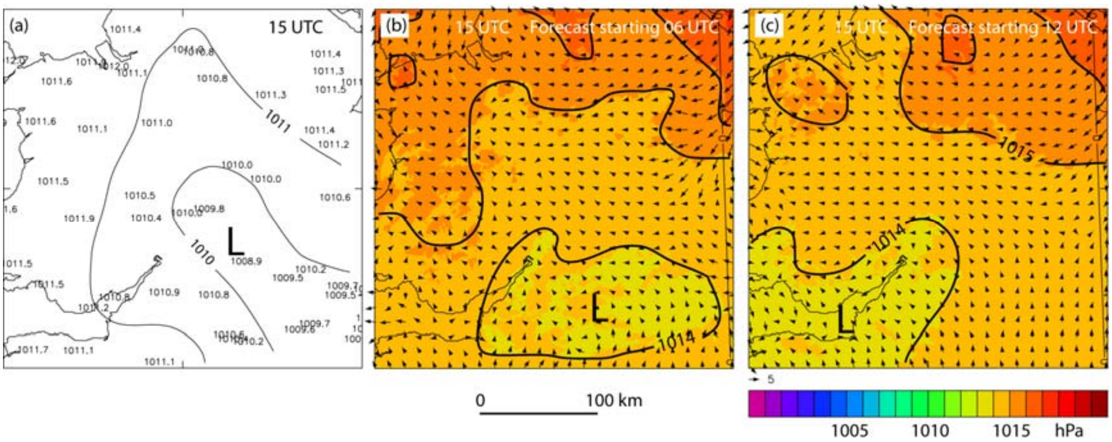


Figure 6-1. Pressure at sea level at 15 UTC 4th July 2006, (a) measured at synoptic stations and contoured by hand, (b) the 1 km forecast from 06 UTC, and (c) the 1 km forecast from 12 UTC.

Recall that the 1 km forecast from 06 UTC did produce heavy showers, some of which were intense, and produced high totals locally. The 1 km forecast from 12 UTC, even though 6 hours closer to the time of the storms, was unable to produce any significant showers in the area of interest. Reasons for the differences in the forecasts as well as insight into why the storms developed can be obtained by comparing the key meteorological fields.

6.2.1 Comparison of surface pressure

Figure 6-1 shows that the surface pressure fields in both forecasts were too high by around 3 hPa. This is unlikely to have had an impact on the development of convection, but is of some concern as typically the model should be not be biased like that. The forecast from 06 UTC had a somewhat more accurate pressure pattern, but given that the gradients were so slack, the significance of that difference is probably not great. Of more concern are differences in the temperature and humidity patterns near the ground. The best way to examine temperature and humidity together is to look at θ_w.

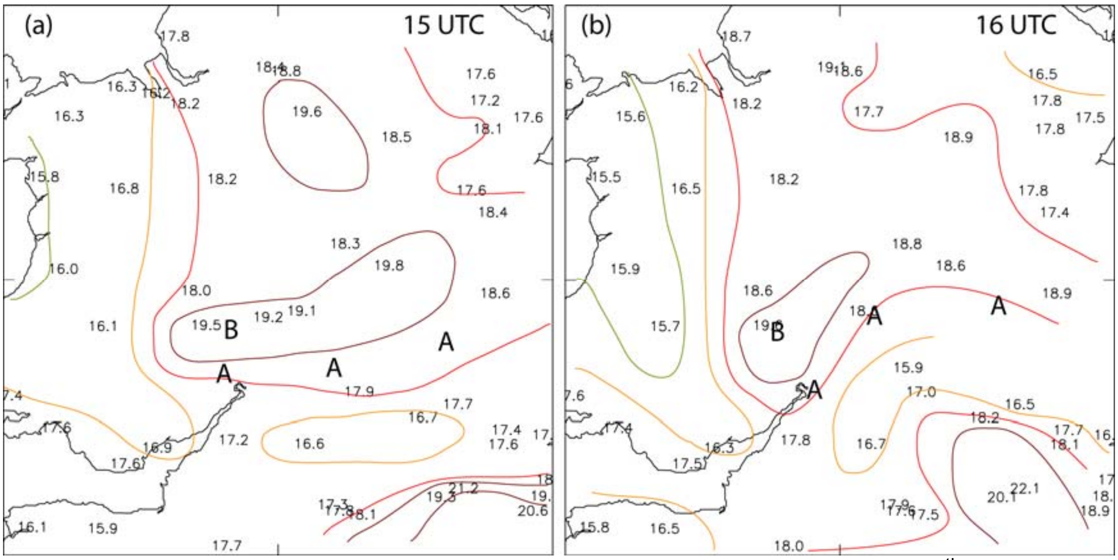


Figure 6-2. Wet-bulb potential temperature (θ_w) at 15 UTC and 16 UTC 4th July 2006 from synoptic stations and contoured by hand. The letters are referred to later in the text.

6.2.2 Comparison of low-level wet-bulb potential temperatures (θ_w)

From a comparison of Figure 6-2 and Figure 6-3 we see that θ_w values near the ground in the forecast from 06 UTC exceeded 18 or 19^{o}C over a wide area, just as observed. There are differences however. The model was too warm and moist towards the southeast of the domain and too dry/cool towards the north. The east-west gradient in θ_w seen in the observations over Wales (towards the left of the domain) is missing in the 06 UTC forecast. We will see that these subtleties in the temperature and humidity probably played an important role in determining where storms developed.

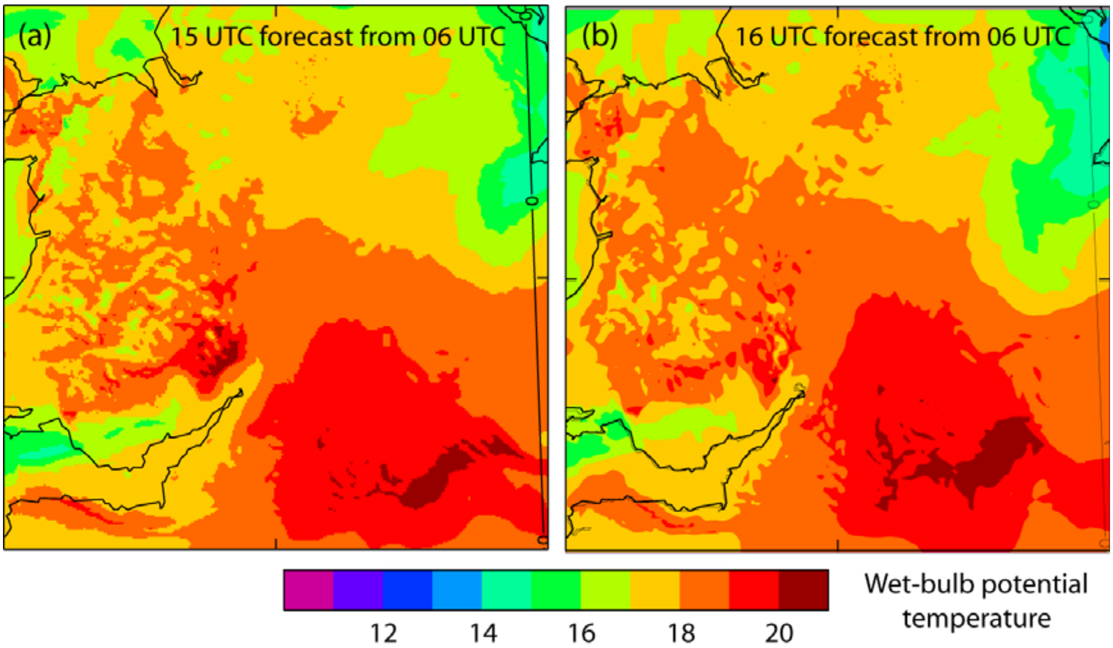


Figure 6-3. θ_w at 975hPa for 15 and 16 UTC 4th July 2006 from the 1 km forecast starting at 06 UTC.

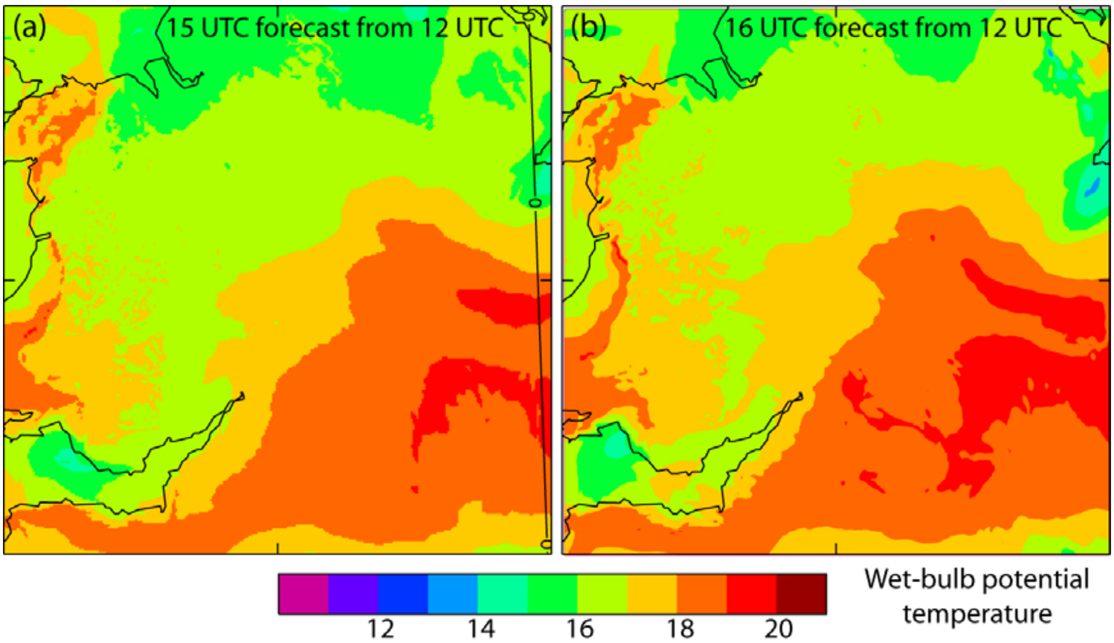


Figure 6-4. θ_w at 975hPa for 15 and 16 UTC 4th July 2006 from the 1 km forecast starting at 12 UTC.

The forecast from 12 UTC had lower values of θ_w near the ground than those observed and those produced by the 06 UTC forecast (Figure 6-4 compared with 6-2 and 6-3). The area above 18°C was very much smaller and largely confined to the southeast of the domain. Overall, the forecast from 12 UTC had θ_w values lower by 1-2 degrees over most of the domain. This is one reason why the 12 UTC forecast was unable to produce many showers, but there are also others.

6.2.3 Potential instability

For convection to develop the atmosphere must be potentially unstable, which means that θ_w must decrease with height. Potential instability does not necessarily mean convection will happen, but it can't happen without it, and the greater the potential instability the more intense a shower or storm is likely to be. Figure 6-5 (a & b) shows that, not only was the low-level θ_w higher in the 06 UTC forecast, but also the mid-level (600 hPa) θ_w was lower. This means that potential instability was greater over most of the domain in the forecast from 06 UTC, which is confirmed by Figure 6-5(c). In some places the potential instability was larger by as much as 2 or 3 degrees. By 16 UTC, several showers had already developed in the 06 UTC forecast and most had developed in areas where the differences in potential instability, (due mainly to differences in low-level θ_w) were greatest between the two forecasts.

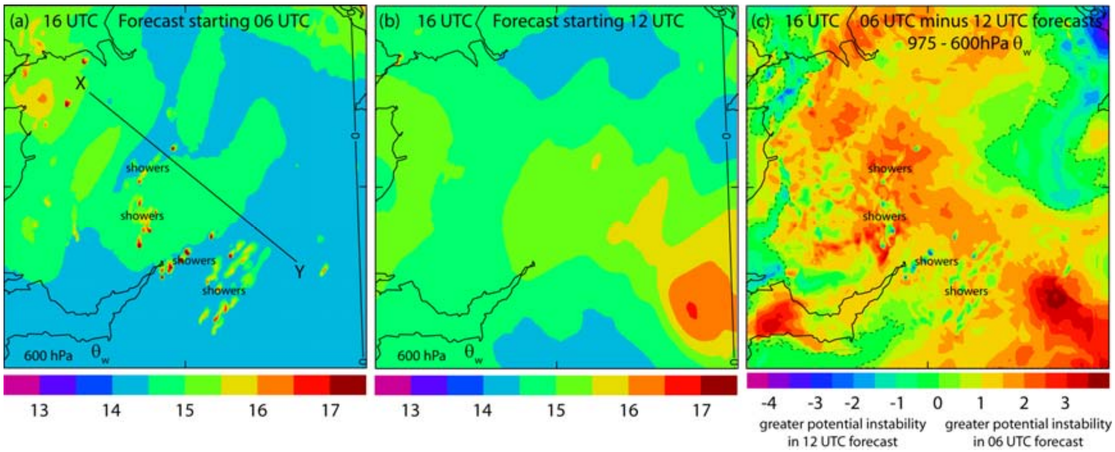


Figure 6-5.(a) and (b) θ_w at 600hPa for 16 UTC 4th July 2006 from the 1 km forecasts starting at 06 and 12UTC. (c) 975 hPa θ_w minus 600 hPa θ_w in the forecast from 12 UTC subtracted from 975 hPa θ_w minus 600 hPa θ_w in the forecast from 06 UTC. This shows the difference in potential instability (between those two levels) in the two forecasts.

We have seen that the forecast from 06 UTC had the most potential instability, but the forecast from 12 UTC was also potentially unstable over a wide area. Convection could have developed in either if the conditions were right. We haven't yet quite got to the bottom of why one produced showers and the other didn't.

6.2.4 The 'lid'

A cross section has been taken from locations x to y in Figure 6-5. Figure 6-6 displays the θ_w along this section from the 06 UTC forecast. The section clearly shows the potential instability with θ_w decreasing from around 18.5 to 19°C at the bottom (1000 hPa) to less than 15°C at around 600 hPa. In some places the decrease is as much as 5°C. The biggest vertical gradient in θ_w occurred at around 800 hPa. This is where the transition between the warmer and moister air below and the colder and drier air above was located.

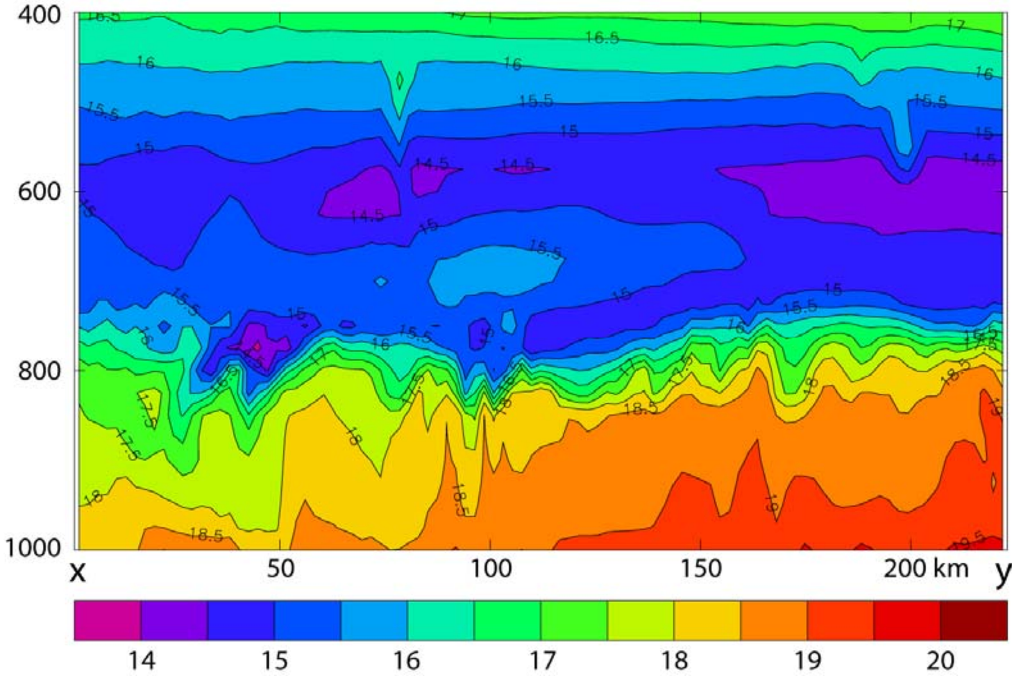


Figure 6-6. Cross section of θ_w along the line x-y in Figure 6-5(a) for 16 UTC 4th July 2006 from the 1 km forecast starting from 06 UTC.

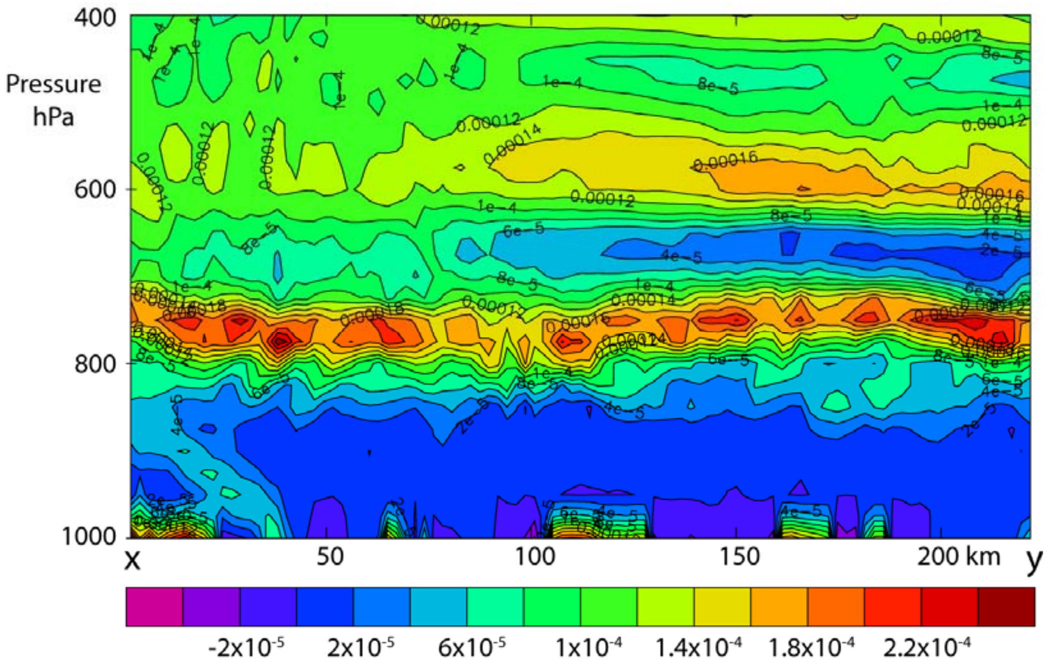


Figure 6-7a. Cross section of static stability (N^2) along the line x-y in Figure 6-5(a) for 16 UTC 4th July 2006 from the 1 km forecast starting from 06 UTC.

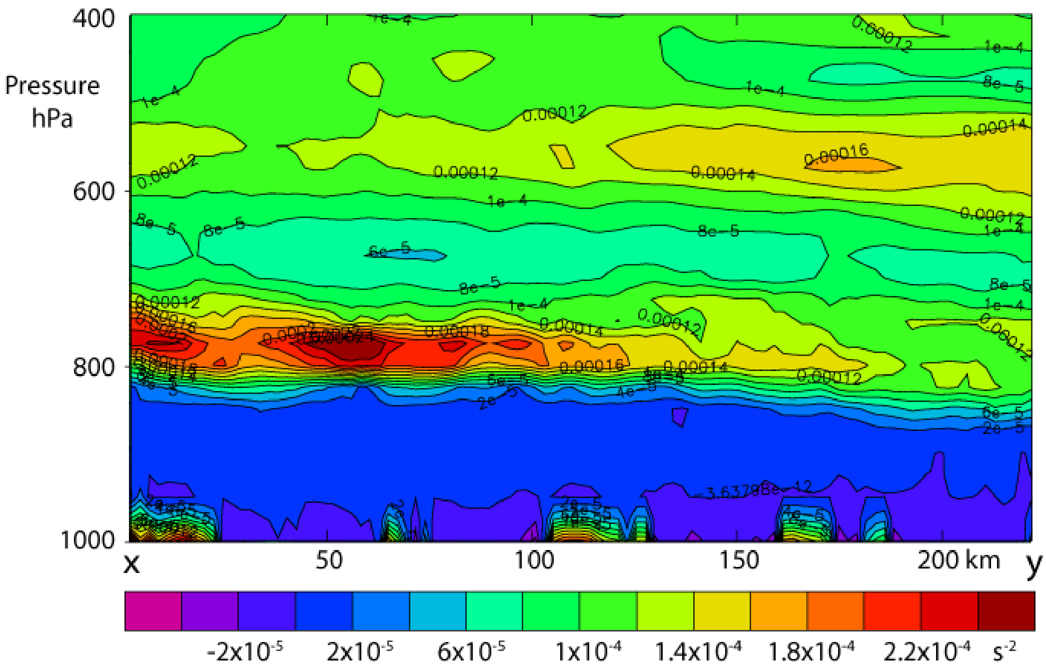


Figure 6-7b. Cross section of static stability (N^2) along the line x-y in Figure 6-5(a) for 16 UTC 4th July 2006 from the 1 km forecast starting from 12 UTC.

At a level just above the strongest θ_w gradient (~800 hPa) is where there was a so-called 'lid' (red/orange strip in Figure 6-7a&b). The lid is a layer where the temperature does not decrease so quickly with height (and may sometimes increase). The effect of a lid can be to trap the warmer and moister air below and prevent convection from occurring even though the atmosphere is potentially unstable. The stronger the lid the more likely it is that convection will be prevented. The two ways in which a lid can be made less of a barrier to convection are (1) lifting the air – this is what an upper-level vortex like those seen in the previous cases can do. (2) heating or moistening the air near the ground – the air near the ground would have been heated in this instance because it was a summer afternoon.

The diagnostic displayed in Figure 6-7 is the static stability (N^2). This shows that the position of the lower layer of more stable air (the 'lid') was mostly between 700 and 800 hPa in both the 06 UTC and 12 UTC forecasts. In the south-eastern part of the section (right side) the lid was lower (below 800 hPa), in the 12 UTC forecast but the values of N^2 were less in that forecast. Further to the northwest the height of the lid was similar in both forecasts but it was stronger (higher values of N^2) in the 12 UTC forecast. A comparison of N^2 between the 06 UTC and 12 UTC forecasts (Figure 6-8) shows the differences between the two. It highlights that the forecasts did differ in this important aspect and this would have had an impact on the likelihood of showers developing. The crucial aspect in the south-eastern part of the section is the lower elevation of the lid in the 12 UTC forecast, which would probably have acted to inhibit convection more (despite the lid being less sharp/strong), especially given the lower values of θ_w below the lid in that forecast. The stronger lid in the 12 UTC forecast in the north-western part of the section would also probably have acted to inhibit convection more.

Figure 6-7a&b also show that there was a more stable layer / lid at around 600 hPa in each of the forecasts. This was not significant as far as the development of convective clouds was concerned because any clouds that were buoyant enough to get past the first lid would then be much warmer than their environment at reaching the level of the second lid and continue to ascend.

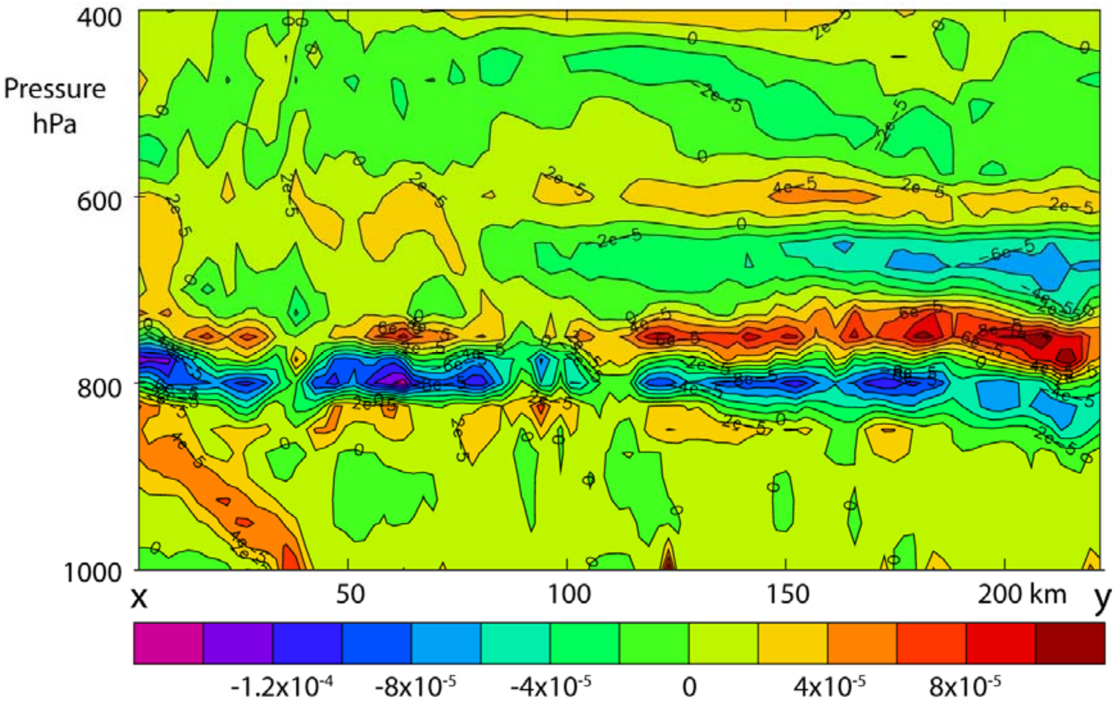


Figure 6-8. Cross section along the line x-y in Figure 6-5(a) for 16 UTC 4th July 2006 of static stability (N^2) from the 1 km forecast starting from 06 UTC minus N^2 from the 1 km forecast starting from 12 UTC.

Another way to visualise the lid is to look at a profile of the atmosphere by the using a tephigram (Figure 6-9). If a tephigram is not familiar, the important thing to see here is that the bulge to the right between 700 and 800 hPa in the temperature profile (red line) from the 12 UTC forecast (Figure 6-9(b)) is greater than in the 06 UTC forecast (Figure 6-9(a)). It shows that the lid was much stronger and sharper in the 12 UTC forecast. The other significant difference between the two is the humidity at low levels (below 800 hPa). Figure 6-9(b) shows that the air was drier at low levels in the 12 UTC forecast (dashed red line further to the left than the dashed blue line) even though the temperature profiles were almost identical. When the air is drier at low levels convection is less likely to break through the lid. The drier air at low-levels in the tephigram from the 12 UTC forecast is the reason for the lower values of low-level θ_w we've already seen at 975 hPa (Figure 6-4). The effect of these differences on the development of convection is shown by the green arrows in the tephigrams. The situation was marginal. During the afternoon the sunshine warmed the atmospheric boundary layer (below the lid) and by 15-16 UTC in the 06 UTC forecast there was enough low-level warmth and humidity for some clouds to get past the lid and then form deep thunderstorms up to 250 hPa (green arrow in Figure 6-9(a)). If the air had been cooler or drier at low levels or if the lid had been stronger or lower there would have been no convection. We see in the 12 UTC forecast that the convection could not break through (dashed green arrow in Figure 6-9(b)).

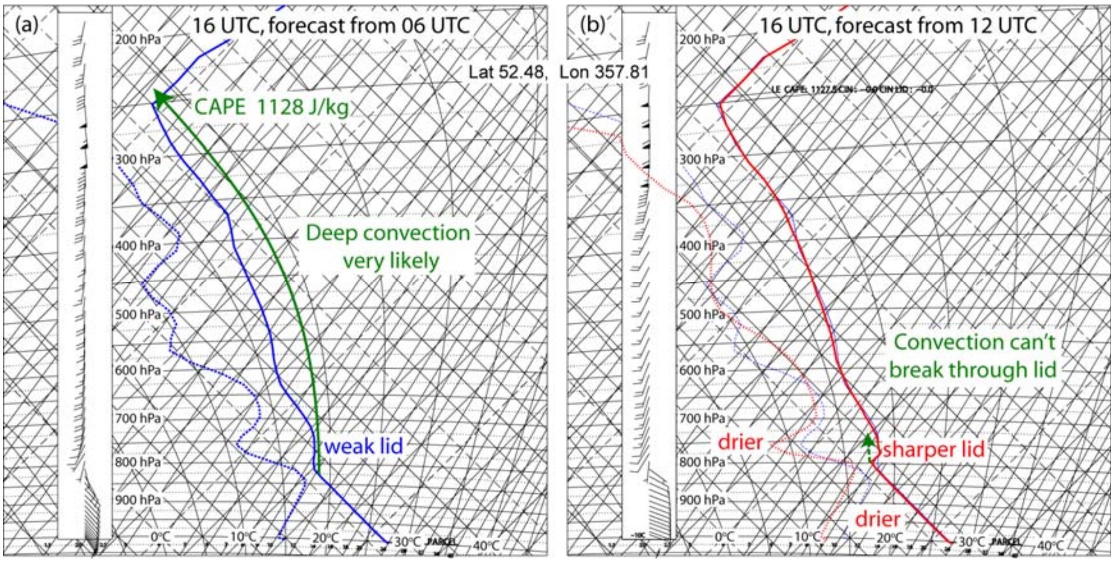


Figure 6-9. Tephigrams for 16 UTC 4th July 2006 at a point mid way along the section x-y (Figures 6-5 to 6-8) from (a) the 1 km forecast starting from 06 UTC, and (b) the 1 km forecast starting from 12 UTC. More explanation is given in the text.

These are profiles from just one location; it is reasonably representative of the area in the northwest part of the section shown in Figures 6-7 and 6-8 (left side). Profiles from other parts of the domain do not look the same, but in general over the domain, profiles from the 12 UTC forecast look less likely to produce convection than those from the 06 UTC forecast.

6.2.5 Summary of differences between the 06 and 12 UTC forecasts

The forecast from 06 UTC produced many more showers because

1. The air at low levels was generally warmer and moister (higher θ_w)
2. The potential instability was greater
3. The 'lid' was less effective

6.3 Why the showers developed where they did

The differences between the 06 and 12 UTC forecasts has provided some reasons for why the showers developed where they did in both the forecasts and in reality.

There did not appear to be any very obvious larger-scale forcing in this case, such as that associated with an upper-level vortex as in the previous cases. Triggering was more to do with the local atmospheric profile, although it is worth saying that a trough line was marked on the synoptic chart (Figure 5-1) by forecasters. This was probably marked simply because the storms seem to have developed along something of a line but it does mean that some upper-level forcing can not be completely ruled out. It is more likely that there was gradual ascent as the air flowed north and this helped to weaken the lid sufficiently along a line.

What seems to be of greatest importance in this case is the low-level variability in θ_w and the strength of the lid. (Typically, if the upper-level forcing is strong, the storms will persist well into the night in association with the dynamical lifting. These storms died out in the evening). As the afternoon progressed the lowest levels of the atmosphere heated up. The showers in the 06 UTC forecast generally developed where the low-level θ_w was highest and this occurred because of warmer and moister air to begin with and then solar heating in addition to that. Towards the southeast of the domain, the low-level θ_w was high in the forecast, but showers did not develop. The reason was because the lid was stronger there and even though the θ_w was high, convective clouds could either not form, or if they did, they could not penetrate the lid.

Observations show that in reality it also looks like the storms generally developed and persisted where the low-level θ_w was high. The area of initial triggering at around 16 to 17 UTC occurred where there was a north to south gradient of θ_w from ~16 to 19^{O}C (labelled 'A' in Figure 6-2). It is possible that some weakening of the lid through ascent occurred along that gradient and then the showers developed and broke through the lid where the θ_w was higher. It may be the reason why the showers broke out in something of a line. The most intense and persistent showers developed where the low-level θ_w was highest (labelled 'B' in Figure 6-2). As in the 06 UTC forecast, showers did not develop in reality where the low-level θ_w was high in the far southeast of the 1 km domain. We can speculate that this is also because of a stronger lid as it was in the forecast.

It appears to have been critical to get the correct pattern of low-level θ_w and the correct strength of the lid for a good forecast. Why the low-level pattern was like it was is unknown, but having examined observations from earlier times, the outflows from previous storms may have played a significant role in determining that pattern. It means that this was not an easy situation to predict. The margins were small between either developing or not developing showers, and the temperature and humidity pattern that determined the most favourable locations may have depended on getting earlier showers correct. There might also have been important local topographical effects but given the uncertainty in the temperature and humidity patterns, they would be difficult to detect.

The reason why the important meteorological fields were more accurate in the 06 UTC forecast than the later 12 UTC forecast is not known. The data assimilation system would have tried to improve the representation at the start of the 12 UTC forecast, but perhaps in trying to put right errors from the previous forecast with an insufficient number of observations made matters worse in this instance (e.g. drying the lower levels, changing the lid). It highlights the difficulty in predicting this type of situation. Note that there was only data assimilation in the 12 km forecasts (not 4 or 1 km) from 06 and 12 UTC so fine structure could not have been changed by the data assimilation system.

Summary of reasons for triggering in this case

To predict showers in the right place required that

1. The spatial variations in temperature and humidity at low levels were sufficiently correct.
2. The spatial variation in the strength of the lid was sufficiently correct.
3. Previous storms had been correctly forecast.
4. The solar heating was correct, which means that patches of cloud were correct.

All this assumes that the larger-scale background state was reasonably accurate. The dynamics of the storms themselves has not yet been mentioned.

6.4 Why was there such a lot of rain

The reasons why the storms developed in the first place has been discussed, but that is not the whole story. In some places extremely high rainfall totals were observed and these totals were large enough to produce flash flooding, yet there didn't appear to be a strong larger-scale dynamical mechanism for maintaining the storms once they had formed. The storms must have been self-maintaining in some way and able to produce rain over the same geographical location for up to an hour or more. The reason the rain fell in the same place might simply have been because the wind was light and the showers moved slowly, but the movement would have had to have been very slow indeed and individual storm cells typically have a lifetime of only 30 minutes or so. Perhaps there is another explanation.

6.4.1 Rainfall analysis

Figure 6-10 shows radar pictures every 15 minutes. The circle shows the Albrighton area where the highest rainfall totals were recorded. The 'x' marks the position of another shower as it progressed from frame to frame. The shower marked with the 'x' moved north by ~50 km and decayed in the hour and 15 minute period. In contrast, the rain within the circle remained very heavy for the hour from 18:30 to 19:30.

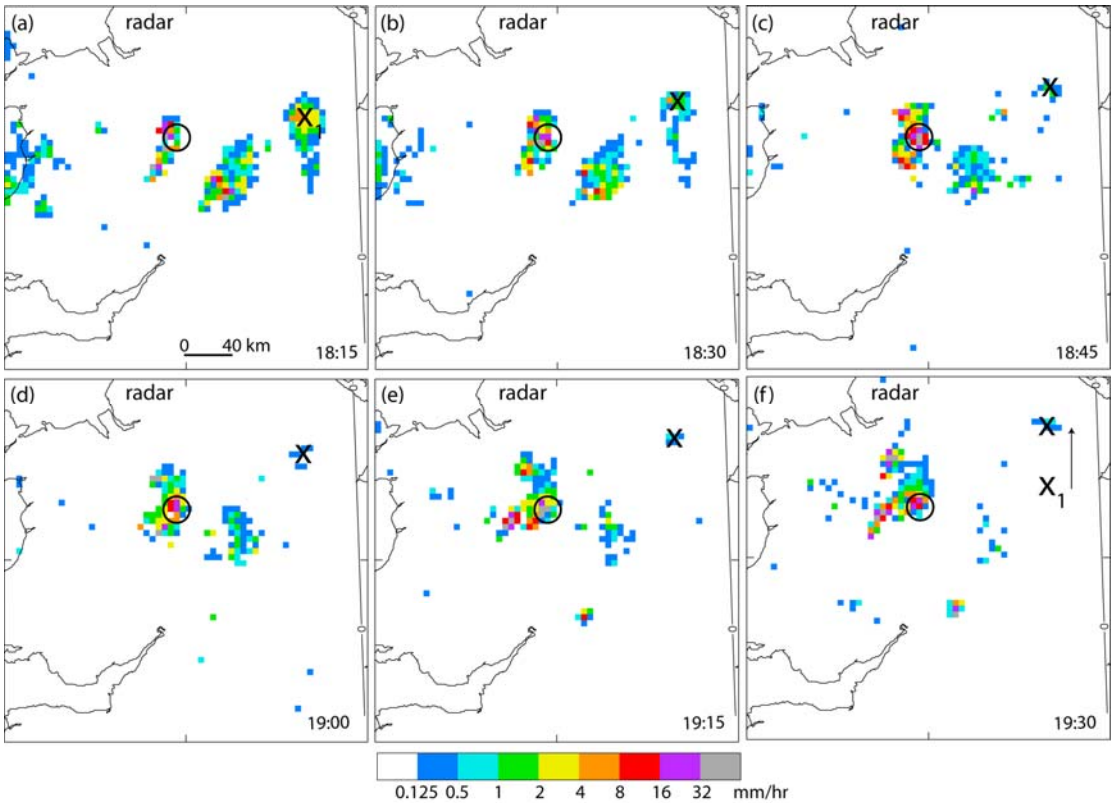


Figure 6-10. Sequence of snapshots of rainfall rates every 15 minutes from the radar network for 18:15 to 19:30 UTC 4th July 2006. The meaning of the circle and 'x' is described in the text.

Figure 6-11 shows the evolution of the rain cells from 17:45 to 19:45. At each time the rainfall areas have been displaced to the right by the distance shown by the arrow. The reason for doing this is to separate out each time frame so that the rain areas do not overlap. Doing this reveals that in the northward direction there are two velocities. First, the blue arrows show rain areas that moved north; second, the red arrow shows an area of rain that didn't move north, and if the frame to frame displacement is taken away, the rain area remained almost stationary. This is rain, within the circle in Figure 6-10, that produced the flash flooding because it poured down in the same place for nearly two hours. The question then is whether the model showed similar behaviour.

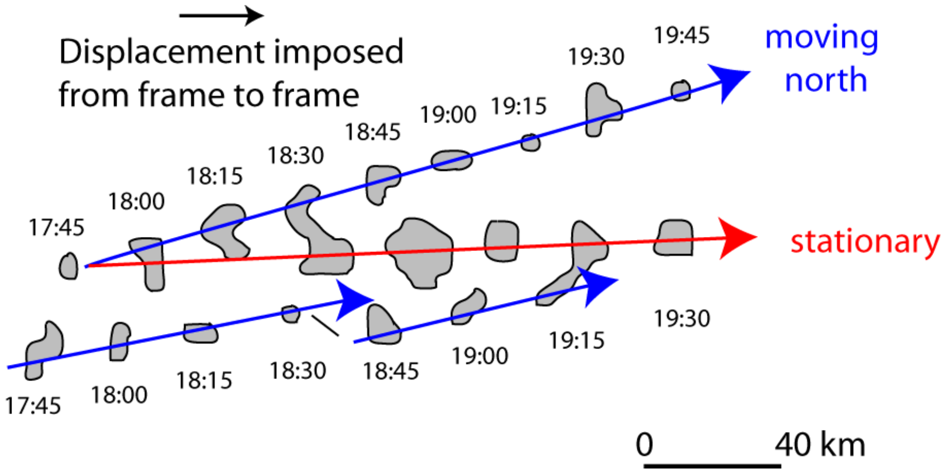


Figure 6-11. Shaded blobs show areas of rain exceeding 8 mm/hr taken from the network radar pictures every 15 minutes. Each time has been separated from the next by the length of the arrow labelled 'Displacement imposed from frame to frame'. Further discussion in the text.

Figure 6-12 shows a time sequence of rainfall rate snapshots from the 1 km forecast from 06 UTC over a small area within which a storm was simulated (close to Albrighton). This storm was smaller than one observed (Figure 6-10) and didn't persist for as long, but the rainfall rates were similar. Again, a circle has been drawn where the highest rainfall amounts were produced. As in the radar sequence, the heaviest rain at the southern end of the storm seems to have remained within the circle over several time frames. It looks like the model storm was behaving like the observed storm.

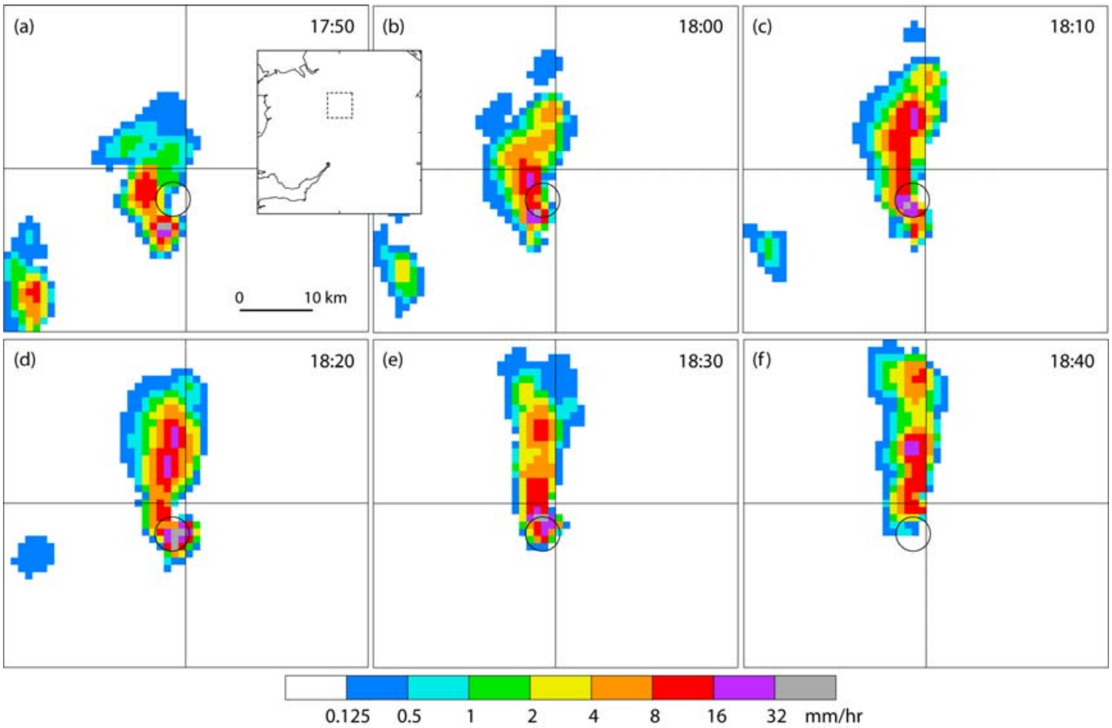


Figure 6-12. Sequence of snapshots of rainfall rates every 10 minutes from the 1 km forecast starting at 06 UTC for 17:50 to 18:40 UTC 4th July 2006. The meaning of the circle is described in the text. The smaller superimposed panel shows the area covered by the frames.

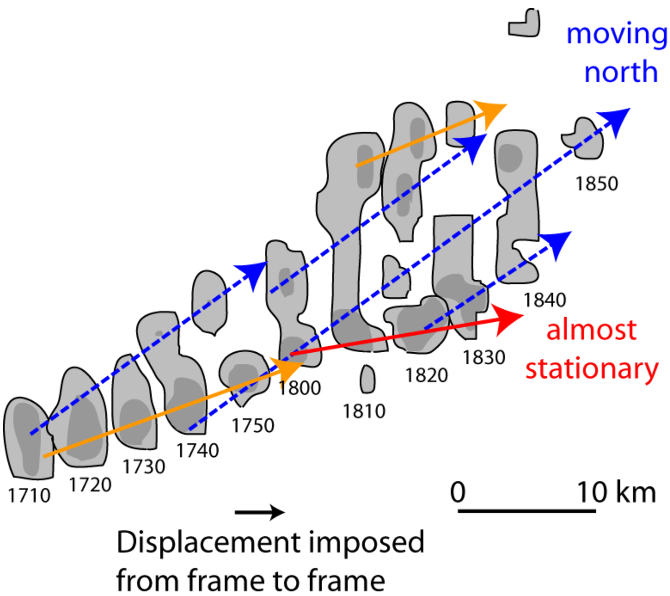


Figure 6-13. Shaded blobs show areas of rain exceeding 8 mm/hr every 10 minutes taken from the 1 km forecast starting at 06 UTC. Darker shades enclose rates >16mm/hr. Each time has been separated from the next by the length of the arrow labelled 'Displacement imposed from frame to frame'. Further discussion in the text.

The same kind of rainfall evolution can be seen in the model as in radar (Figure 6-13). Some of the cells were moving north (blue dashed arrows). Some of the cells were moving north but not as quickly (orange arrows). One of the rainfall cells was moving only very slowly north (red arrow). This cell was almost stationary (when the frame to frame displacement is removed) and produced up to 30 mm of rain in the ~30-minute period that it didn't move much. It would appear that whatever was happening that day could also be reproduced to some extent by the 1 km model.

6.4.2 New cell development (backward propagation)

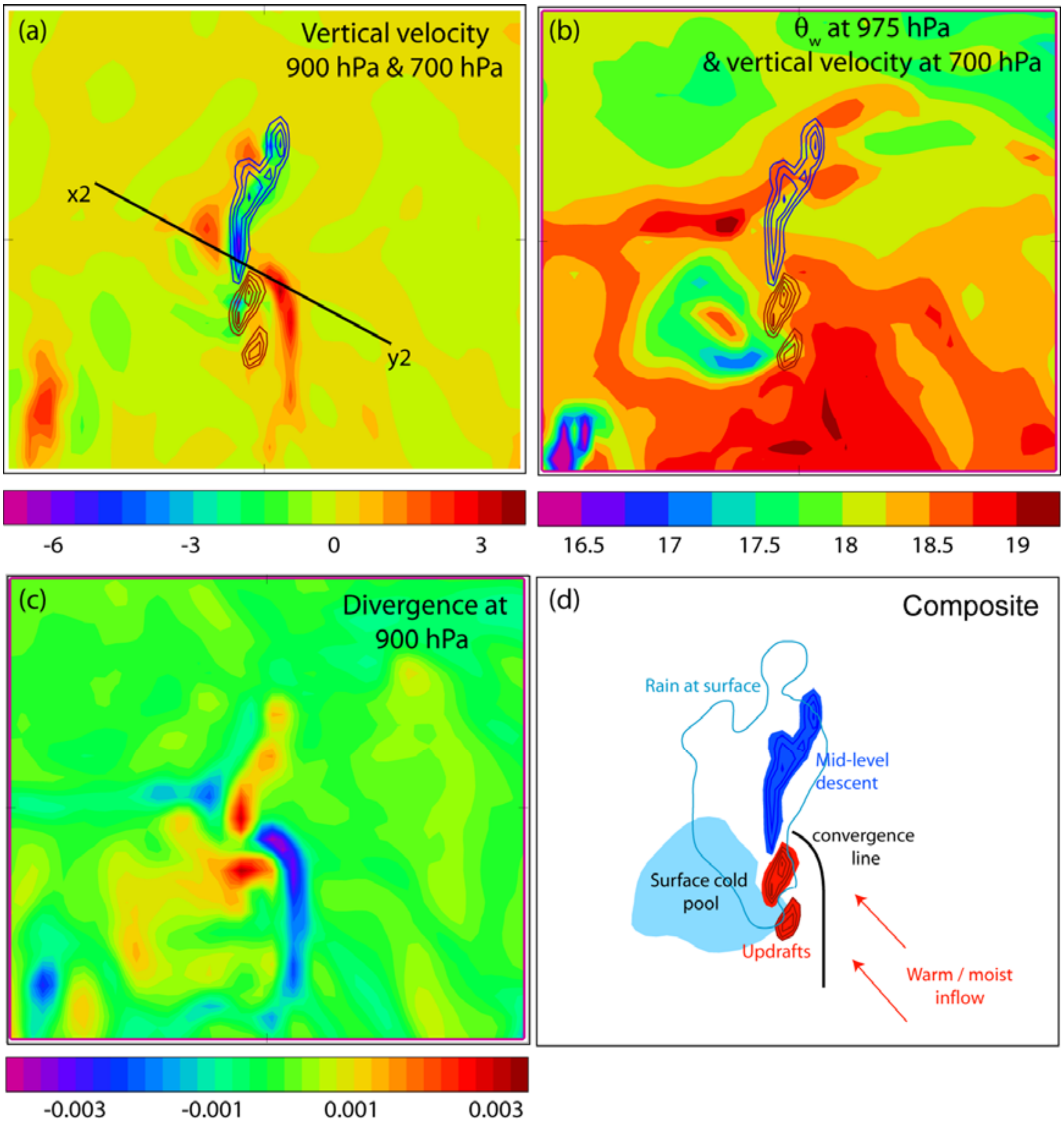


Figure 6-14. Diagnostics from the 1 km model forecast starting from 06 UTC for 18 UTC 4th July 2006, (a) Vertical velocity at 900 hPa (shaded) and 700 hPa (contoured, with blue=descent, red=ascent), (b) θ_w at 975hPa with vertical velocity at 700 hPa (as panel (a)), divergence at 900 hPa, (d) a composite picture.

An explanation of what was happening can be deduced from the panels in Figure 6-14. First though, we should think about the vertical wind profile. The tephigram (Figure 6-9) shows that the wind was from the southeast and light (~2-3m/s) below ~800 hPa and ~10m/s (~20kts) from the south above ~800 hPa. This means that a storm would tend to move north because the steering level (level at which the wind velocity is close to the storm velocity) would be above 800 hPa. It also means that rain would fall to the north of the main updraft into the area the storm is moving in to, and this would tend to kill off the storm. We should also think again about the storm environment. The updraft had to break through a strong lid and then, although there was plenty of Convectively Available Potential Energy (CAPE) to allow deep convection up to 250 hPa (as seen by the green arrow in Figure 6-9(a), the air above the lid was quite dry. Any mixing between the rising cloud plume and environmental air would dry the plume and weaken the storm. We should therefore expect that a single storm cell in this environment should move north and then decay fairly quickly.

Now back to Figure 6-14. We see in panel (a) that the downdraft at 700 hPa is located to the north of the updraft at 700 hPa. The downdraft was generated by evaporative cooling of the air in the rain, which then became denser than the air below and sank. In panel (b) we see that a cold pool in θ_w has formed at the ground. This is the downdraft at the surface spreading out. The downdraft-air moved south and then east relative to the movement of the storm, so the cold pool at the surface formed largely to the east of the updraft, but also started to undercut the updraft as the cold pool spread out. In panels (a) and (c) we see that there was both convergence and ascent at 900 hPa to the east of the updrafts. This was due to the edge of the cold pool meeting the warm moist flow from the southeast.

This strip of ascent at 900 hPa provided enough lift for the lid to be weakened and a new cell to develop that was fed by the high-θ_w air. This new cell, although forming to the southeast of the initial storm cell, produced rain in the same place as the original storm. A composite picture is shown in panel (d).

The composite cross-section along x2-y2 in (Figure 6-15(d)) shows the cloud from the original storm and the new cell forming as the downdraft below the tilted updraft spread out and undercut the warm moist inflow. It also shows the rain falling into another layer of cloud that was formed by ascent at the edge of the cold pool. This layer of cloud enhanced the rainfall by the seeder-feeder mechanism (just like in the orographic enhancement situation described in the stage 1 report).

It would seem that the 1 km model was capable of producing locally high rainfall totals (though not as high as observed) because it could generate a new storm cell at the back of the original cell, and the propagation speed was such that the new cell produced rain in the same place as the original cell.

A 12-km model would not be able to produce this backward propagation because the convection is not resolved. Neither would a nowcasting system that just advected existing rain.

In order for the backward propagation to have occurred in the 1 km model several processes needed to have been represented with sufficient accuracy. The initial updraft has to be of the right magnitude and then the cloud formation and the microphysical interactions that produce the rain should also be sufficiently correct so that a realistic downdraft can form. If the updraft is too strong the downdraft may be too strong and the positioning of the new cell will be changed and the positioning of the rain will also be changed. If the updraft is too weak the downdraft may not be sufficiently strong to produce the convergence necessary for a new cell to break through the lid. If the processes that produce the cloud and rain are wrong then the downdraft may, again, be too strong or too weak with the same consequences. If the background flow is different then a different updraft and downdraft will be needed to get rain falling in the same place. If the lid is too strong a new cell (or even the original cell) may not form. If the lid is too weak, widespread convection and secondary cells may develop.

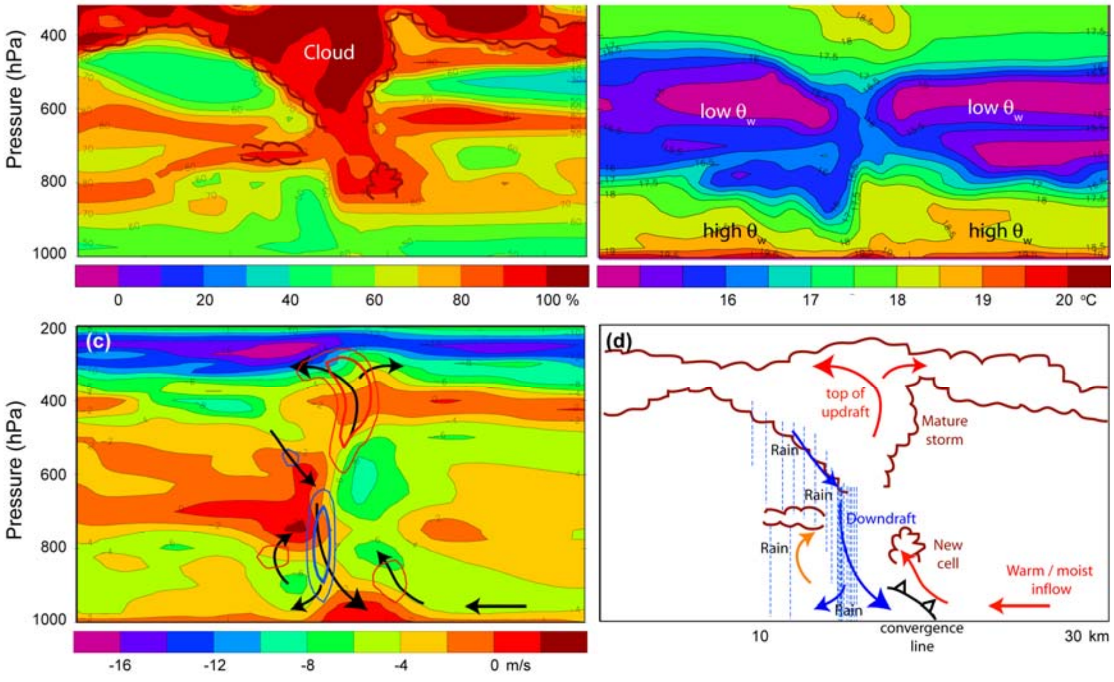


Figure 6-15. Cross section along the line x2-y2 in Figure 6-14. Diagnostics from the 1 km model forecast starting from 06 UTC for 18 UTC 4th July 2006, (a) Relative humidity, (b) θ_w, (c) along section wind (shaded), vertical velocity (contoured every 2 m/s) and arrows superimposed to depict key flows, (d) a composite picture.

Given all this uncertainty, it is remarkable that the model is able to produce something realistic at all. The fact that it does is probably an indication that there is some tolerance allowed in the representation of the physical and dynamical processes. It also indicates that the 1 km model is able to do a reasonably good job at representing the storm dynamics and physical processes. A 1 km model can therefore represent this type of backward propagating storm, but in this case the uncertainties in the local pre-storm environment and then the need for an accurate representation of the storm dynamics meant that the predicted rainfall totals were not as large as observed and at a different location. The uncertainties associated with this type of event present a big challenge for the future, but at least we have a starting point – a model that we know is capable of producing a realistic storm evolution. However, we do need further research to improve the representation of the pre-storm environment through improved data assimilation methods at fine scale and further research to improve the turbulent and microphysical processes within the model.

6.5 Summary

- The storm that produced the flooding in the village of Albrighton on the 4th July 2006 occurred primarily because cumulus clouds that formed as a result of solar heating of warm moist (high-θ_w) air near the ground, were able to break through a so called 'lid' and produce locally deep convective storms.

- The low-level pattern in θ_w and the strength of the 'lid' were crucial for determining where showers would form in this case.

- The rainfall totals were very large in some places because the storms were able to generate new storm cells behind, which went on to produce rain in the same place. The 1 km model was able to reproduce this behaviour. A 12 km model or advection nowcasting system could not.

- The storms died out in the evening as the sunshine became weaker. This is further evidence that there was no significant larger-scale dynamical forcing in this instance.

- The 1 km forecast from 06 UTC was able to generate many more showers than the 1 km forecast from 12 UTC because the air near the ground was generally warmer and the lid weaker. Why this was the case is not clear.

- The case has highlighted the need for further research into improving data assimilation at fine scales as well as further research into investigating and improving turbulent and microphysical processes within the model.

- This type of situation is inherently difficult to predict and a probabilistic approach to presenting output from a ~1 km NWP model is essential.

7. Detailed study 3: The Albrighton flood, 4th July 2006.

Part 3. 1 km sensitivity studies

This is the last of three sections about this case. The purpose of this section is to show the differences that scientifically defensible changes to the model grid or formulation make to the 1 km model forecasts.

The sensitivity experiments are:

(1) Change from 76 to 38 vertical levels.
(2) Change to a 1.5 km horizontal grid spacing with a 300x300 grid-square domain. This domain is therefore larger than the 300x300 1 km domain.
(3) Change to a 1.5 km horizontal grid spacing on the same sized domain as the 1 km model (300x300 km).
(4) Add more diffusion to the larger 1.5 km domain.
(5) The additional of Smagorinsky-Lilly type turbulent mixing to the 1 km forecast.
(6) The use of the convection parametrisation scheme to represent the effects of sub-grid-scale clouds in the larger 1.5 km domain.

All the sensitivity experiments are applied to the forecast from 06 UTC (rather than 12 UTC) because it was much better.

Justification for these experiments will be given as they are presented. Clearly, these experiments do not cover anything like all the possibilities. The idea is not to tune the model to this case, rather to examine the way the model behaves with these changes in this meteorological situation to see whether more general conclusions can be made.

7.1 Change from 76 to 38 vertical levels

It was shown in the section 6 that the triggering of the storms was partly dependent on the strength of the 'lid'. The purpose of this experiment is to see whether a poorer vertical resolution and therefore a less well resolved lid has a significant impact on the forecast. Vertical resolution may also have an impact on other aspects as well as the lid of course. Increased vertical resolution has been shown to give a better representation of slanted frontal circulations in a high-resolution model (Lean and Clark 2003 however for convective situations it is not so clear whether it matters so much. This a good test because the showers are locally triggered and the lid is thought to be such a critical factor.

Figure 7-1 shows how the rainfall accumulations over the period of interest (17 to 21 UTC) have been changed by the halving of the number of vertical levels. The overall amount of rain has decreased markedly and the rain in the Albrighton area has disapeared. Given that the amount of rain in the 1 km model was already too little, it appears that the reduction in the number of vertical levels has made the forecast worse. Some caution is needed here though; because the background meteorological fields are not exactly correct and the errors are not known (i.e we don't know what the true atmospheric profiles looked like), so we should not draw broad conclusions from a comparison against the observed rainfall in a single case. Nevertheless, the key result here is that the change to lower vertical resolution does make a big difference and probably for the worse.

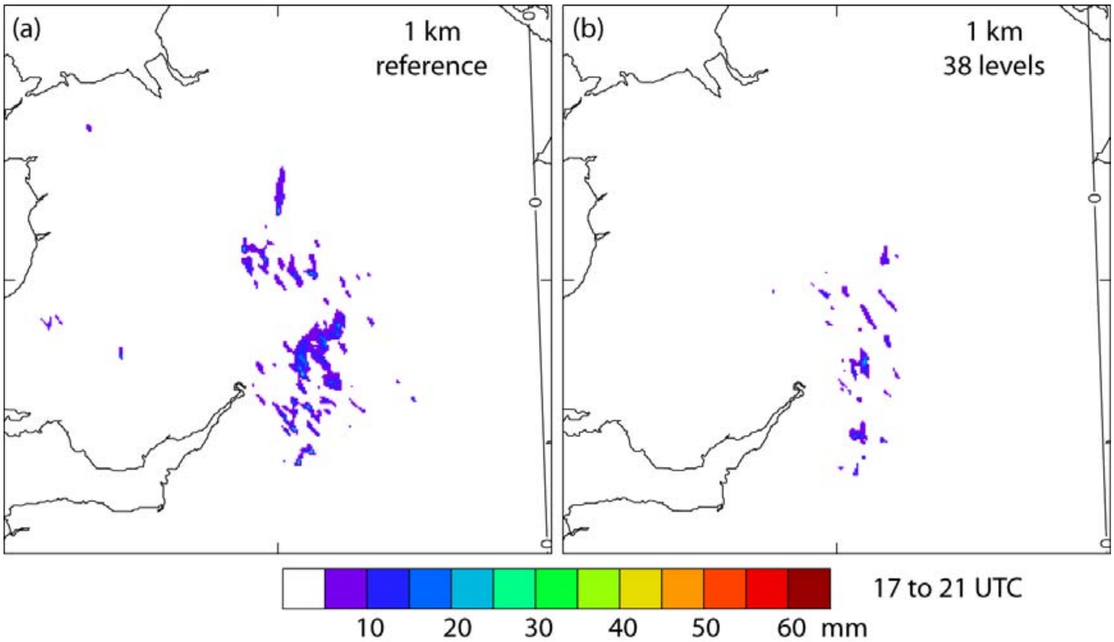


Figure 7-1. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) with 76 vertical levels, and (b) with 38 vertical levels.

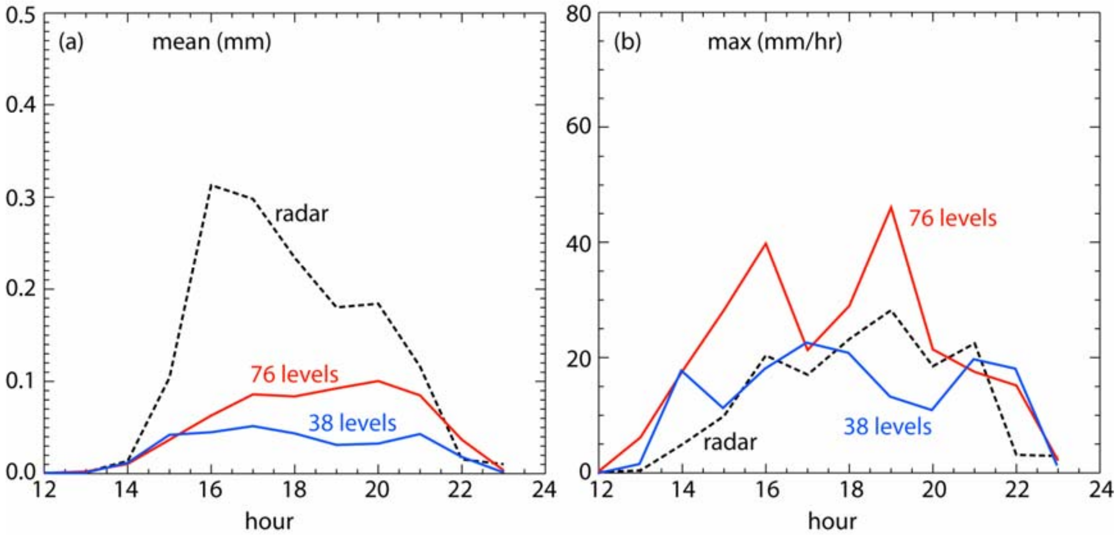


Figure 7-2. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar and the 38 and 76 level 1 km forecasts from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

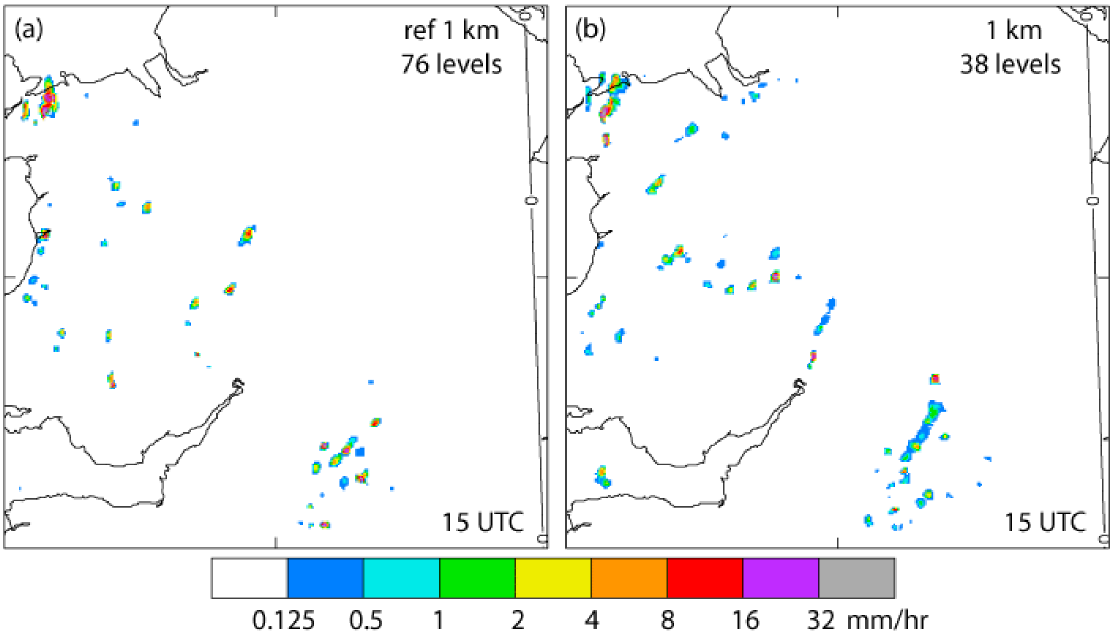


Figure 7-3. Rainfall rates at 15 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) with 76 vertical levels, and (b) with 38 vertical levels.

The reduction in rainfall in the 38-level forecast is also seen as a reduction in the hourly accumulations (Figure 7-2(a). Up until 15 UTC the forecasts gave the same accumulations, which suggests that the early triggering was similar. Figure 7-3 shows that the 38-level forecast did generate a similar number of showers early on (15 UTC), but after that the 38-level forecast produced less rain. This is suggestive that the initial breaking through the lid was not very different and casts doubt on the notion that representation of the lid was by far the most important aspect. The rainfall rates in Figure 7-3 indicate that the 38-level model developed less intense cells, and a closer inspection of the output confirmed that behaviour. Later, when the generation of new cells was important, the weaker cells in the 38-level forecast were less likely to trigger new cells, which were also probably weaker than in the 76-level forecast and this feedback resulted in less rain

It is worth having a look at how the representation of the lid can be changed by reducing the vertical resolution and this is demonstrated by the tephigrams in Figure 7-4. The stable layer (lid) between 900 and 850 hPa in (a) and 825 and 750 hPa in (b) has more detail in the 76-level forecast. In (a) the boundary layer (air below the lid) is deeper and slightly warmer with 76 levels (up to 900 hPa instead of 925 hPa). In (b) the lid is more stable and, again, the base of the lid is higher and moister with 76 levels. These differences may be significant in marginal situations when the difference between having convective storms or not is finely balanced. This situation appears to have been marginal in some places, but the differences between the 76 and 38 level forecast profiles shown here (and others not shown) are not great and either 76 or 38 level profile would have allowed convection with just a little more warming from sunshine. Perhaps the differences between the representation of the lid was not the major reason for the difference in rainfall in this case.

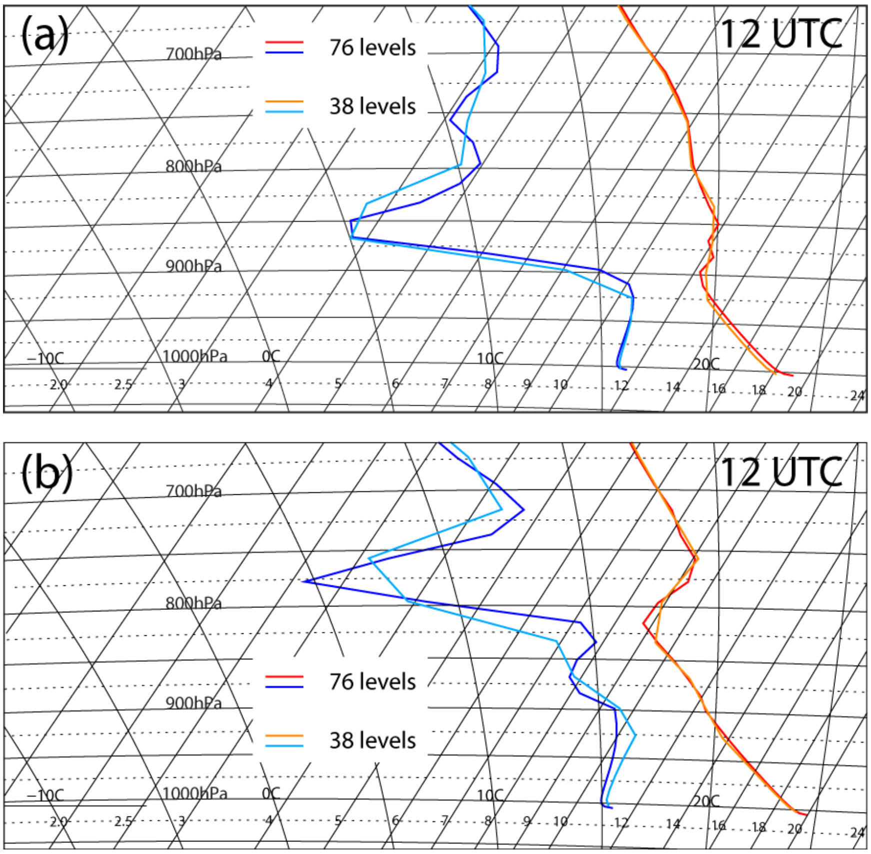


Figure 7-4. Tephigrams at locations along the cross-section x-y in section 6, from the 1 km model forecasts with 76 and 38 vertical levels.

7.2 Change from 1 km to 1.5 km with larger domain

It is probable that the first operational ‘storm-resolving’ model at the Met Office will have a horizontal grid spacing of 1.5 km. A 1.5 km model will allow a larger domain for the same computational cost as a 1 km model and it is therefore makes it more viable to have a domain that covers the whole UK. It is thought that the differences between 1.5 km and 1 km will usually be small. This event might be one of those in which differences are most evident because convection is locally driven, hence the reason for this experiment. The 1.5 km forecast examined here also started from 06 UTC and was on a 300x300 grid-point domain centred at the same location as the 1 km domain. This means that the domain was larger – 450x450 km.

Figure 7-5 shows that the 1.5 km model produced considerably more rain and significantly higher totals than its 1 km counterpart, but the general area where the rain fell was similar.

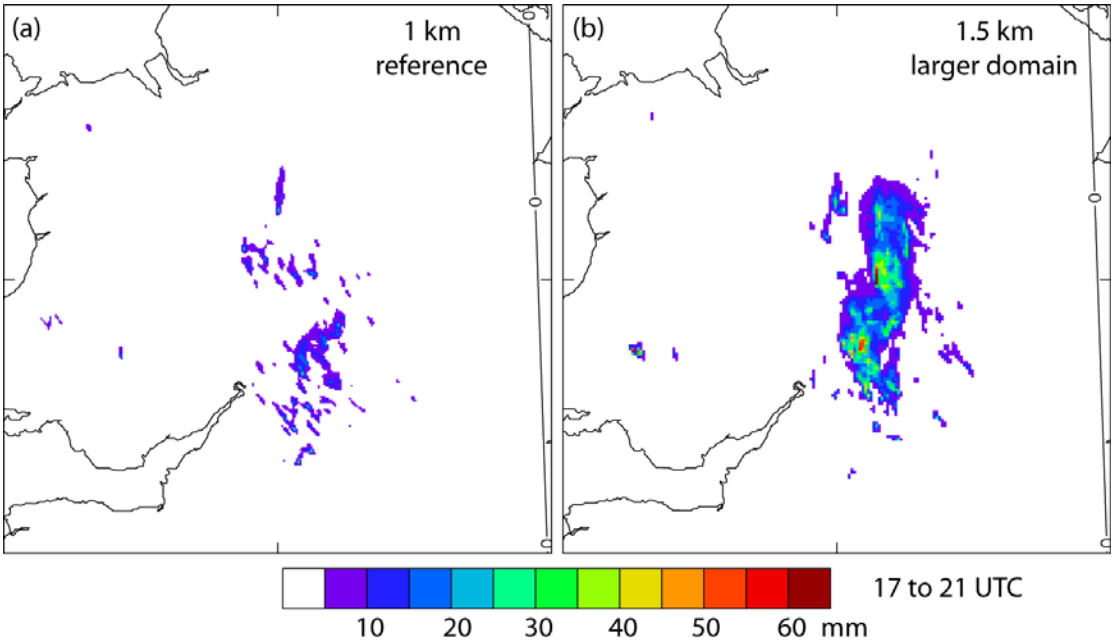


Figure 7-5. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain).

7.2.1 The differences between the two forecasts

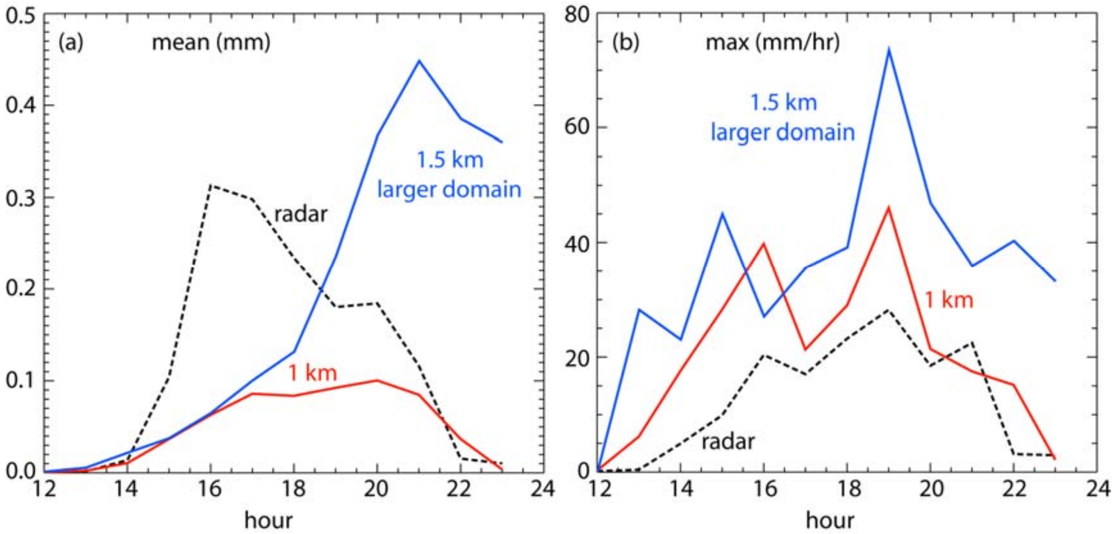


Figure 7-6. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar, the reference 1 km forecast (300x300km domain) and the 1.5 km forecast (450x450 km domain) from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

The differences in hourly mean rainfall amounts (over the 1 km domain) were very large (Figure 7-6). Up until around 17 UTC they were similar and then the values shot up in the 1.5 km forecast and continued high through the evening. Whereas the 1 km forecast produced too little rain the 1.5 km forecast produced too much. Both had their peak in rainfall amount too late and this was slightly worse at 1.5 km. The persistence of the rain into the evening in the 1.5 km forecast was overdone. The showers should have decayed and died out as the sun went down. Figure 7-6(b) shows that the rainfall rates in the 1.5 km model were higher than in the 1 km model especially early in the afternoon and late in the evening.

7.2.2 Why the differences

We saw in section 5 that the 4 km model was unable to produce any convection until much later than it occurred because it takes longer to resolve convective storms on a 4 km grid, especially in situations when the showers are localised and not particularly large, as was the case here. When the showers did develop in the 4 km model they were too few and too large (if they can develop at all before the suns heating has gone). The 1 km model on the other hand can have a tendency to initiate a rash of heavy showers too early because they are easier to resolve on a 1 km grid and because there may be insufficient turbulent mixing to weaken them initially. The 1.5 km model has a behaviour that lies between those two (though closer, of course, to the 1km model). So, in the early afternoon, fewer showers developed in the 1.5 km forecast than in the 1 km forecast, but they were heavier (Figure 7-7). That is why the rainfall rates were higher but the mean accumulations about the same up to 15/16 UTC in Figure 7-6.

From mid afternoon onwards the biggest differences emerged. Recall from section 6 that the showers in the 1 km model were able to reproduce some of the observed behaviour in that they could generate new daughter cells. Any individual storm cell was not able to persist for long due to a number of factors; the strength of the lid, the dryness of the air above the lid, the wind profile (which meant that the downdraft could cut off the updraft) and the lack of larger-scale forcing. All this made for a hostile environment in which the storms had to grow, despite there being enough Convectively Available Potential Energy (CAPE) for clouds to potentially become very deep. The persistence of a thunderstorm depended on new cells forming. A cell needed to be strong enough to produce a strong enough downdraft to provide enough low-level convergence for a new cell to break through the lid. The 1.5 km model produced fewer but stronger cells and because those cells were stronger they were much more likely to produce new daughter cells. By evening, when the sun was weaker, neither of the forecasts would have been able to produce new cells by any other means. This meant that the showers in the 1 km forecast gradually died out as cells weakened. However, in the 1.5 km forecast it appears that enough new daughter cells were able to form and then to go on to generate further cells. By this stage there might have been cells forming as outflows collided and the result was a large storm complex, or multicell storm (Figure 7-8). Once this larger entity had developed it was able to ‘suck in’ air from the surrounding environment and continue to produce new cells well into the evening as it tracked north. In effect, the multicell storm had created its own larger-scale forcing and the lack of solar heating no longer mattered.

The two very different behaviours show just how difficult a situation this was to predict and highlights the need to get the balance between model resolution and sub-grid-scale mixing and microphysics correct.

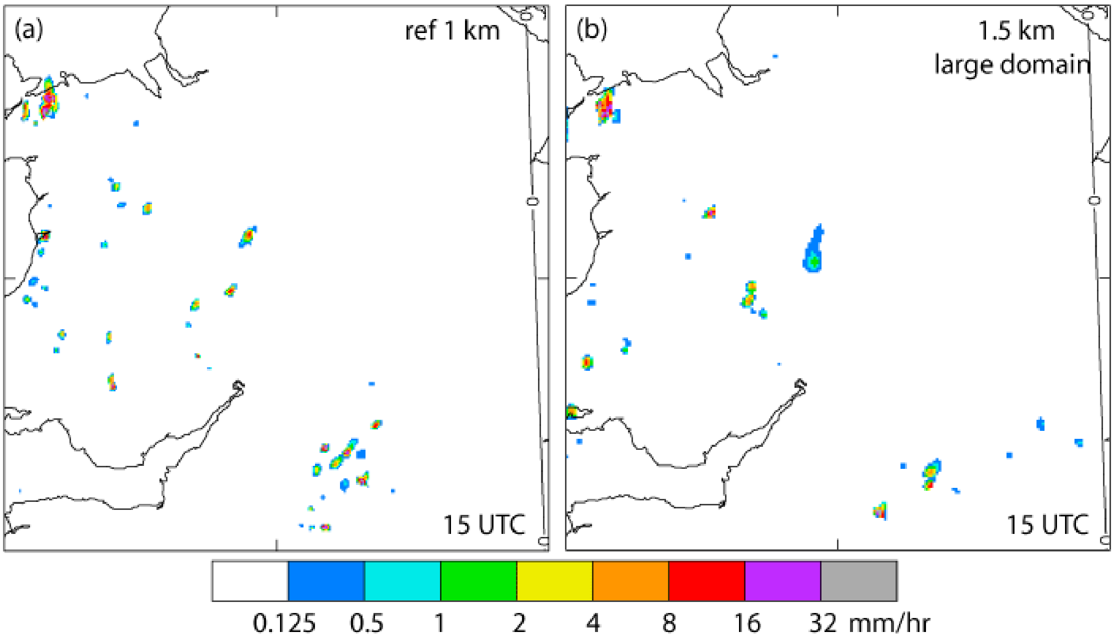


Figure 7-7. Rainfall rates at 15 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain).

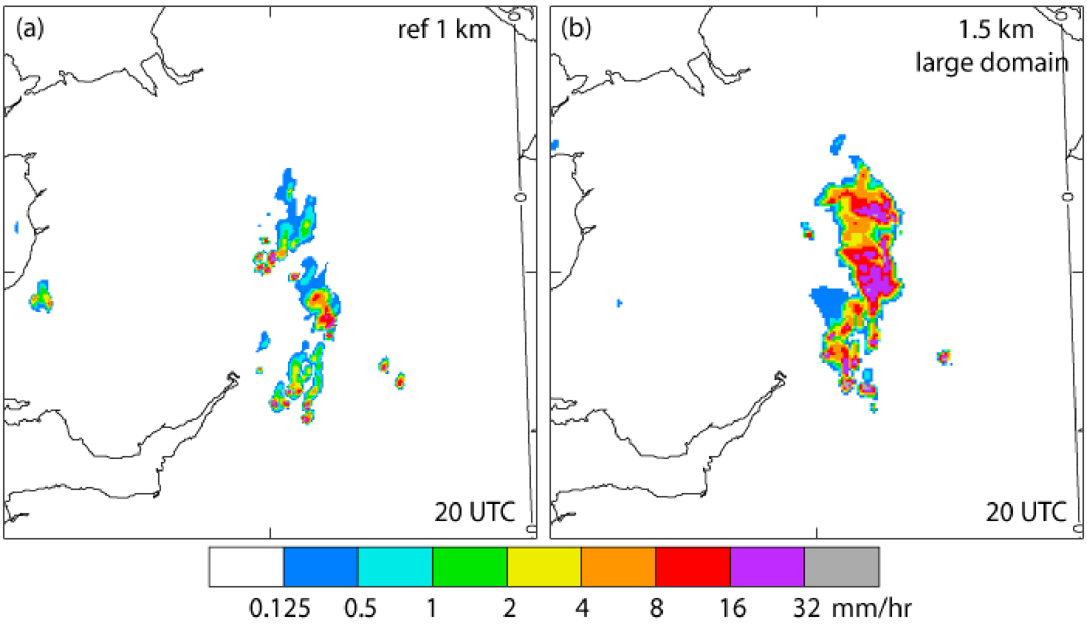


Figure 7-8. Rainfall rates at 20 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain).

7.3 Change from 1 km to 1.5 km with the same sized domain

.

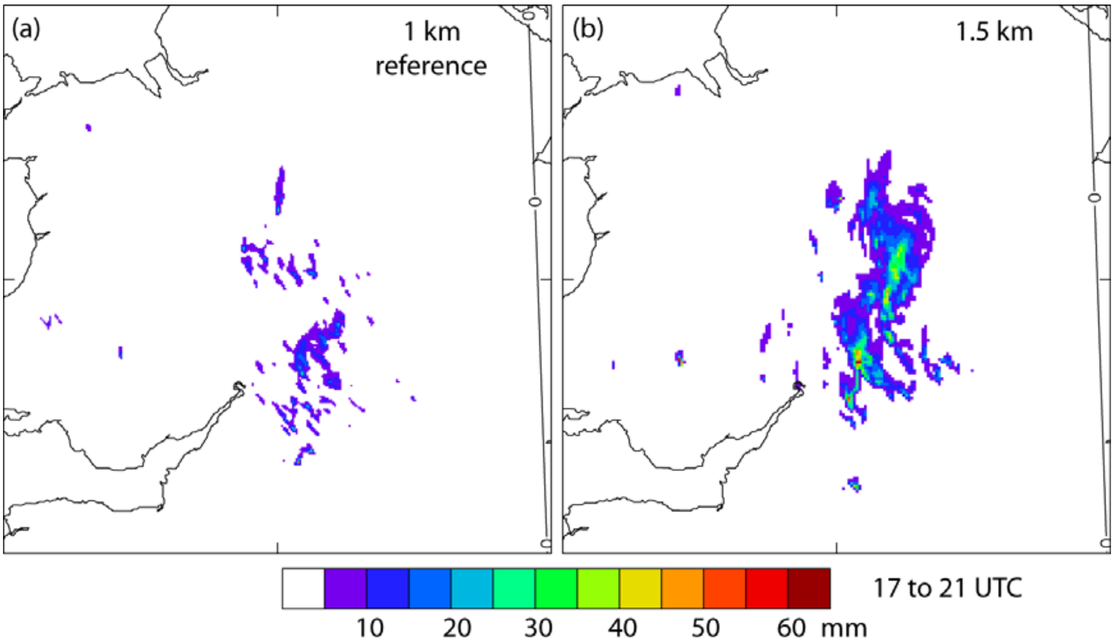


Figure 7-9. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (300x300 km domain).

It was shown in the Storm Scale Modelling project (Roberts 2005) that the boundaries of the domain can have a detrimental impact on the forecast in a high-resolution model. This was particularly true in convective situations with a strong flow coming into the domain. The reason being that the flow entering the domain from a coarser-resolution model needs time (and therefore distance) to generate showers. In general it is thought to be better to have a larger domain, not just for the reason above, but also because there can be other spurious effects at the boundary.

For this case it was considered worthwhile to examine the difference between the 1 km forecast and a 1.5 km forecast with the same number of grid points (larger domain) and a 1.5 km forecast with the same domain. This situation does not have a strong flow through the domain, but as we've seen there is uncertainty as to how much convection will develop, so it is of interest to see how the forecasts compare as small differences may have large feedbacks.

The 4-hour accumulations produced by the 1.5 km forecast on the 1 km domain were much greater than those produced by the 1km forecast (Figure 7-9) and similar to those produced by the 1.5 km forecast on the larger domain (see Figure 7-5(b)). There were differences between the two 1.5 km forecasts in detail, but overall rainfall patterns were very alike. Figure 7-10 shows that the 1.5 km forecast on the 1 km domain gave similar mean hourly rainfall amounts as the 1 km forecast up to 16 UTC and then like the 1.5 km forecast on the larger domain the rainfall totals increased much more rapidly after that.

There are some differences between the hourly mean rainfall amounts in the two 1.5 km forecasts (Figure 7-6 and 7-10). The rainfall on the smaller domain increased more rapidly between 16 and 17 UTC and then the peak at 21 UTC was lower. Why this was the case is not known. Some differences would be expected in this situation just because of the nature of the convection that has been discussed earlier. It is possible that the bigger increase in rainfall was associated with more gravity wave activity on the smaller domain, but that is speculation.

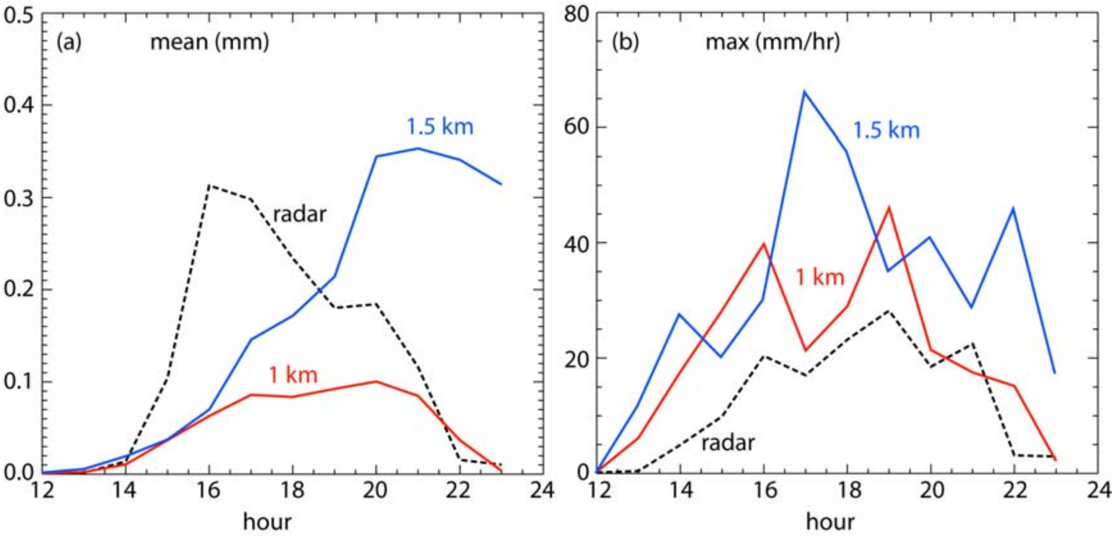


Figure 7-10. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar, the reference 1 km forecast (300x300km domain) and the 1.5 km forecast (300x300 km domain) from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

7.4 Representing sub-grid-scale turbulence

On some occasions the showers in the 1 km model can tend to be too small, too numerous and develop too early; the 'measles' effect. Such behaviour was also observed on this occasion. It has already been recognised that one of the probable causes of this behaviour is a lack of mixing of air at the edges of the resolved clouds. Typically, if a model grid square is representing a shower, adjacent grid squares will be representing the environment outside the shower. In reality, the cloudy air and the drier air will mix at the turbulent boundary, which means that the transition should not be so sharp and there should be a gradient across grid squares. There needs to be a way of representing this turbulent mixing that the model is not able to resolve.

An adaptation of the Smagorinsky-Lilly formulation for representing sub-grid-scale turbulence (Smagorinsky 1963, Lilly 1967) has recently been developed for implementation in the Met Office Unified Model (Halliwell 2006). The effect is to apply additional diffusion locally; particularly where there is stronger wind shear (as is the case in the vicinity of updrafts in showers).

In this experiment a comparison was made between the 1 km forecast from 06 UTC and an equivalent 1 km forecast, but with the Smagorinsky-Lilly type 3D turbulence included. The purpose of doing this was not to test the new scheme exactly the way it is going to be implemented in a future system as that is not yet decided; work is ongoing and there have been new developments since this experiment was performed. The objective was to see just how sensitive this type of event could be to the inclusion of the sub-grid turbulent mixing.

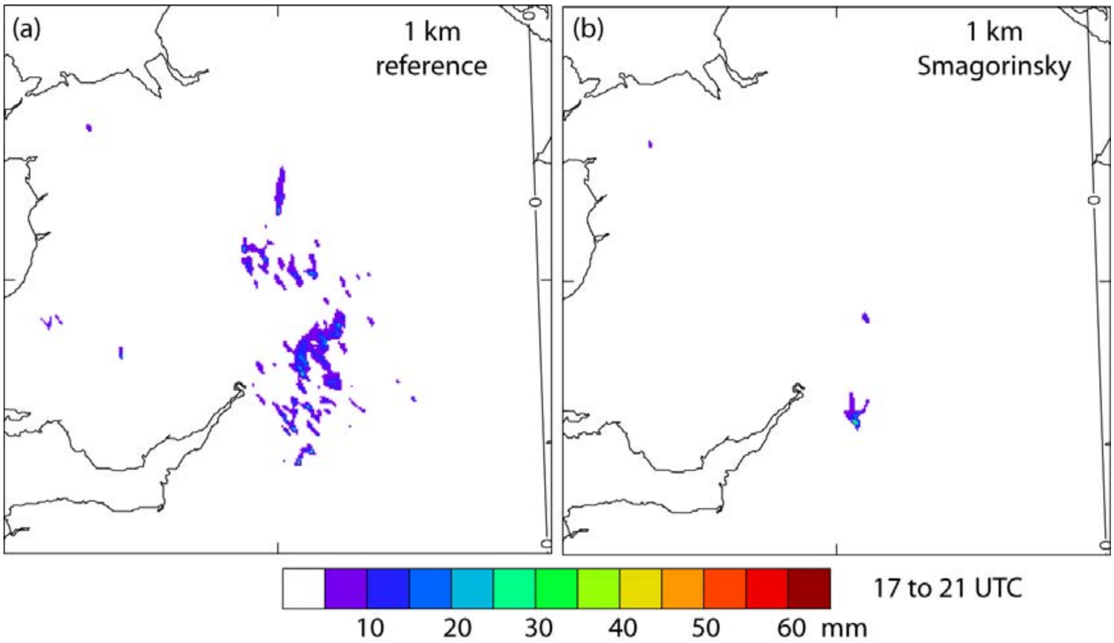


Figure 7-11. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast, and (b) 1 km forecast with Smagorinsky-Lilly type sub-grid-scale turbulent mixing.

It can be seen in Figure 7-11 that the effect is dramatic. On this occasion the sub-grid turbulence has removed nearly all of the convection. The mixing has firstly reduced the initial number of clouds that are able to produce rain and weakened any that did develop. After that, very few secondary cells were able to form. The large reduction in both the amount of rain and the peak rainfall rates is clearly seen in Figure 7-12.

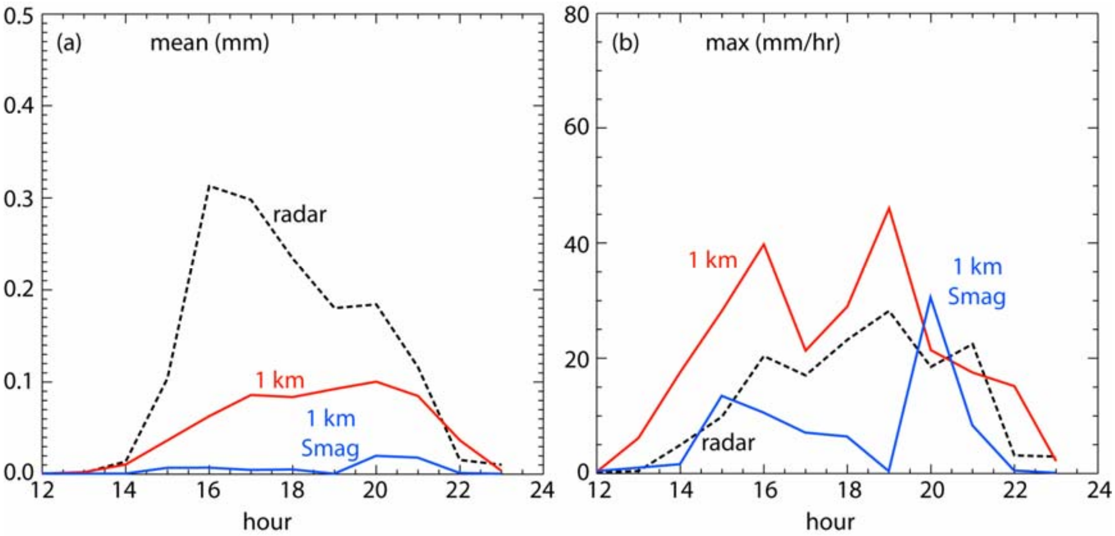


Figure 7-12. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar, the reference 1 km forecast (300x300km domain) and the 1 km forecast with Smagorinsky-Lilly type sub-grid-scale turbulent mixing from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

The behaviour is not surprising given the nature of the event. The small showers in the 1 km model were weakened by the turbulent mixing, which meant that the downdrafts were weaker and the development of new cells at the edge of the downdrafts was therefore less likely. In addition, any new cells would also be weaker. The feedback is the opposite to that seen by changing the grid spacing from 1 km to 1.5 km.

This result doesn't mean that applying a sub-grid-scale turbulent mixing scheme is the wrong thing to do. This was not a tuning exercise. It does mean, however, that the it can have a big impact in some situations and care needs to be taken to make sure that the correct amount of turbulent diffusion is used for a given grid spacing.

7.5 Additional diffusion at 1.5 km

Some numerical diffusion is deliberately added to the high-resolution simulations to give a representation of sub-grid-scale mixing and reduce the chance of numerical instabilities developing at single grid squares. This diffusion is applied to every grid square; it is not selective like the Smagorinsky-Lilly type.

The amount of diffusion that can be applied in this way is determined by the grid spacing and becomes more restricted as the grid spacing becomes shorter. If too much diffusion is added the model is likely to produce spurious grid-scale features and perhaps fail. The previous 1.5 km forecasts have used the same amount of diffusion that was applied at 1 km. This was the largest that is considered safe at that resolution. In this experiment the amount of diffusion has been increased to the largest thought to be appropriate for a1.5 km grid spacing.

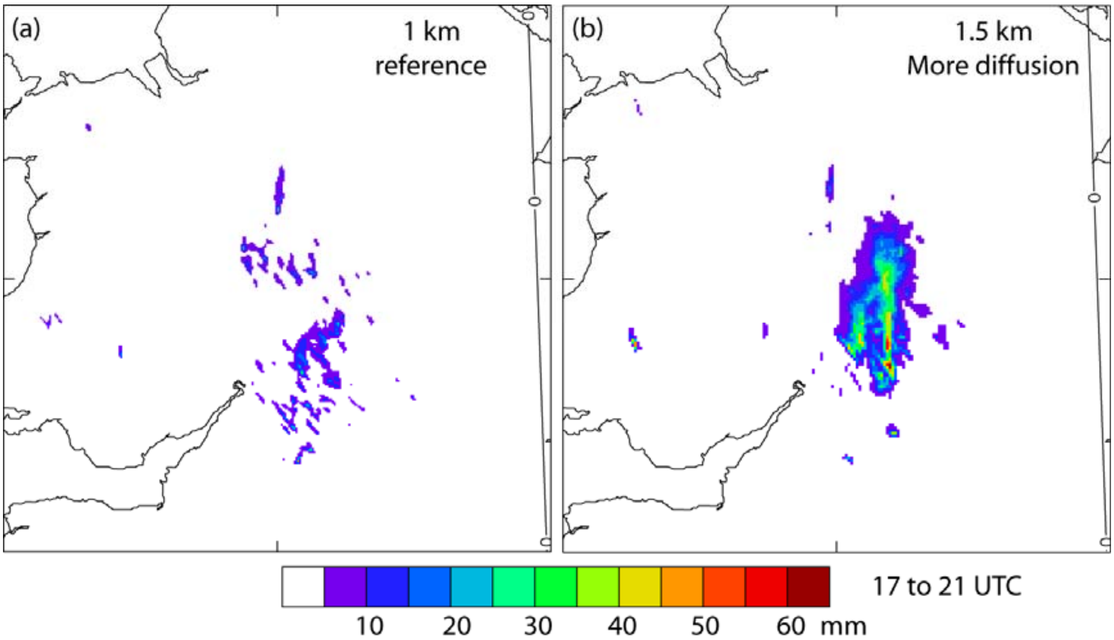


Figure 7-13. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain) with more diffusion.

Figure 7-13 shows that the 1.5 km forecast with more diffusion behaves much like the 1.5 km forecast with less diffusion (Figures 7-5 & 7-9). The main difference compared to the other 1.5 km forecasts is that the accumulation pattern is somewhat smoother, but the location of the rain has not really changed. The mean accumulations are a little less than for the other 1.5 km forecasts and the peak rainfall rates are not as high (Figures 7-6 & 7-10).

For this case, adding more diffusion at 1.5 km has not had a very large impact. Given the particular sensitivity of this case it is reasonable to assume that it would only have a limited effect in general.

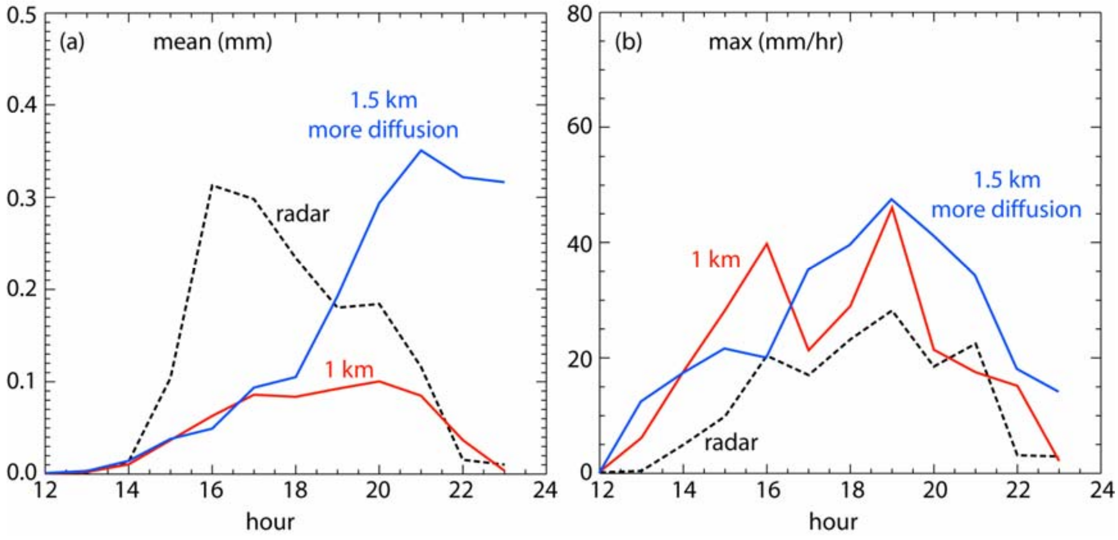


Figure 7-14. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar, the reference 1 km forecast (300x300km domain) and the 1.5 km forecast (450x450 km domain) with additional diffusion from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

7.6 Using the convection scheme at 1.5 km

The operational UK 4 km model uses the convection parametrization scheme (Gregory and Rowntree 1990 + enhancements – version 4A) to represent the effect of showers that can not be resolved on the grid. To make sure that the convection scheme does not operate too much and inhibit the model from representing the larger showers on the grid, it has been strongly restricted according to Roberts (2003) (with c=1200, t=0.5 – the same as the 'very weak conv scheme' in section 4). Deng and Stauffer (2006) also show the value of using a convection parametrization scheme at 4 km (although it is not used in the same way).

In the 1 / 1.5 km models the convection scheme is usually switched off because it is thought that the grid spacing is small enough to resolve the showers sufficiently well. However, even on a 1 /1.5 km grid there will be clouds that can not be resolved as well as the sub-grid-scale mixing within the clouds that has already been discussed.

In this experiment the convection scheme was used in the 1.5 km forecast (larger domain) in exactly the same way it is used in the operational UK 4 km model.

The rainfall accumulations produced by this 1.5 km forecast (Figure 7-15) were much closer to the other 1.5 km forecasts (Figures 7-5, 7-9 & 7-13) than to the 1 km reference forecast. There were, however, some noticeable differences between this forecast and the other 1.5 km forecasts. The higher rainfall amounts were more fragmented and there wasn't the development of a large storm complex that persisted into the night. The mean rainfall accumulations dropped away sharply after 22 UTC (Figure 7-16). In addition, the peak rainfall rates were reduced.

This appears to be a more realistic forecast. The effect of the convection scheme has been to weaken the storm cells enough to limit the formation of new cells at the outflow boundaries and allow the storm system to eventually decay late in the evening. What it hasn't done though, is change the general area over which the storms are produced because that is determined more by the initial low-level temperature and humidity fields.

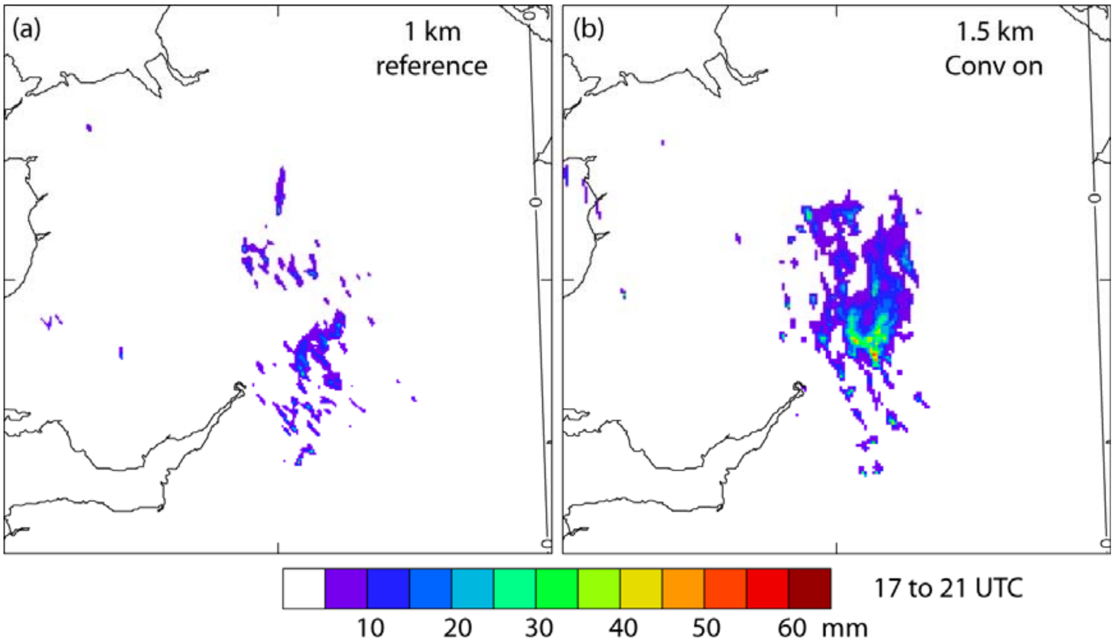


Figure 7-15. Rainfall accumulations over the period 17 to 21 UTC 4th July 2006 from 1 km forecasts starting from 06 UTC (a) reference 1 km forecast (300x300 km domain), and (b) 1.5 km forecast (450x450 km domain) with a restricted version of the convection parametrisation scheme switched on.

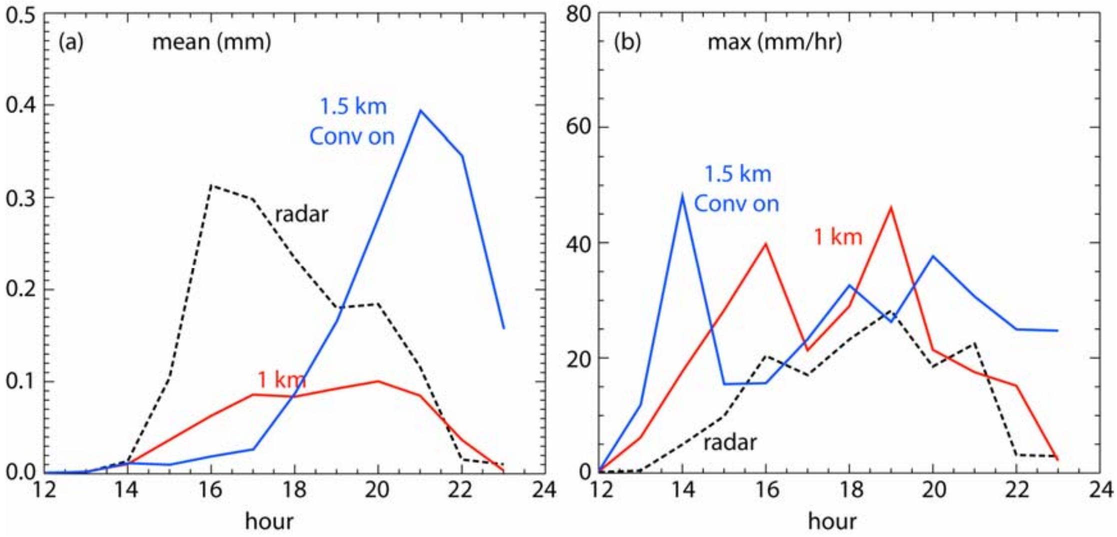


Figure 7-16. (a) Mean hourly rainfall accumulations over the 1 km model domain for the period 12 to 23 UTC 4th July 2006 from radar, the reference 1 km forecast (300x300km domain) and the 1.5 km forecast (450x450 km domain) with a limited version of the convection parametrization scheme from 06 UTC. (b) Same as (a) for snapshot maximum rainfall rates every hour.

Once again, this experiment has shown the sensitive nature of this case and that there is a need to represent the sub grid clouds somehow, even in a 1.5 km model, either with the convection scheme in some form, or more likely with the Smagorinsky-Lilly type sub-grid-scale turbulence set up for this resolution. One of the difficulties with using the convection scheme in this way is that even though the later representation of the storms was better, the initial triggering was delayed too much. This can be seen by comparing the mean accumulation graph (Figure 7-16) with the other mean accumulation graphs. It is a consistent signal when the restricted convection scheme is used.

7.7 Verification scores

7.7.1 Comparison against radar

The Fractions Skill Score (FSS) measure (Roberts & Lean 2008, see appendix 1 and section 4) has been used to compare the skill of the rainfall accumulation forecasts discussed in this section with radar as a function of spatial scale. The FSS values plotted against spatial scale are shown in Figure 17-7(a) for a threshold of 4 mm. The 1 km reference forecast was closer in skill, over all scales, to the 1.5 km forecasts than the other 1 km forecasts. The reason for this was that the under-prediction at 1 km affected the skill just about as much as the over-prediction at 1.5 km. The most skilful forecast at large scales was the 1.5 km forecast with extra diffusion because the over-prediction was less than in the other 1.5 km forecasts. The effect of the bias is seen most dramatically in the 1.5 km forecast with 38 levels and with the sub-grid-scale turbulence. The under-prediction of those two forecasts has resulted in much lower skill at all scales.

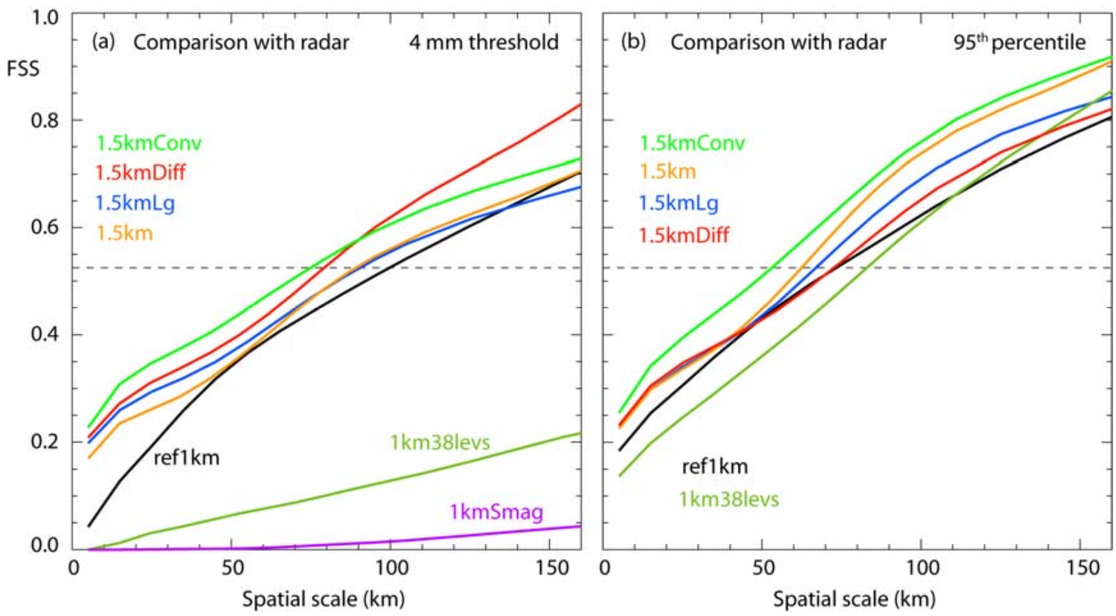


Figure 7-17. Graphs showing the change of forecast skill with spatial scale compared to radar for rainfall accumulations over the period 17 to 21 UTC 4th July 2006, for all the model experiments, using the Fractions Skill Scores (FSS) as the measure of skill. (a) Using an accumulation threshold of 4mm, (b) using as a threshold the 95th percentile value taken from all points in the domain (including zeros) (see appendix in section 4 for more information about the FSS)

The 95th percentile threshold has been used to remove the effect of the bias on the scores (because the number of pixels exceeding the threshold is the same) (Figure 7-17(b). Now the difference in skill between the forecasts is less pronounced. This reflects what was observed; that the biggest differences between the forecasts were in the amount of rain rather than a change in the position of the rain. The most skilful forecast was the 1.5 km with the convection scheme, which achieved the target level of skill (dashed line) at around 50 km. The least skilful forecast was the 1 km with 38 levels, which achieved the target level of skill at around 80 km.

7.7.2 Comparison against each other

The same verification technique can be used to compare the forecasts with each other to examine how a particular change impacts the forecast over different spatial scales. This is probably a more useful exercise for this case than comparing with radar because it is the model behaviour in representing the showers that is under scrutiny, and not so much where they occurred, since that is largely determined by the initial fields.

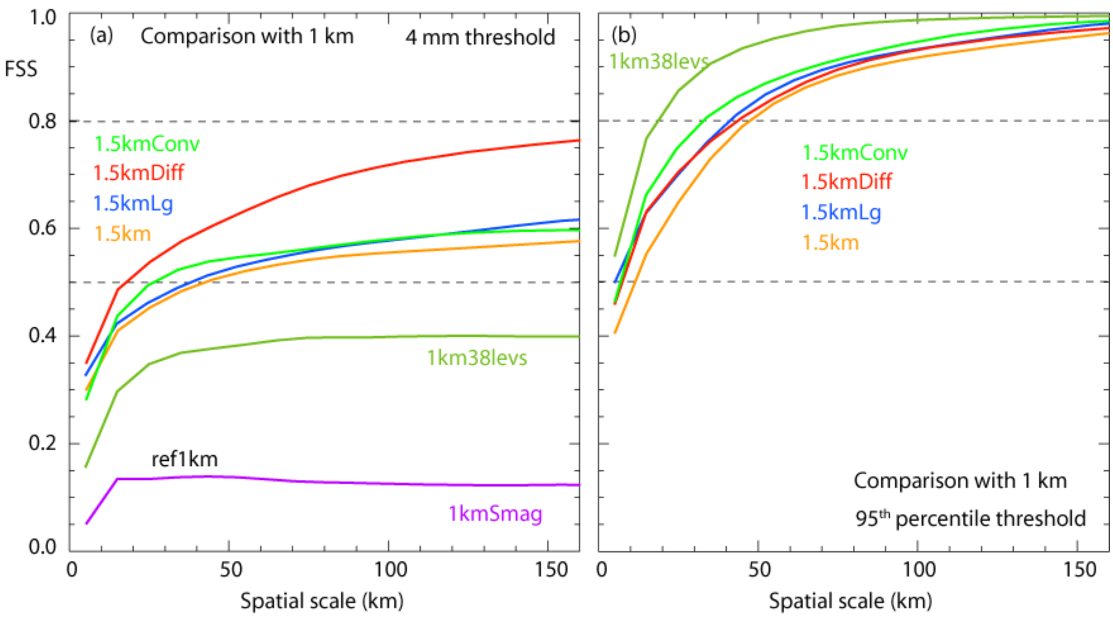


Figure 7-18. Graphs showing the change in the Fraction Skill Score (FSS) with spatial scale compared to the 1 km reference forecast for rainfall accumulations over the period 17 to 21 UTC 4th July 2006, for all the model experiments. (a) Using an accumulation threshold of 4mm, (b) using as a threshold the 95th percentile value taken from all points in the domain (including zeros) (see appendix in section 4 for more information about the FSS). The dashed lines discriminate between a good agreement between forecasts (FSS>0.8), a reasonable agreement (0.5<FSS<0.8) and a poor agreement (FSS<0.5).

Figure 7-18 shows how much agreement there was between the reference 1 km forecast and the other forecasts. For the 4 mm threshold the impact of all the changes was large at all scales (FSS does not exceed 0.8 at any scale). This was again due to the change in the amount of rain rather than its location. By far the biggest difference occurred when the sub-grid-scale turbulence was introduced and the amount of rain was much less (purple line). The differences compared to the 1.5 km forecasts were less because the increased amount of rain at 1.5 km was less dramatic than the decrease in rain from the two 1 km forecasts.

There was a much better agreement between the experimental forecasts and the reference forecast in terms of the spatial distribution of the rain (Figure 7-18(b)). The biggest differences were at the smaller spatial scales (FSS < 0.5 for most forecasts). At scales larger than ~50 km, FSS > 0.8 for all the experiments, which is a good agreement and again indicates that the major impact of the changes was on the amount of rain rather than its location. In terms of the location of the higher rainfall accumulations, it was the 1 km forecast with 38 levels that was closest to the reference 1 km forecast at all scales. The purple line is missing (inclusion of sub-grid scale turbulence) because there was not enough rain to fill 5% of the domain (i.e. meet the 95th percentile threshold).

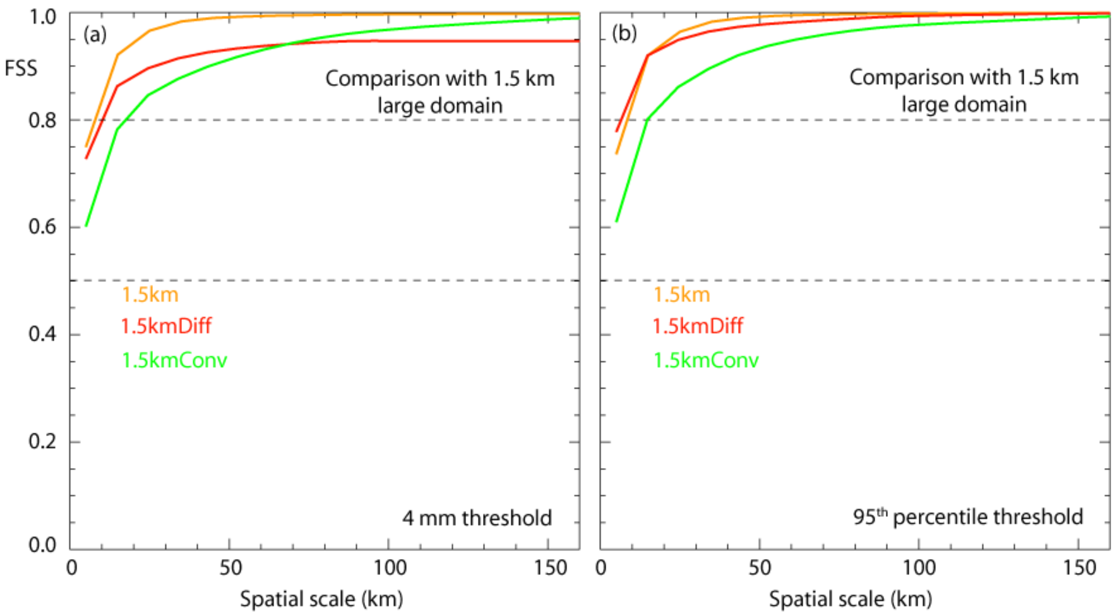


Figure 7-19. Graphs showing the change in the Fraction Skill Score (FSS) with spatial scale for the 1.5 km forecast (300x300 km domain), the 1.5 km forecast with additional diffusion and the 1.5 km forecast with convection parametrization compared to the 1.5 km forecast (450x450 km domain) for rainfall accumulations over the period 17 to 21 UTC 4th July 2006. (a) Using an accumulation threshold of 4mm, (b) using as a threshold the 95th percentile value taken from all points in the domain (including zeros) (see appendix in section 4 for more information about the FSS). The dashed lines discriminate between a good agreement between forecasts (FSS>0.8), a reasonable agreement (0.5<FSS<0.8) and a poor agreement (FSS<0.5).

It is also informative to compare just the 1.5 km forecasts. A comparison of the 1.5 km forecast on the larger domain with the other 1.5 km forecasts is shown in Figure 7-19. Notice that the curves for the 4 mm and 95th percentile thresholds are similar because the amount of rain did not differ so much between the 1.5 km forecasts. For both thresholds the 1.5 km forecast on the smaller domain (1 km domain) (orange line) is identical at scales larger than 50 km (FSS = 1), and even at scales as small as ~10 km the FSS is around 0.8 (defined as a good agreement in section 4). This shows that the change in domain size did have some impact at smaller scales, as might be expected for such a sensitive case, but reassuringly, the impact wasn't great and mostly affected the smallest scales. The introduction of more diffusion at 1.5 km (red line) had some affect at smaller scales but then the line goes flat (4 mm threshold) with a FSS of ~ 0.95 beyond around 30 km. The flattening of the line shows an impact at all scales due to there being a little less rain in the forecast with more diffusion. The biggest difference occurred when the convection scheme was introduced. Again, though, the impact was mostly at smaller scales. At scales larger than ~40 km the agreement was 'good' (FSS > 0.8) and at scales smaller than ~40 km the agreement was 'reasonable', so forecast differences were small at all but the small scales. At the largest scale FSS = 1 for the 4 mm threshold, showing that the number of pixels above 4 mm did not change with the introduction of the convection scheme. It therefore appears that the use of the convection scheme can improve the forecast over the smaller to intermediate scales (Figure 17-7), without changing the overall rainfall pattern very much.

7.8 Summary

- The change to 38 vertical levels reduced the amount of rain. The number of initial cells was very similar, but they were weaker and the subsequent development of new cells was less.
- The change to 1.5 km increased the amount of rain. The storms were larger and stronger and that resulted in the initiation of more new cells and eventually a large storm complex that persisted into the night when it should have decayed.
- The change in the size of the domain at 1.5 km had a small impact at small scales and almost no impact at larger scales. I.e. the area of rain was not changed.
- The addition of the sub-grid-scale turbulence scheme at 1 km reduced the amount of rain considerably. The showers were very much fewer and weaker.
- The increase in the amount of diffusion at 1.5 km smoothed the rainfall pattern and reduced the amount of rain a little, but overall signal was unchanged.
- The introduction of a restricted version of the convection scheme at 1.5 km had a positive impact. It reduced the number and intensity of storm cells and then stopped the formation of the unrealistic larger storm complex. However, it also delayed the first initiation too much.
- These studies have shown that there can be a great sensitivity to changes to the horizontal and vertical resolution and representations of sub-grid scale mixing or convection in this type of case. The biggest impact is on the amount and duration of the rain. There was little change to the general area of rainfall, which was controlled more by the initial fields.

8. Other Cases G to J

During the course of the project some other notable flood-producing events occurred. Four of them are documented briefly here. They were all simulated by the 1 / 1.5 km model after the event.

8.1 Case G 10th May 2006

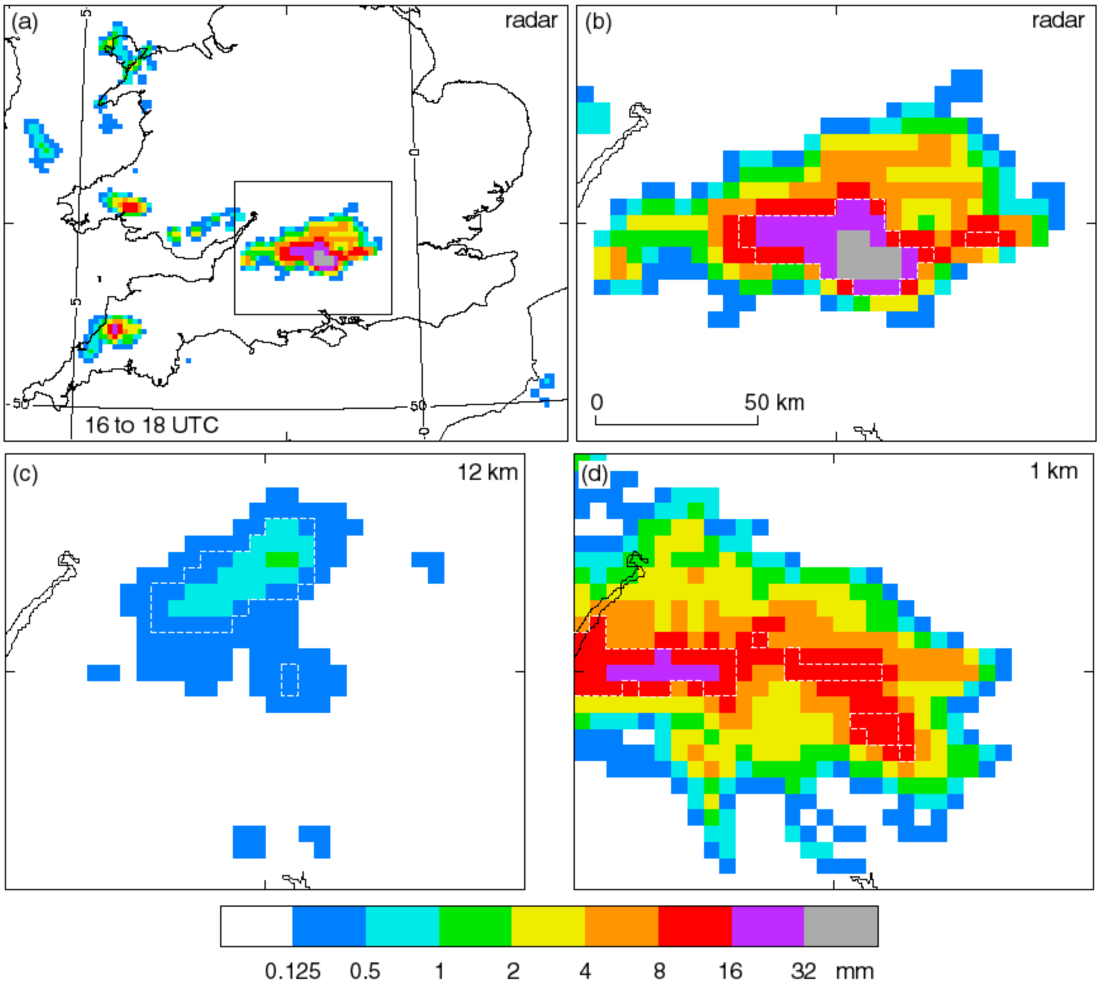


Figure 8-1. Rainfall accumulation over the period 16 to 18 UTC 10th May 2006, from (a & b) radar, (c) 12 km forecast from 12 UTC, (d) 1 km forecast from 12 UTC.

On the afternoon of the 10th May 2006 some heavy thunderstorms developed in Southern England and organised into a large storm complex. This storm system gave large hailstones in places and caused flash flooding in South Wales in the evening. The operational 12 km forecasts produced small rainfall amounts and nothing that could be described as being very much like the storm that occurred. Because it was such a marginal situation the operational 4 km model did produce a storm in some forecasts, but not in most. The storms that were generated in the 4 km model were roughly in the correct location, but not extensive enough. Simulations with the 1 km model were run after the event and the performance was good. One of these forecasts is shown and compared to the equivalent 12 km forecast (Figure 8-1).

The 1 km model produced rainfall totals locally that were as high as seen by radar, but when averaged to the radar grid they were less because the individual storm cells were not large enough. Nevertheless, the signal from the 1 km model for a severe storm over southern England would have been useful. The 1 km model also managed to generate back-building storm cells to the east of the storm complex as it moved westward. This behaviour was also observed in reality.

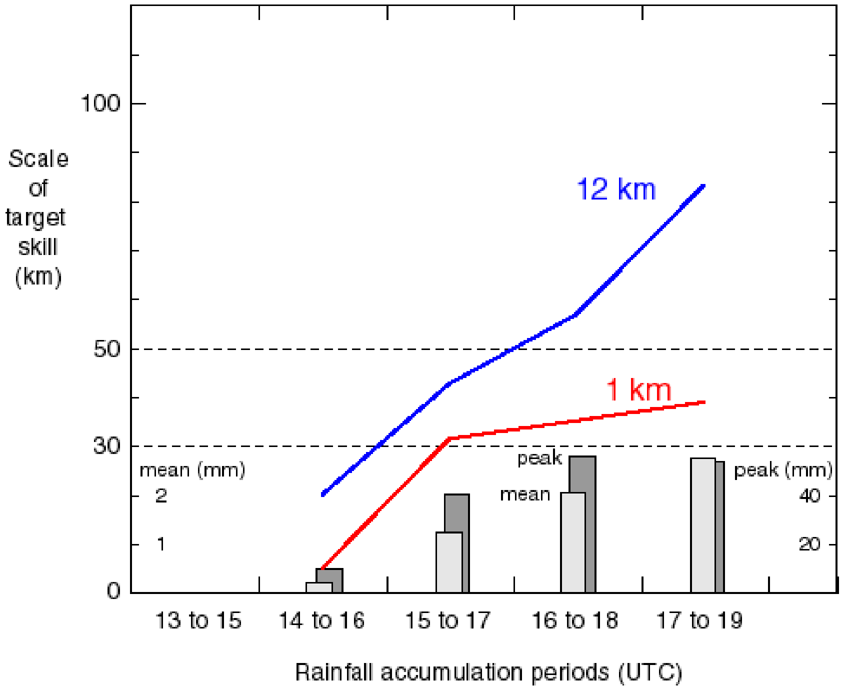


Figure 8-2. Graph showing the scale at which a target or useful level of skill was reached by the 12 km forecast (blue line)and 1 km forecast (red line) for 2-hourly rainfall accumulations (lower is better). See appendix 1 for more information. The light shaded bars show the mean 2-hour rainfall accumulations (over small domain in Figure 8-1) and the grey bars the peak rainfall rates from radar (5x5km grid squares).

The graph in Figure 8-2 shows that the 1 km forecast was spatially more accurate that the 12 km forecast. This method for determining the spatial accuracy (Roberts and Lean 2008, see appendix 1) can be used to determine the scale over which a forecast had useful skill. In this case the 1 km model had useful skill over a scale of 30 to 40 km at the times when the storm was at its peak compared to 40 to 80 km from the 12 km model.

8.2 Case H 13th August 2006

In the afternoon of the 3rd August 2006, localised thunderstorms developed into bands to the west of London. Along one of those bands new storm cells kept regenerating and this lead to very large rainfall amounts and flash flooding.

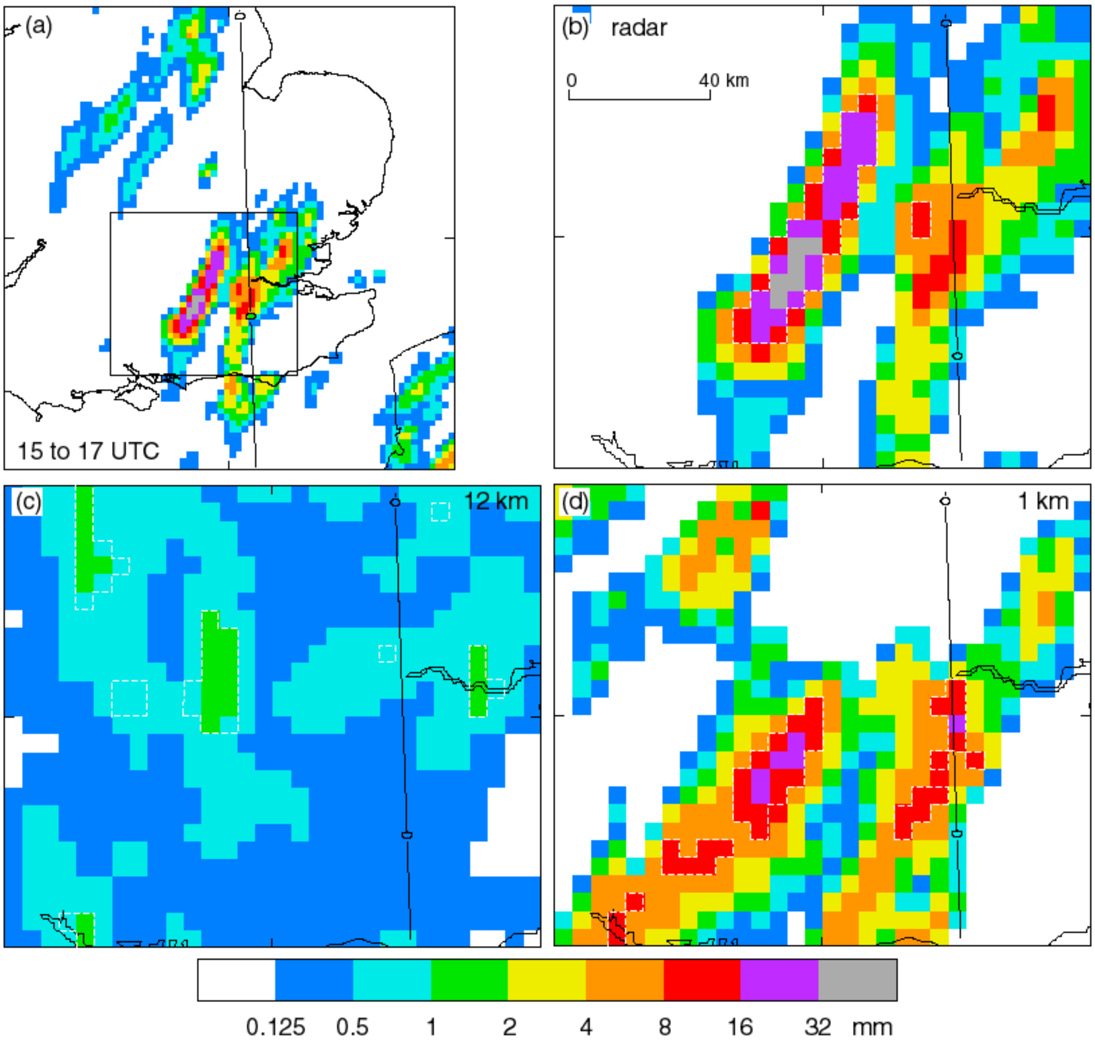


Figure 8-3. Rainfall accumulation over the period 15 to 17 UTC 13th August 2006, from (a & b) radar, (c) 12 km forecast from 12 UTC, (d) 1 km forecast from 12 UTC.

As for the previous case the 12 km forecasts were poor and gave no indication of very large rainfall totals (Figure 8-3(c)). The 4 km forecasts were better and did give an indication of heavy showers and the possibility of large totals. The 1.5 km model when run after the event gave similar results to the 4 km model when run from early on in the day, but the latest 1.5 km forecast (from 13 UTC) did a very good job of generating two bands of convection and high rainfall totals (Figure 8-3(d)). Just like the previous case the totals were locally very high in the 1.5 km forecast and gave a good signal that flood-producing rain was possible, but they were not as high as observed when averaged to a 5 km grid. Again the reason was that the storms were somewhat too small. It is also worth noting that the 1.5 km model was able to generate new storm cells in the same location as previous cells. This is the sort of storm dynamics that needs to be captured if flood-producing storms are going to be predicted from these types of events.

Figure 8-4 shows that, in terms of spatial accuracy, the 1.5 km forecast achieved a useful level of skill at scales of less than 10 km early on, which increased to more than 50 km later. The equivalent 12 km forecast achieved the same level of skill at scales of 50 to 80 km throughout (and never produced high rainfall totals).

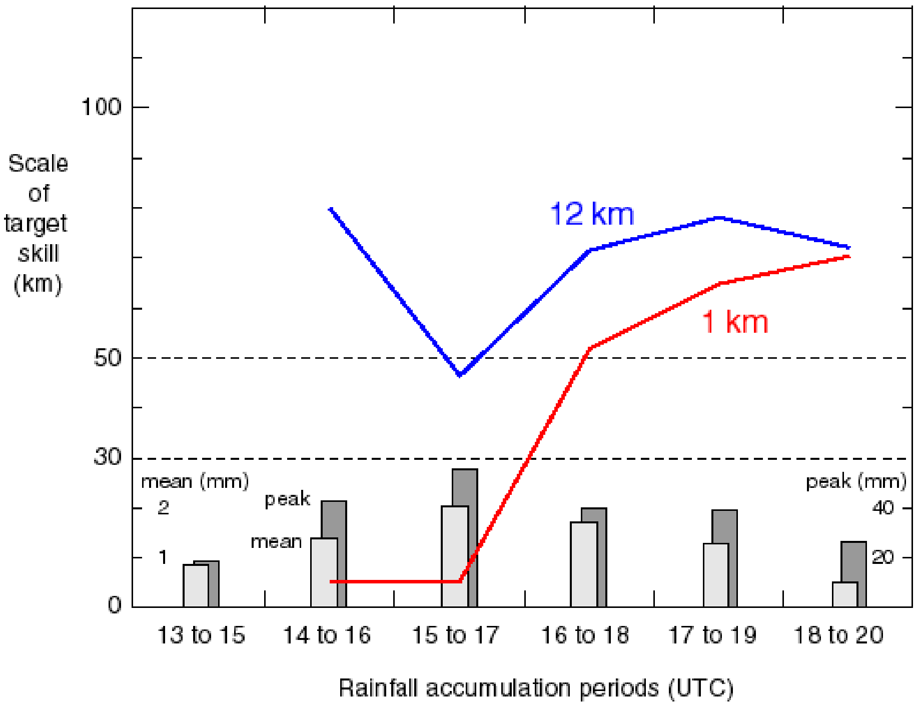


Figure 8-4. Graph showing the scale at which a target or useful level of skill was reached by the 12 km forecast (blue line)and 1 km forecast (red line) for 2-hourly rainfall accumulations (lower is better). See appendix 1 for more information. The light shaded bars show the mean 2-hour rainfall accumulations (over small domain in Figure 8-1) and the grey bars the peak rainfall rates from radar (5x5km grid squares).

8.3 Case I 25th June 2007

This was one of the major flood events in 2007. Very high rainfall totals occurred as an active frontal system became almost stationary over Yorkshire for the whole day. Figure 8-5 shows that rain gauge measurements for the 24th and 25th exceeded 100 mm in the Hull and Sheffield areas. Nearly all of that rain fell on the 25th. The radar accumulations for the 25th are considerably lower, especially in the east, because the Ingham radar was not working that day. Severe flash flooding and river flooding occurred in parts of Hull and Sheffield and it became a major news item.

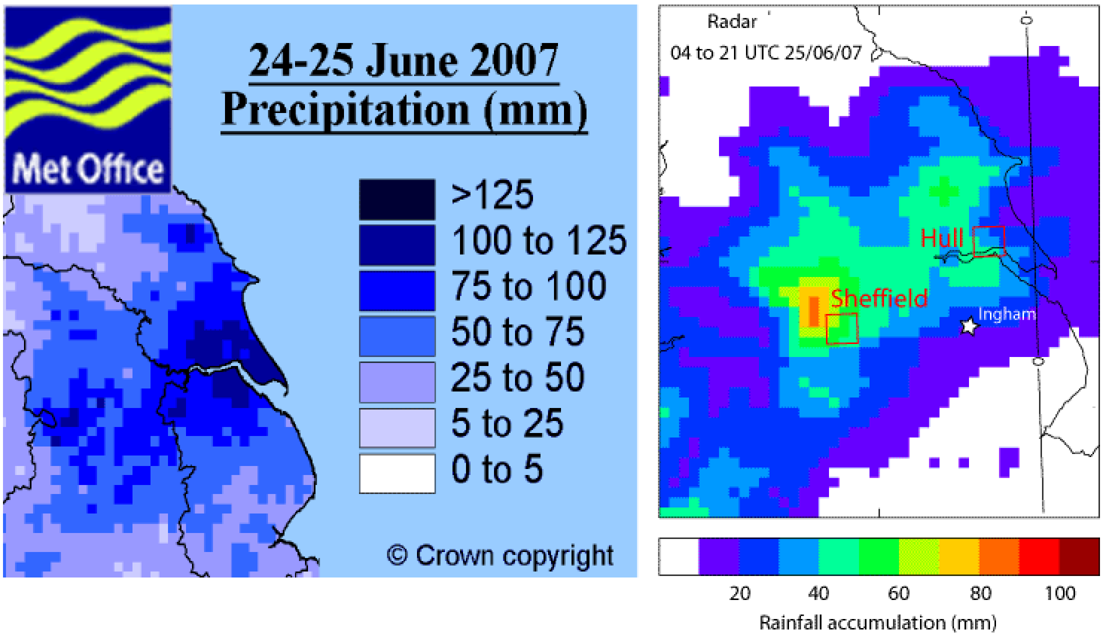


Figure 8-5. (a) Rainfall accumulations from gauge measurements for the 24th and 25th June 2007 courtesy of the National Climate Information Centre (NCIC). (b) Rainfall accumulations from radar over the period 04 to 21 UTC 25th June 2007. The Ingham radar is marked with a star. The red boxes enclose the Hull and Sheffield urban areas.

This event was different in nature to the two previous cases because the rainfall was frontal and not convective. The rainfall amounts were high because heavy (but not torrential) rain fell in the same places for a long time. This is the sort of event we would expect a 12 km model to be able to capture. It is able to resolve frontal rain areas provided that they are broad enough – and this one was. The forecast uncertainty in this type of case should come mostly from getting the position of the front correct.

Figure 8-6 shows that the operational NAE model (12 km) forecasts were able to give a very good indication of very high rainfall amounts 24 to 36 hours in advance (in fact even longer than that). Hence, the Met Office was able to issue early severe weather warnings. The five panels in Figure 8-6 show the variability in the location of the front from forecast to forecast, and that this indeed was the major source of uncertainty. This is why the warnings were issued in probabilistic terms. The rainfall amounts were consistently high from run to run (except for the forecast from 18 UTC 24th).

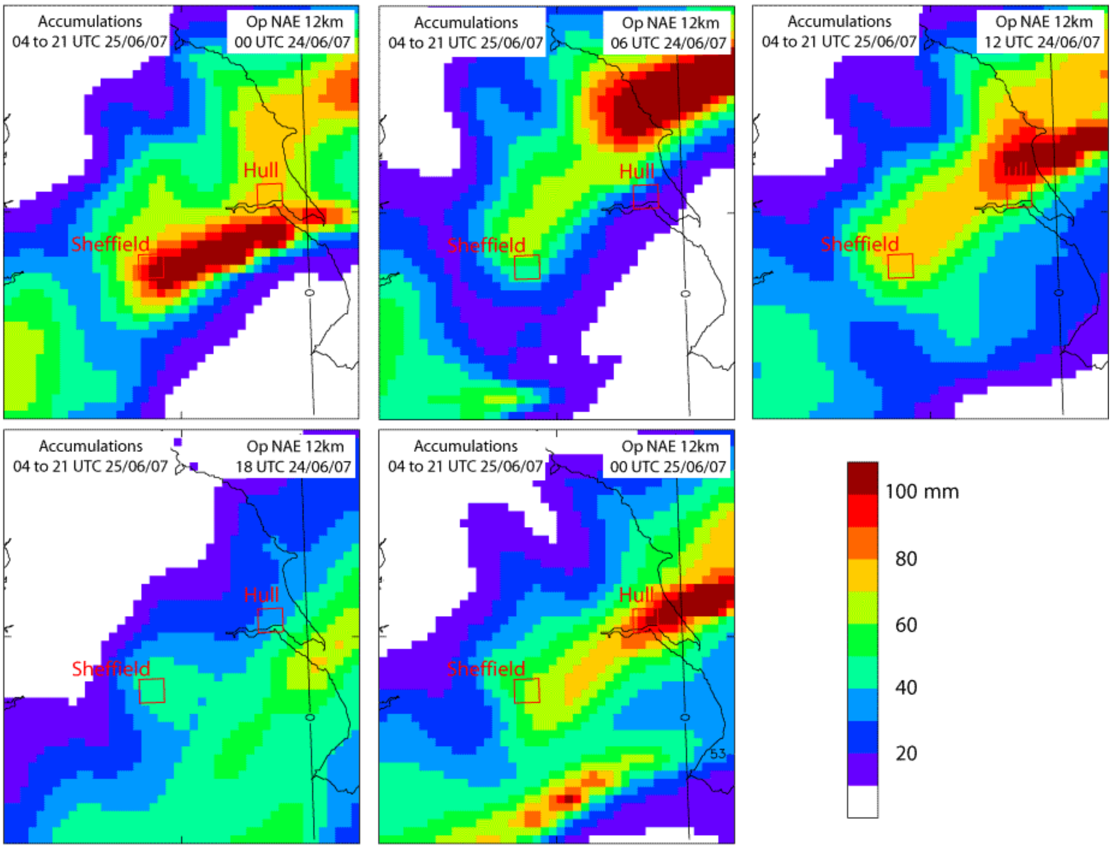


Figure 8-6. Rainfall accumulations over the period 04 to 21 UTC 25th June 2007 from the North Atlantic European (NAE) 12 km operational forecasts starting from 00, 06, 12 and 18 UTC on the 24th and from 00 UTC on the 25th. The red boxes enclose the Hull and Sheffield urban areas.

In this type of situation it is expected to be more difficult for a high-resolution model to add much improvement to the forecast than would be the case for convective situations. Two forecasts from the UK 4km model and from a 1.5 km version are shown in Figures 8-7 and 8-8. The first thing to notice is that they both give essentially the same signal and that they are both similar to the 12 km forecasts. As stated previously, this is not surprising as frontal rain can be resolved at 12 km and the forecasts were already good.

There were some differences though, and these should be noted. Both the 4 km and 1.5 km forecasts were able to produce a peak in rainfall amounts to the west of Sheffield and along the coast to the north of Hull. The 1.5 km forecasts did this more than the 4km forecasts. A comparison with the rain gauge accumulations in Figure 8-5 shows that the additional rain in those areas was closer to reality. The reason for the higher totals was orographic enhancement of the rain due to the seeder-feeder mechanism where the easterly wind blew on to the Pennines west of Sheffield and the North York Moors north of Hull. The mechanism is just the same as it was for the Carlisle flood (Case D). It would appear that the high-resolution forecasts were able to add important local detail that matters for small catchments in hilly areas, provided that the frontal rain is reasonably well positioned in the first place.

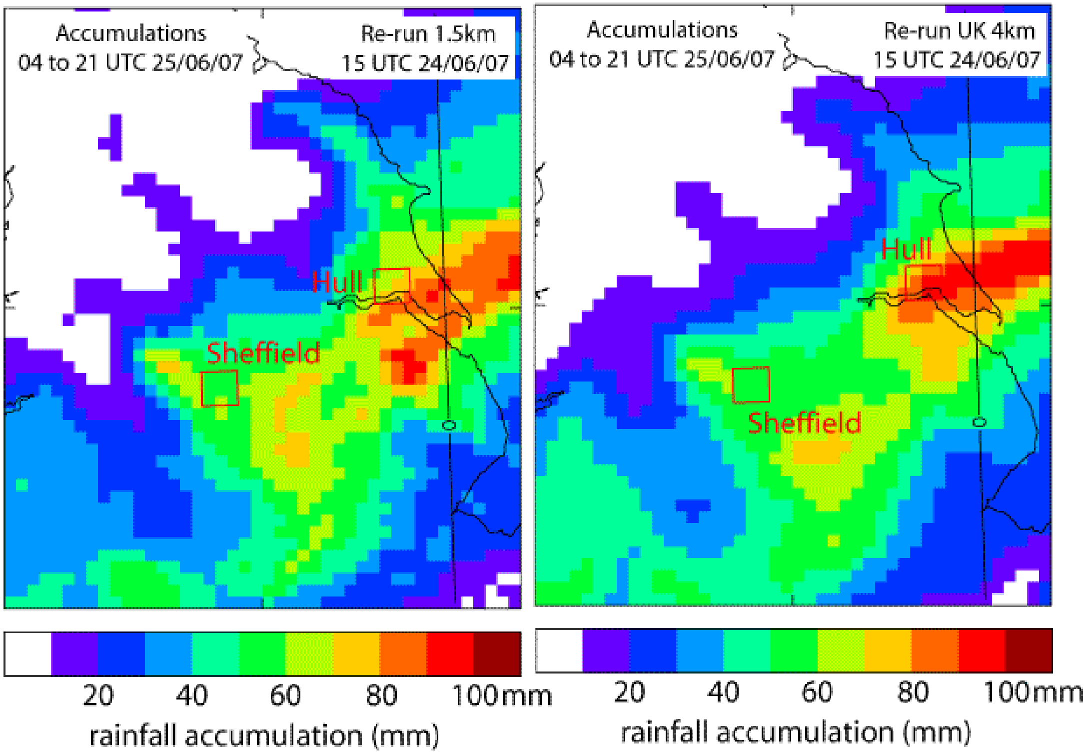


Figure 8-7. Rainfall accumulations over the period 04 to 21 UTC 25th June 2007 from simulations starting at 15 UTC on the 24th using (a) the 1.5 km model and (b) the 4 km model. The red boxes enclose the Hull and Sheffield urban areas.

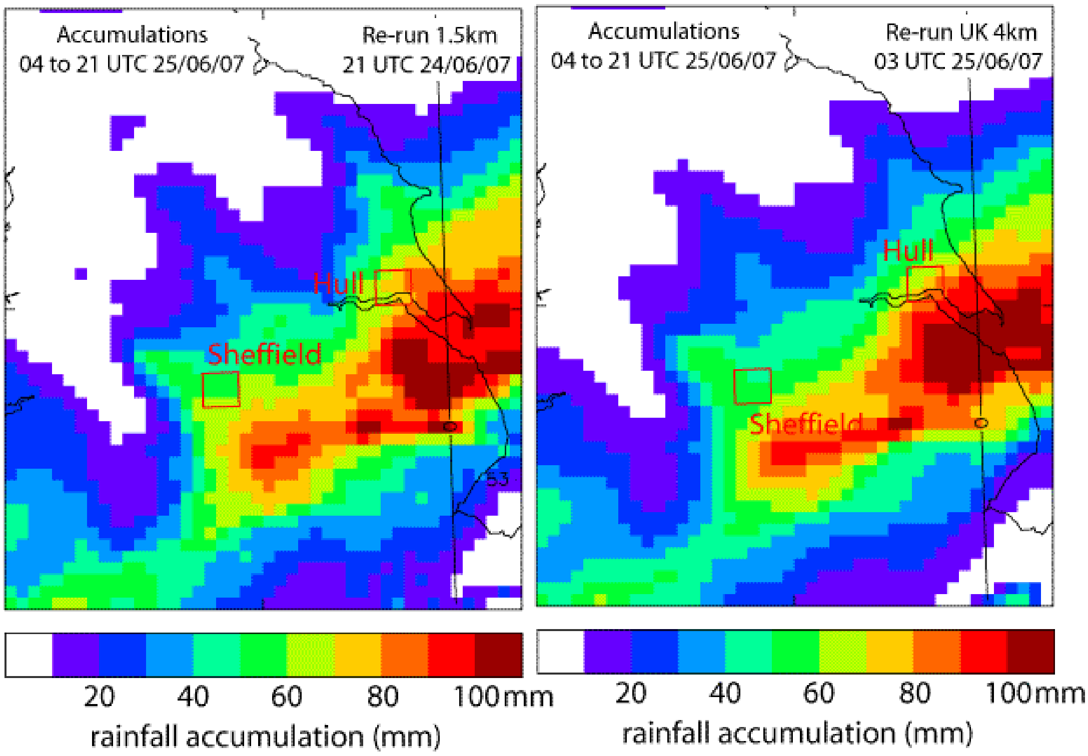


Figure 8-8. Rainfall accumulations over the period 04 to 21 UTC 25th June 2007 from simulations starting at 21 UTC on the 24th using (a) the 1.5 km model and (b) the 4 km model. The red boxes enclose the Hull and Sheffield urban areas.

8.4 Case J 20th July 2007

This was another of the major flood events in 2007. A band of heavy and thundery rain moved slowly north across Central Southern England and the West Midlands during the day. Figure 8-9 shows that in excess of 100 mm of rain was measured over a large area and most of this fell on the 20th. The rainfall caused serious flash flooding over a very wide area during the day and this turned into serious river flooding over subsequent days. In addition to the main band of thundery rain, there were also thunderstorms further east (to the west of London) early in the morning that also caused flash flooding.

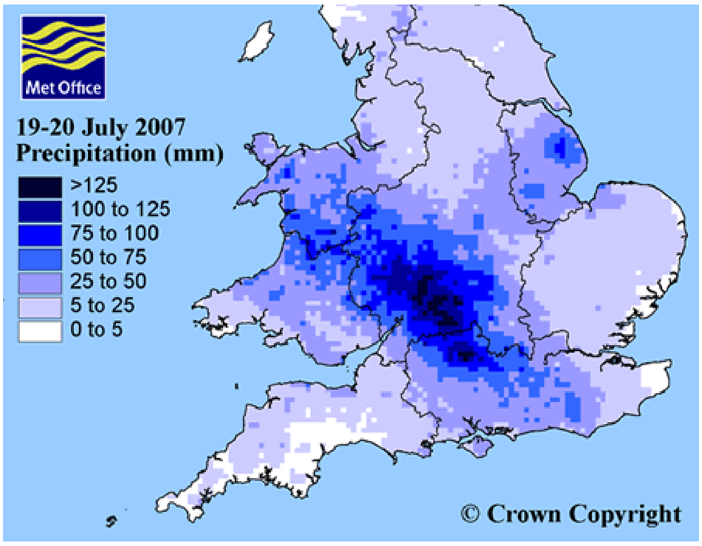


Figure 8-9. Rainfall accumulations from gauge measurements for the 19th and 20th July 2007 courtesy of the National Climate Information Centre (NCIC).

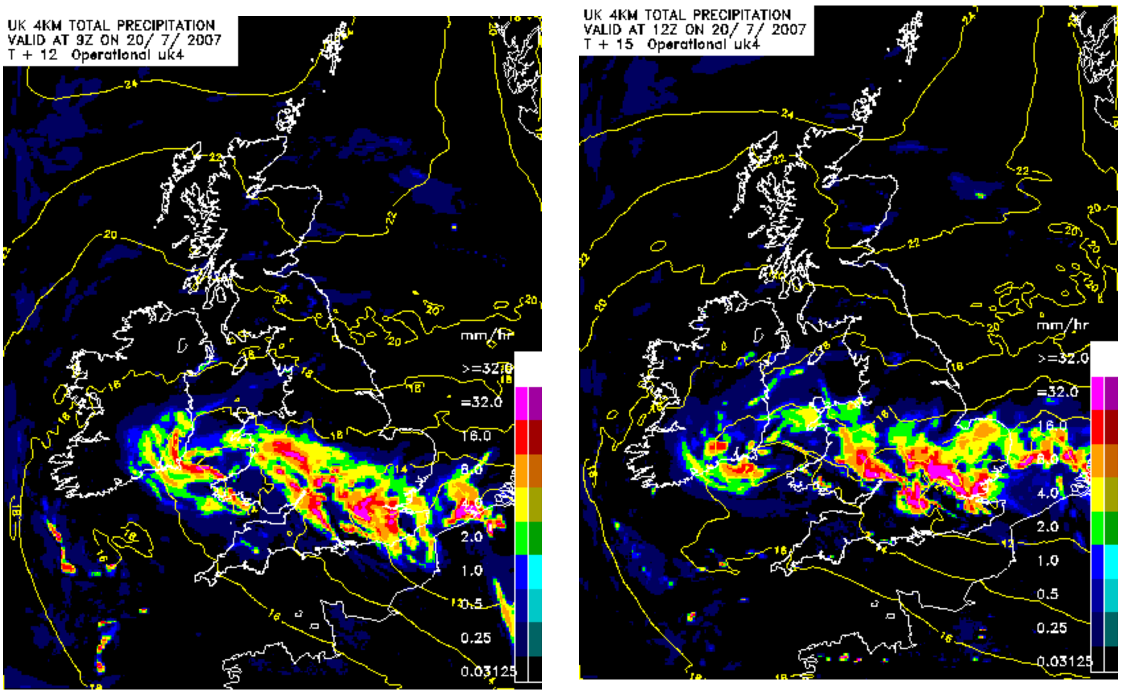


Figure 8-10. Rainfall rates at 09 and 12 UTC 20th July 2007 from the operational UK 4 km forecast starting at 21 UTC on the 19th.

The Met Office forecasts of this event were excellent. Even a few days in advance there was a good signal. The 4 km model forecasts were particularly useful. An example of the very high rainfall rates produced by the 4 km model is shown in Figure 8-10.

The 1.5 km model was run after the event. The rainfall accumulations from these forecasts are shown in Figures 8-11 to 8-13. The rain was so heavy over such a wide area that the radars are thought to have underestimated the amounts because of attenuation of the beams. Given that the radar was probably underestimating, the 1.5 km simulations are extremely good and if that model had been run operationally, it would, like the 4 km model, have given a very good indication of what was to come.

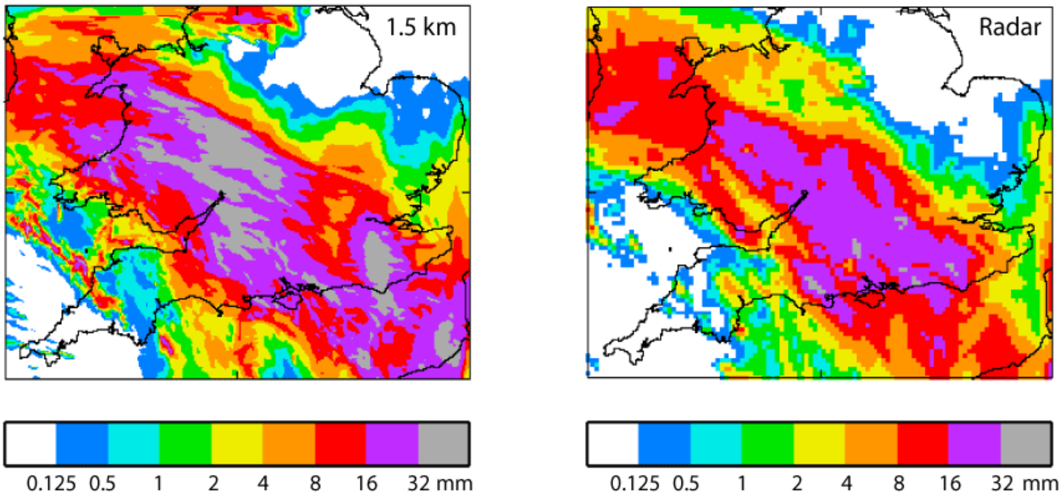


Figure 8-11. Rainfall accumulations over the period 00 to 12 UTC 20th July 2007 from (a) radar, and (b) a 1.5 km forecast starting at 00 UTC on the 20th. Courtesy of Humphrey Lean, Joint Centre for Mesoscale Meteorology (JCMM), Met Office.

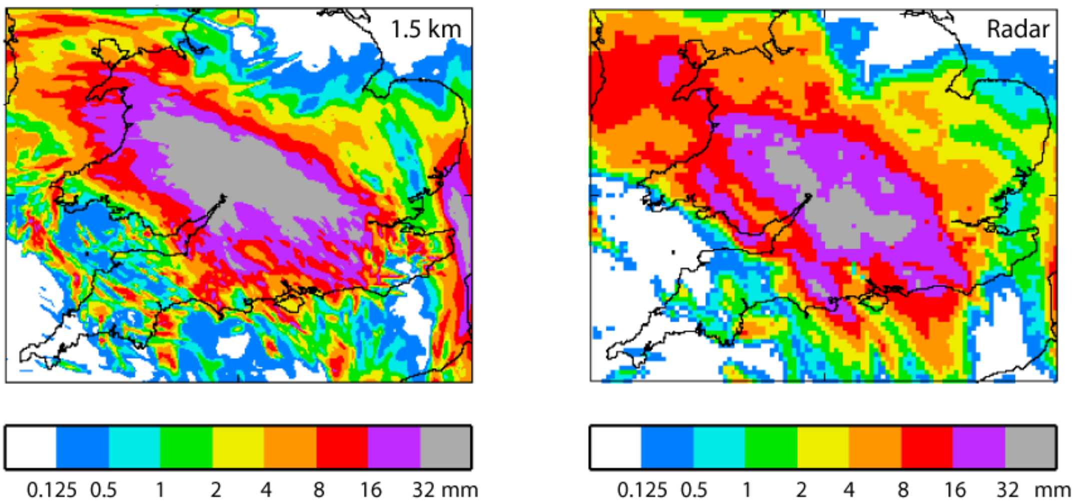


Figure 8-12. Rainfall accumulations over the period 03 to 15 UTC 20th July 2007 from (a) radar, and (b) a 1.5 km forecast starting at 03 UTC on the 20th. Courtesy of Humphrey Lean, Joint Centre for Mesoscale Meteorology (JCMM), Met Office.

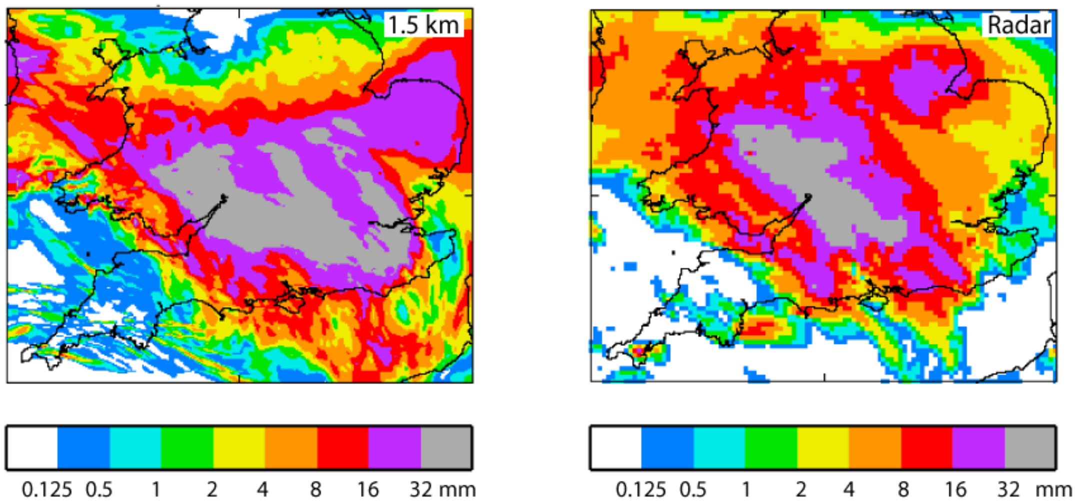


Figure 8-13. Rainfall accumulations over the period 06 to 18 UTC 20th July 2007 from (a) radar, and (b) a 1.5 km forecast starting at 06 UTC on the 20th. Courtesy of Humphrey Lean, Joint Centre for Mesoscale Meteorology (JCMM), Met Office.

9. Probabilistic output products

One of the findings and main themes of this work has been that a ‘storm-resolving’ model (~1/1.5 km) is able to produce more accurate forecasts of extreme events (convective in particular) than the 12 km or 4 km models on scales that are useful for flood warning. Such a model should be (and is being) developed for operational use because it is the best route towards improved forecasts of high-impact weather.

Once we have such a model though, we need to know how to use it. We need to recognise that there will always be uncertainty in the forecasts and this may become more apparent as we continue to expect more precision. Some forecasts will naturally be more skilful than others because some situations are inherently more predictable than others. The output should be therefore be presented in ways that are useful for specific applications, but also take account of and give an indication of the forecast uncertainty. The best way to assess forecast uncertainty is to use an ensemble forecast system and make the assumption that the more each of the forecasts in the ensemble differ from one another, the more uncertain the forecast will be. The problem we face in the near future is that we will not have an ensemble of storm-resolving forecasts until computer resources allow that possibility. In the meantime another approach is needed.

This section shows probabilistic outputs that could have been generated from 1 / 1.5 km forecasts of some of the events studied in this project using a fuzzy neighbourhood approach rather than requiring an ensemble. The method to generate the probabilistic forecasts this way has been described in the Storm Scale Modelling Project (Roberts 2004, 2005) and is similar to Theis *et al* 2005. It also forms the basis of the Fractions Skill Score verification method (Roberts 2005, Roberts and Lean 2008, appendix 1) that has been used in this work; the advantage being that verification scores can be directly translated into appropriate scales over which to compute the probabilities. The purpose here is simply to give an idea of a subset of what can be done rather than go into the methodology in detail, but a very brief outline is given below.

The probability generation assumes that a rainfall pixel in the forecast is equally likely to occur at any other nearby pixel within a distance defined by some spatial scale below which we think the model is not skilful. In other words, the forecast of a shower over a town in an 18-hour forecast should be equally as skilful as the forecast of a shower over a town say 10 km away because we don’t believe the model is accurate enough to pinpoint individual showers at specific places with that lead time. Probabilities are generated by examining the neighbourhood around each model pixel and computing the fraction of pixels that exceed some threshold we are interested in (e.g. > 40 mm of rain in 3 hours). Some additional filtering is also then done using a recursive filter. The size of the neighbourhood will be determined by the spatial scale we believe is the smallest that has useful skill for a particular model, application and forecast time. This scale can be determined by use of the Fraction skill Score verification method.

9.1 Examples of probabilities using two neighbourhood sizes

Figures 9-1 to 9-3 show probabilities that have been generated for cases B (3rd August 2004), F (4th July 2006) and case I (25th June 2007). The probabilities shown are for exceeding particular rainfall accumulation thresholds at each model pixel. They are typically small when presented in this way because it is giving the probability of something occurring at a particular grid square and even if an event is quite likely the probability that it will happen at a specific locality may still be small. E.g. even if it is very likely to rain heavily somewhere in your county it may still not be likely close to your house.

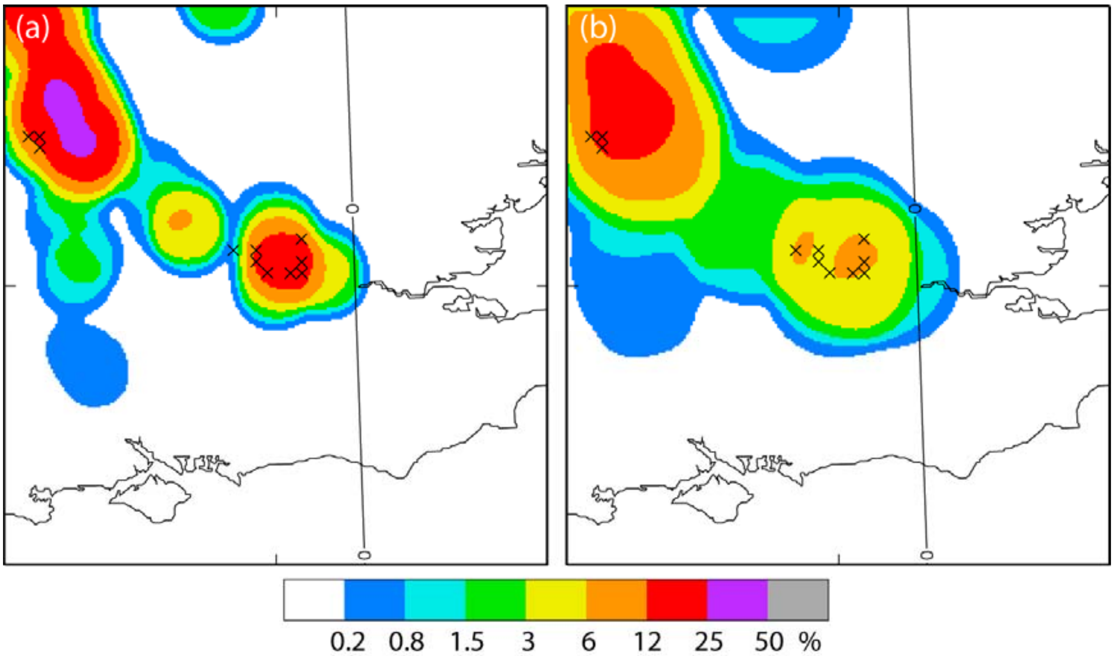


Figure 9-1. Probability that rainfall accumulations will exceed 50 mm in the period 13 to 19 UTC 3rd August 2004 from a 1 km forecast starting from 09 UTC. (a) Using a neighbourhood of 30 km, (b) Using a neighbourhood of 60 km. The crosses mark where 50 mm was exceeded from radar measurements.

The probabilities for case B (Figure 9-1) with a 30 km neighbourhood give a good indication of where the high totals were likely to occur. This reflects the fact that it was a good forecast. The probabilities with a 60 km neighbourhood are smoother, smaller and cover a greater area.

The probabilities for case F (Figure 9-2) with a 30 km neighbourhood are perhaps too focussed on where the threshold was not exceeded. This reflects the fact that it was not such a good forecast as for case B. Again, the probabilities with a 60 km neighbourhood are smoother, smaller and cover a greater area and in this instance are more appropriate.

The probabilities for case I (Figure 9-3) with a 30 km neighbourhood fit very well with what was observed. The exception is the small area to the north over the North York Moors. The 60 km neighbourhood is, as expected, smoother and broader, but still doesn’t cover all the eventualities. In this instance the 30 km neighbourhood was excellent for nearly all of the rain area, but to capture the additional bit to the north would have needed a neighbourhood bigger than 60 km.

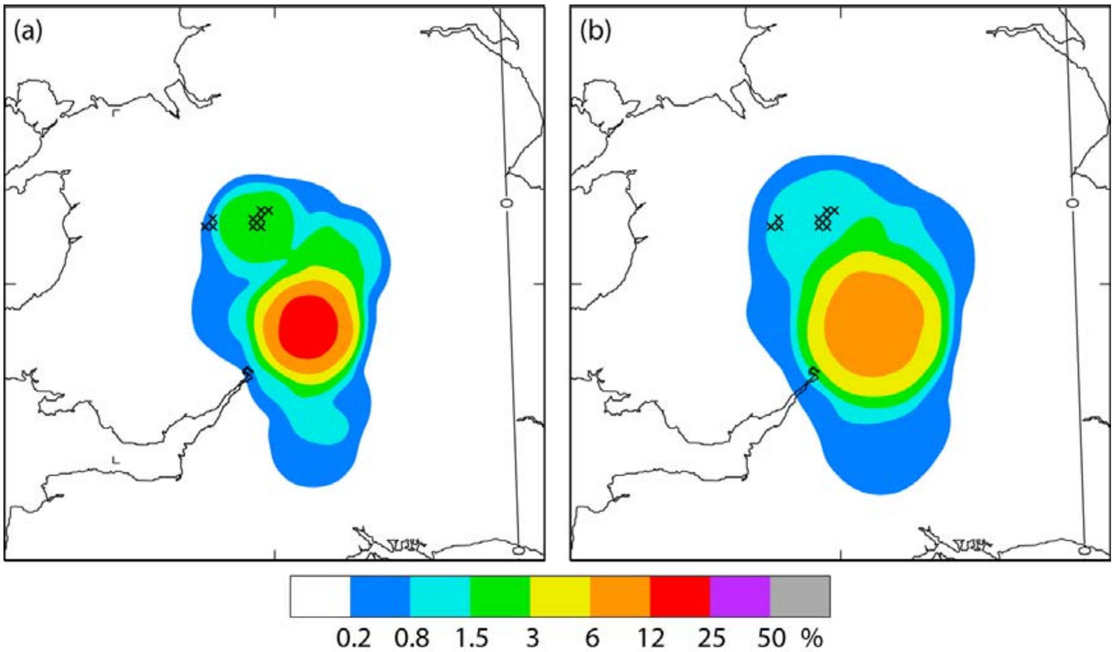


Figure 9-2. Probability that rainfall accumulations will exceed 25mm in the period 17 to 19 UTC 4th July 2006 from a 1 km forecast starting from 06 UTC. (a) Using a neighbourhood of 30 km, (b) Using a neighbourhood of 60 km. The crosses mark where 25 mm was exceeded from radar measurements.

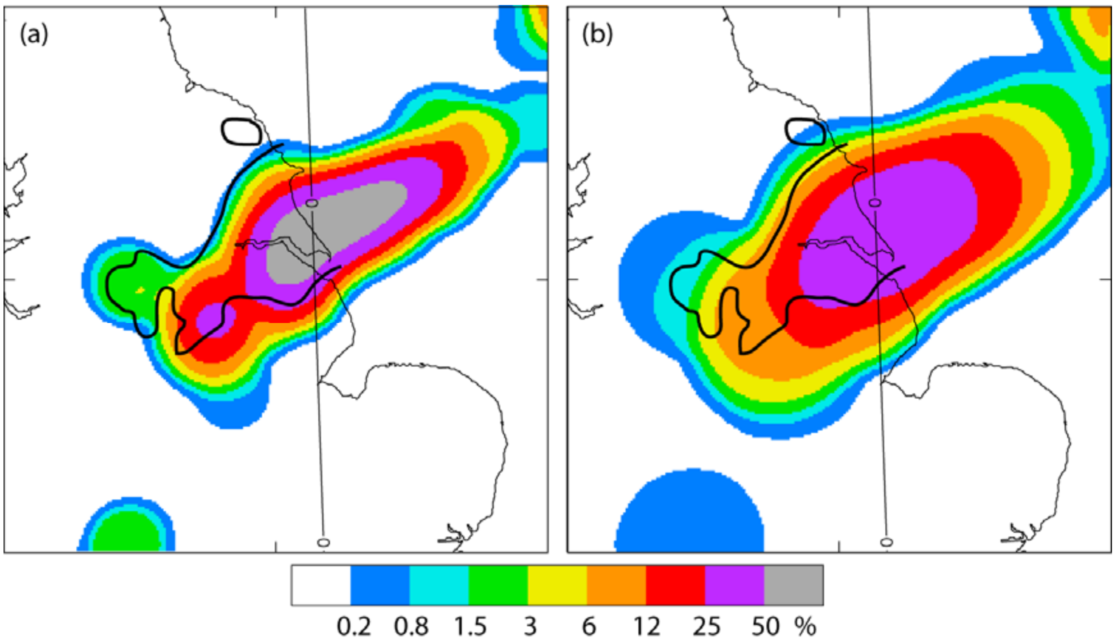


Figure 9-3. Probability that rainfall accumulations will exceed 75mm in the period 06 to 18 UTC 25th June 2007 from a 1 km forecast starting from 03 UTC. (a) Using a neighbourhood of 30 km, (b) Using a neighbourhood of 60 km. The black line encloses where rainfall exceeded 75 mm (from gauges) over the 24th and 25th.

This method of presentation shows what can easily be achieved from a high-resolution forecast (4 km as well as ~1 km). Even when rainfall is misplaced in the forecast, provided that it is not too much in error, this type of approach can give a very useful warning for the areas at risk.

9.2 Probabilities from a time-lag ensemble

Figure 9-4 shows the probabilities of rainfall accumulations exceeding 75 mm using a neighbourhood of 60 km from two different 1.5 km forecasts 3-hours apart. Both of the individual forecasts were good, but neither gave a probability distribution that was centred where the totals >75 mm were observed (using the 60 km neighbourhood). However, when they are combined the new probability picture, although broader, gives a better fit to what was observed. This combining of the latest forecast with previous forecasts is called a time-lag ensemble. Work by Mittermaier (2007) at the Met Office has shown that there is greater skill in using a time-lag ensemble with the UK 4km model than taking the latest forecast. The combination of a time-lag ensemble with this probabilistic approach could prove very useful for warnings of high rainfall amounts.

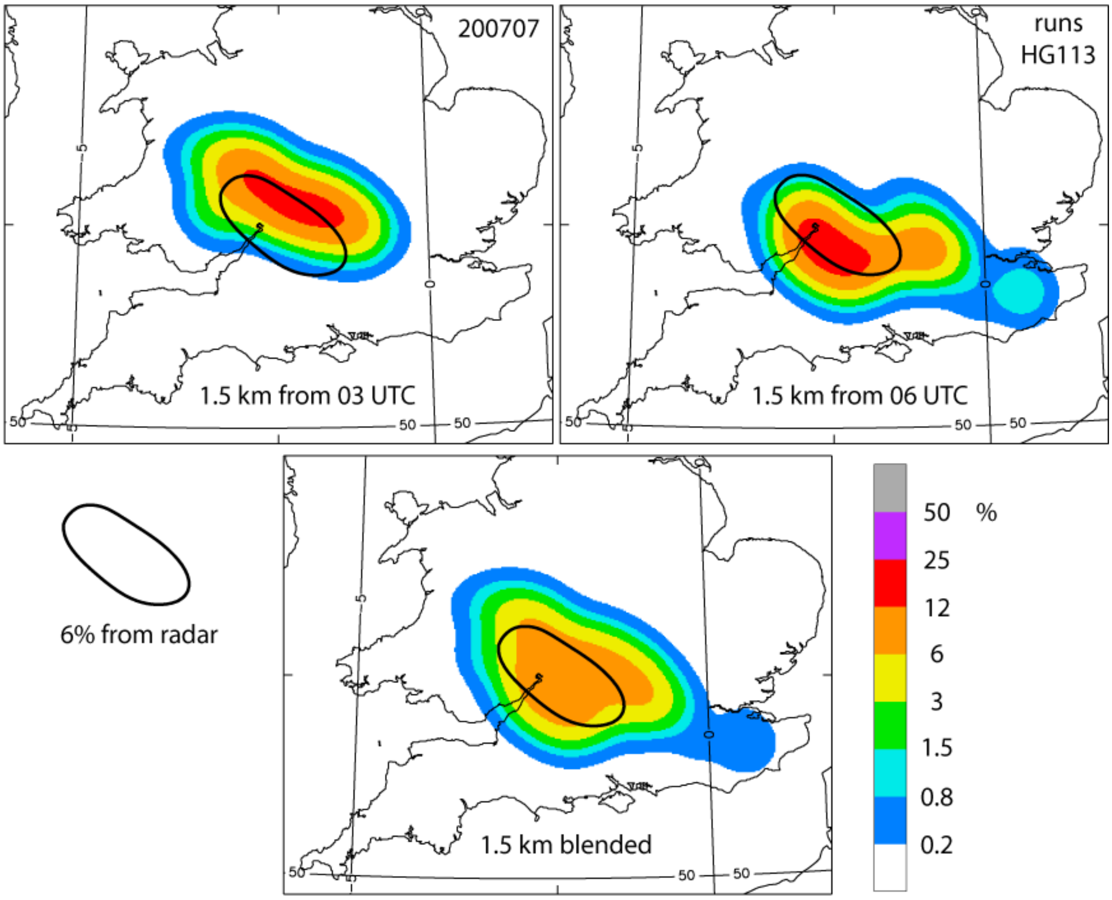


Figure 9-4 Probability that rainfall accumulations will exceed 75mm, using a 60 km neighbourhood, in the period 06 to 15 UTC 20th July 2007 from 1.5 km forecasts starting from (a) 03 UTC, (b) 06 UTC, and (c) the two forecasts combined with equal weighting. Black contour encloses the probabilities >6% when the same processing was applied to radar data.

9.3 Products for areas rather than pixels

Products can also be produced for specific areas. In Figure 9-5, Environment Agency warning areas are used, but, for example, a tiling of square areas could be used instead.

Figure 9-5(a) is designed to give an indication of the highest mean rainfall totals that could occur over a small area/catchment ($24km^2$) somewhere within each warning area given that the positioning of the rainfall could be in error by up to 30 km in any direction (same as 60 km neighbourhood). The product warns of the possibility of more than 100 mm over small catchments within a large number of the warning areas.

Figure 9-5(b) shows the probabilities that an area of $24km^2$ somewhere within each warning area will have an average accumulation of more than 50 mm. Again, it shows that there is a large swathe of warning areas over which the probabilities of this occurring exceed 25%. The probabilities are higher for this type of product than for the pixel-based probabilities shown earlier because there is more likelihood of exceeding a threshold somewhere within an area than at a specific pixel. E.g. it is easier to get any score on a dartboard than to land on the treble 20 (even when trying to!).

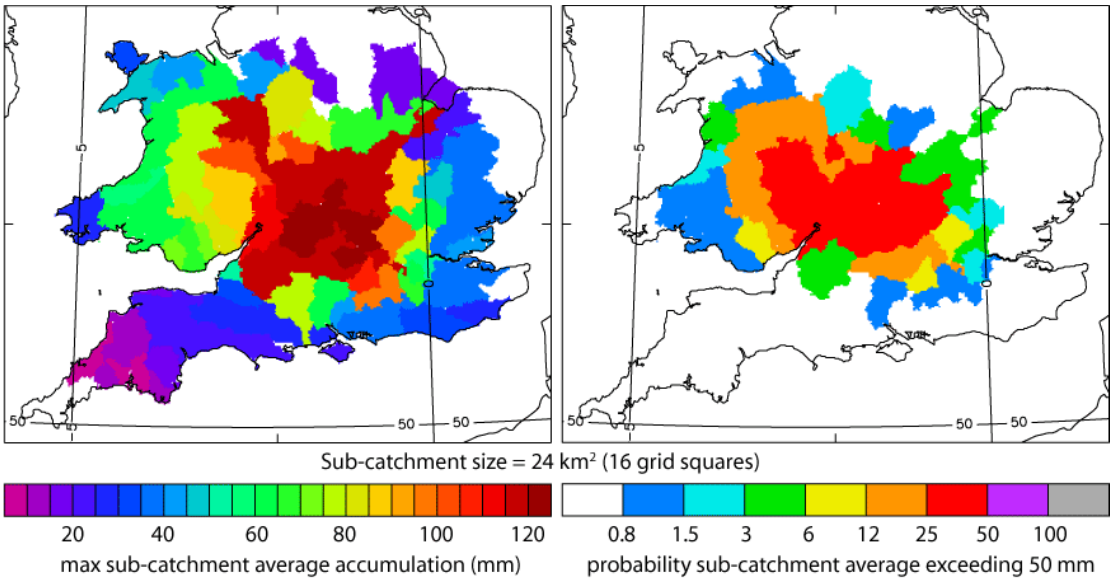


Figure 9-5. Diagnostics from the 1.5 km model forecast starting from 03 UTC 20th July 2007 for the period 06 to 15 UTC. (a) The maximum possible mean rainfall amount over a $24km^2$ area within each EA warning area given a forecast uncertainty of 60 km. (b) The probability of exceeding a mean accumulation of more than 50 mm over a $24km^2$ area within each EA warning area given a forecast uncertainty of 60 km.

10. Recent and ongoing UK and European projects aimed at understanding and predicting flood-producing rainfall

10.1 Chronological description

This section provides a summary of the relationships between recent and ongoing research projects directed at understanding and predicting flood-producing rainfall events over the UK and Europe. Central to the prediction part is the use of high-resolution Numerical Weather Prediction (NWP) models. The funding for these projects can be broadly split into four sources (a) The Public Weather Service Customer Group's research programme, (b) Department for the Environment Food and Rural Affairs (Defra) in cooperation with the Environment Agency (EA) (c) Natural Environmental Research Council (NERC) and (d) European / Global funding bodies. Context is provided by linking the research to cutting-edge developments in the operational NWP capability at the Met Office.

The story is picked up in 2001, but it should be recognised that this research is only possible because of the considerable amount of work that went in to both the fundamental understanding of meteorology and into research and development in numerical weather prediction over preceding decades both in the UK (including the Met Office) and worldwide. It should also be pointed out that parallel advances in the science of hydrology and development of hydrological models are vital for improving flood prediction, but are not the subject of this overview.

2001 - 2002

In 2001 phase 1 of the Extreme Events Recognition Project (FD2201) began. The objective was to examine the nature of extreme rainfall events including typical meteorological conditions in the run-up and to look at the susceptibility of river catchments to such events. This was timely, as it followed the serious floods that occurred in Autumn 2000. The outcome was a technical document that categorised typical meteorological situations that led to flood-producing rainfall and when they were most likely to occur, along with the main factors that lead to serious flooding. It documented a procedure for determining the susceptibility of river catchments to particular rainfall events and has provided a training dataset that could be used in the development of hydrological models. There were recommendations that new events should be routinely analysed in the same framework and that work should be done to examine how this understanding of particularly significant meteorological scenarios could be used with NWP model forecasts to provide early warning of the potential for flood-producing rainfall.

At the same time as the Extreme Events Recognition Project was going on, work to develop a new formulation of the dynamics in the Met Office Unified Model (UM) (Davies et al 2005) was reaching fruition after several years of research and development and this culminated in the release of a new operational version of the Unified Model, with the so called 'New Dynamics' on the 7th August 2002. It was this 'semi-implicit', 'semi-lagrangian' and 'non-hydrostatic' formulation of the UM that was to open up the possibility of predicting flood-producing rainfall using storm-resolving NWP models. The 'non-hydrostatic' part meant that the rapid accelerations of air in thunderstorms could be modelled and the ‘semi-lagrangian’ part meant that the calculations could be completed quickly enough to make operational forecasting a real possibility within a decade given the expected upgrades in supercomputer resources.

The potential of the new formulation of the UM was known and lead to the initiation of two new mutually beneficial projects. The High Resolution Trial Model (HRTM) project began in June 2002 at the Met Office, Joint Centre for Mesoscale Meteorology (JCMM) located in the Department of Meteorology in Reading (although work on test versions of the new UM had begun in 2001). It was funded through the Public Weather Service Customer Group's research programme. Its aims were initially to develop and test 4 km and 1 km versions of the UM (i.e. grid spacing of 4 and 1 km) over limited sized domains (centred on southern England) in order to evaluate their performance and recommend the best configuration for operational implementation. The Storm Scale Numerical Modelling project (FD2207), which was half funded by EA/ Defra was initiated in December 2001. It was set up to investigate, through a number of case studies, whether a storm-resolving NWP model could provide more accurate forecasts of flood-producing (mostly convective) rainfall events up to 12 hours ahead. It was particularly concerned with the applicability for flood warning and how the output should be presented for that purpose. To do that, a link between JCMM and the Joint Centre for Hydro-meteorological Research (JCHMR) at Wallingford was forged. (JCHMR is a collaboration between the Centre for Ecology and Hydrology (CEH) and the Met Office unit at Wallingford). The types of events were classified according to recommendations in the Extreme Events Recognition project.

2003 - 2004

The Storm Scale Numerical Modelling project finished in 2004. It concluded that a storm resolving model does indeed have the potential to deliver more accurate forecasts of high-impact rainfall events, but with the caveats that (1) appropriate post-processing of the output is essential to utilise the information on scales that are considered predictable and (2) to fulfil this potential a considerable amount of research is still needed, particularly in the area of data assimilation at high resolution and in understanding differences in predictability from day to day. The Storm Scale Model project benefited considerably from the HRTM project. The reference forecasts for each case study were based on the state-of-the-art configurations set up and maintained by the HRTM project. In turn, the Storm Scale Modelling project provided for the HRTM project a novel method for quantitatively evaluating the performance of rainfall forecasts and a modification to the way small convective cloud are represented in the 4km model.

The final report from the Storm Scale Numerical modelling project proposed three new projects (along with the need for NWP model development and data assimilation research) that would be taken up later.

1. The examine the ability of a storm-resolving NWP model to predict extreme events
2. To optimize short range precipitation forecasts by blending nowcasts (forecasts based on expected movement of rainfall for a few hours ahead) with storm-resolving NWP forecasts.
3. To incorporate forecasts from storm-resolving NWP forecasts into hydrological models for flood warning.

On the 5th December 2003 the North Atlantic European (NAE) model was introduced at the Met Office and declared fully operational on the 22nd September 2004. This model was eventually to replace the UK mesoscale model. It was to have the same horizontal grid spacing (~12km) but cover the North Atlantic and Western Europe rather than just the UK and North Sea (although to begin with it had a coarser grid spacing than ~12km). The introduction of the NAE model was essential if storm-resolving models were to eventually cover the whole of the UK (like the mesoscale model) and it can more accurately represent the smaller-scale weather patterns (e.g. fronts, organised clusters of showers or storms) that move into the UK.

Following on from phase 1, the Extreme Events Recognition Project (code FD2208, funded by Defra/EA & Met Office) moved into phase 2 in January 2004. Phase 2 focused on both the understanding of extreme rainfall events and the development and trialling of possible new ways of forecasting them and providing warnings. The areas of research were wide and varied: to develop a dataset of extreme events that can be used as an educational tool and for evaluating the extreme event performance of hydrological models, document new extreme (and less extreme events), develop and evaluate an extreme events forecasting system, evaluate a vorticity indicator for extreme events and report on the EA requirement for a decision-support scheme. It represents a step forward in the move towards a fully integrated rainfall prediction through to flood warning system.

On the 16th August 2004 the famous flood of the village of Boscastle (north Cornwall) occurred. Simulations of this case were run using the UM with grid-spacings of 4 km and 1 km as part of the HRTM and Storm Scale Numerical Modelling projects. The results showed that the finer resolution models would have been able to provide a warning of high rainfall totals on this occasion, whereas the mesoscale model (12 km) and the Nimrod nowcasting system (forecast based on expected movement of rainfall for a few hours ahead) were unable to do so. This hastened the urgency to set up an operational 4 km version of the UM for the UK.

Two other research programmes should also be mentioned here.

1. FLOODsite. This is a programme funded by the European Union and others to the tune of €14 Million, to run from 2004-2008, and involves more than 30 European universities/ research organisations. The purpose of this research is to make a difference to the way flood risks are managed within member states of the EU and more broadly. It is hoped that implementation of the research results will ultimately benefit the citizens of Europe, through a reduction in flood risk and an improvement in resilience in the face of flooding. The project is mentioned because it is an important component in European research towards managing flood risk, but it is not specifically involved with research into rainfall predictions from NWP models.
2. Flood Risk Management Research Consortium (FRMRC) - Engineering and Physical Sciences Research Council (EPSRC) funded (£5-6 million; 2004-08), also partially funded by EA/Defra and involving more than 20 British universities/research organisations. 'The aim of the FRMRC is to undertake an integrated programme of research to support effective flood risk management by
    a. establishing a programme of "cutting edge" research to enhance flood risk management practice worldwide
    b. short-term delivery of tools and techniques to support short term improvements in flood risk management in the United Kingdom; and,
    c. development and training of the next generation of flood risk management professionals through their involvement in and exposure to the consortium's research.

    It has collaboration with the Defra/EA Joint R&D Programme on Flood and Coastal Erosion Risk Management, Natural Environment Research Council (NERC) and the Scottish Executive and UK Water Industry Research. As with

FLOODsite the FRMRC is involved in using rainfall predictions, but is not sponsoring research in this area. Nevertheless it should be mentioned.

2005

There were a number of very significant advances in operational NWP at the Met Office in 2005. On the 22nd February 2005 the NAE model was upgraded to have a 12-km grid spacing (the same as the UK mesoscale model). On the 13th April 2005 a semi-operational version of the 4-km model was introduced. This was based on the configuration recommended by the HRTM project. To begin with, the 4-km model did not have its own data assimilation system and had to begin each forecast with NAE fields, but data assimilation was incorporated on the 13th December (using 3DVAR and nudging methods).

Once the 4-km model was handed over for operational development, the focus of the HRTM project turned towards a 1.5km model for future implementation and was tied in closely with advanced research into data assimilation methods for a model with that resolution. It was also in 2005 that the development of a variable resolution NWP model at JCMM was beginning to make very good progress (and would continue to do so). Variable resolution provides a very credible alternative to nesting high resolution domains inside coarser-resolution domains. A variable-resolution domain means that the grid spacing can be very small (highest resolution) in the middle of the domain where the forecast matters most and become coarser towards the outer parts of the domain. The main benefits are (1) as shown by the Storm Scale Numerical modelling project the presence of the edge in a nested high-resolution domain can have a serious detrimental impact on the representation of convective rainfall and the best way round that (without variable resolution) is to have a much larger, and therefore much more costly, high resolution domain, and (2) with a nested system the model is run twice (or more times) over the area of interest, first at coarser resolution, then at finer resolution and this adds to the computational cost and delay in forecast availability.

At the same time as the resolution of the deterministic NWP models was being improved, the Met office introduced its own ensemble prediction system. The global version of the Met Office Global and Regional Ensemble Prediction System (MOGREPS) became operational on the 14th June 2005 and the regional version, covering the same domain as the NAE but with a grid spacing of 24 km, became operational on the 17th August. The regional version has 24 members and runs for 54 hours.

A large proportion of extreme rainfall events are convective (or partly convective). (The Extreme Events Recognition project showed this.) A field campaign called the Convective Storms Initiation Project (CSIP), which was funded by the National Environmental Research Council (NERC) was set up because it was recognised that there was insufficient observational information about the conditions that lead to the initiation of convection over the UK. From an NWP model perspective, if the initiation of a storm is not captured correctly then the subsequent forecast is unlikely to be sufficiently accurate either, so it is vital to understand what mechanisms are most important and then to get a more accurate representation of the key elements at start (through data assimilation methods).

The main CSIP observational period ran for most of June, July and August 2005 (following a pilot campaign the previous July). The observational area was focussed on central southern England and centred on the Chilbolton radar facility. A large array of ground-based instruments, from the National Centre for Atmospheric Science (NCAS)

Universities' Facility for Atmospheric Measurement (UFAM), the UK Met Office and the Institute for Meteorology and Climate Research (IMK) Karlsruhe were deployed. In total there were 18 Intensive Observation Periods (IOPs). The result is a valuable dataset which has been used in several other projects.

In November the 2005 this project ‘Modelling Extreme Rainfall Events’ (MERE) (FD2210) was started. It is half funded by Defra/EA and half funded by the Public Weather Service Customer Group's research programme. The purpose of MERE is to investigate the ability of a storm scale configuration of the Met Office NWP model to predict extreme rainfall events and to determine what it is about the meteorology of these situations that the model must capture in order to produce useful predictions for flood warning. The work follows the first proposal for further research made in the Storm Scale Numerical Model project.

Forecasting extreme events had become a primary concern, so, also in 2005, the Natural Environment Research Council (NERC) research programme Flood Risk from Extreme Events (FREE) began and has £6M worth of funding. This programme ‘brings researchers in the hydrological, meteorological, terrestrial and coastal oceanography communities together in an integrated research programme for the first time.’ It aims to research what causes and propagates floods in order to help forecast and quantify flood risk. The programme has eight main objectives; the ones that are most relevant for mention here are the first three:

Objective 1. ‘To develop and extend the science underpinning integrated modelling frameworks enabling models to work sensibly and more effectively together.’ In the context of storm-resolving models, it is to work on the integration of high-resolution NWP models and hydrological models.

Objective 2. ‘To identify and spread scientific improvements in model initialisation, data assimilation and the processing of forecast ensemble outputs across modelling communities.’ For high-resolution modelling, this is about improving data assimilation methodology and observational capability.

Objective 3. ‘To understand and quantify the propagation of uncertainty within a changing environment and within rapidly changing catchments.’. For flood-producing rainfall this is do with understanding where the uncertainties lie in the NWP-model forecasts and the hydrological models and where there might be feedbacks in uncertainty particularly for small catchments.

The FREE programme will interact with FRMRC, the Joint Defra/Environment Agency Flood and Coastal Erosion Risk Management programme, the Met Office's research and development programmes, the DTI's Floods Foresight project, and the European Union's Framework 5 and Framework 6 initiatives.’ The FREE project that is relevant to high-resolution NWP precipitation forecasting is ‘Exploitation of new data sources, data assimilation and ensemble techniques for storm and flood forecasting’

Meanwhile across Europe, another collaborative research programme was also beginning. The European Cooperation in the field of Scientific and Technical research (COST) which is supported by the EU framework programme has instigated Action 731: Propagation of Uncertainty in Advanced Meteo-Hydrological Forecast Systems. ‘The main objective of the Action is to address issues associated with the quality and uncertainty of meteorological observations from remote sensing and other potentially valuable instrumentation. It will also consider their impacts on hydro-meteorological outputs from advanced forecasting systems.’ Twenty European National Met Services, including the Met Office have signed up to be a part. The funding provides for the collaborative exchange of ideas and data rather than specific research projects.

2006

On the 14th March 2006, Four Dimensional Variational data assimilation (4DVAR) was introduced operationally into the NAE model. This more sophisticated data assimilation methodology, along with the improvements in the use of satellite data it allows, brought significant improvements to the performance of the NAE model. This had the knock on effect of improving the UK 4km model which uses information from the NAE at its boundaries. The improved performance of the NAE meant that the UK mesoscale model could be retired on the 31st October 2006.

In July 2006 a new project was initiated within the Joint Defra/EA Flood and coastal erosion risk management R&D Programme. This was the Probabilistic Flood Forecasting Scoping Study (FD2901). It was set up to document sources of uncertainty in the flood forecasting process in recognition of the fact that forecasts are not certain and should include some indication of uncertainty, which may involve the use of probabilities. It examined current approaches within the Environment Agency to assess uncertainty in flood forecasts, international research on ensemble flood forecasting techniques, possible applications of decision support systems, and risk based forecasting techniques in other (non-water) sectors. The contractor was Atkins Environment. This project was set up at a time when it was becoming clearer that a more probabilistic approach was necessary in both the meteorological and hydrological components of a flood warning system. As far as rainfall predictions were concerned, the Met Office had developed the capability to produce short-range ensemble rainfall predictions through STEPS and 2-day 24km NWP ensembles from MOGREPS. Methods for producing probabilistic output from storm-resolving NWP forecasts had been presented in the Storm Scale Numerical Model project.

At around the same time the Use of Probabilities project was set up to develop experimental probabilistic output for trial use. Phase 1 of the Use of probability forecasts project (May 2006 to December 2007) established an initial user requirement for probabilistic precipitation forecasts in the Environment Agency. A questionnaire aimed at identifying the learning needs and product requirements of Agency flood forecasters was distributed to participants in September 2006. The feedback received was presented at the Phase 1 workshop, held in November 2006. In phase 2 (2007/2008) a basic web visualisation capability was developed within the Met Office to support a trial of selected probabilistic products in the Environment Agency based on output from MOGREPS, the UK 4 km model and STEPS. Operational implementation is planned for 2008.

At the end of October the Met Office started to trial an ‘on-demand’ 1.5km model. Nine regional 300x300km domains were set up to patchwork over the whole of the UK. The 1.5km model could be run on one of these domains once each day (starting from 03 UTC) for 18 hours. The decision on whether the model should be run and on what domain would be made by the forecaster in charge on the day. In December the on-demand system was made quasi-operational.

2007

This was the year in which there were two very serious flood events that prompted a Cabinet Office Review; one on the 25th June over parts of northern England (particularly Sheffield and Hull) and one on 20th July over southern England and the Midlands (worst affected was the Severn Valley). In addition, there was another case of flooding at Boscastle (North Cornwall), although much less serious this time (and a near miss for Boscastle), a flash flood with large hail in parts of London and a flash flood at Filey (North Yorkshire) and flash floods in the Midlands and northern England

(15^{th} June). These events were a reminder of the need to take seriously the impact of flooding and the need to continue research into improving forecasts and warnings.

Two new projects were commissioned within the Joint Defra/EA Flood and coastal erosion risk management R&D Programme Incident Management and Community Engagement (IMC) Theme. They both follow on from the Probabilistic Flood Forecasting Scoping Study and are labelled FDK(07)01 and FDK(07)06.

The project ‘Probabilistic Fluvial Forecast Modelling (FDK(07)01)’ was set up to evaluate the impact of adopting probabilistic flood forecasting operationally for hydrological and hydraulic fluvial flood forecasting models, and to recommend practical ways of reducing model runtime and an operational statistical framework for processing multiple model runs with probabilistic inputs and interpretation of the hydrological model components. It will use ensemble rainfall forecasts derived from single high-resolution NWP forecasts and output from MOGREPS for selected cases to provide the probabilistic rainfall component. It is realised that generating probabilities from a deterministic forecast has limitations, but it will still provide a useful test-bed for examining how hydrological models can deal with a realistic estimate of precipitation forecast uncertainty.

The project ‘Blending convective scale NWP with ensemble nowcasting (FDK(07)06’ was set up to explore the possibility of generating high resolution ensemble precipitation forecasts with a maximum range of two days by blending stochastic precipitation nowcasts (e.g. from STEPS) with high resolution, NWP-model forecasts and ensemble forecasts of precipitation produced by MOGREPS. This project is based on the recognition that there is a need for probabilistic rainfall predictions, but with an understanding of the limitations of current forecast systems. At present MOGREPS can provide ensemble output but does not have the necessary resolution to represent convective storms; a storm-resolving model can represent the storms but is too computationally expensive to run as an ensemble, and nowcasting systems such as STEPS can be run as an ensemble but are only useful for a limited forecast period. A pragmatic approach is to combine these three different types of forecast in a way that will produce a seamless probabilistic product.

Also commissioned by the Environment Agency under the Making Space for Water programme was the ‘Feasibility study into expanding flood warning to cover other flood risks such as groundwater’ (RF5). This ran from February to October 2007. Work undertaken in the Met Office showed that, through the use of ‘storm-resolving’ NWP models, it will be possible in future to provide warnings of pluvial as well as fluvial flooding events.

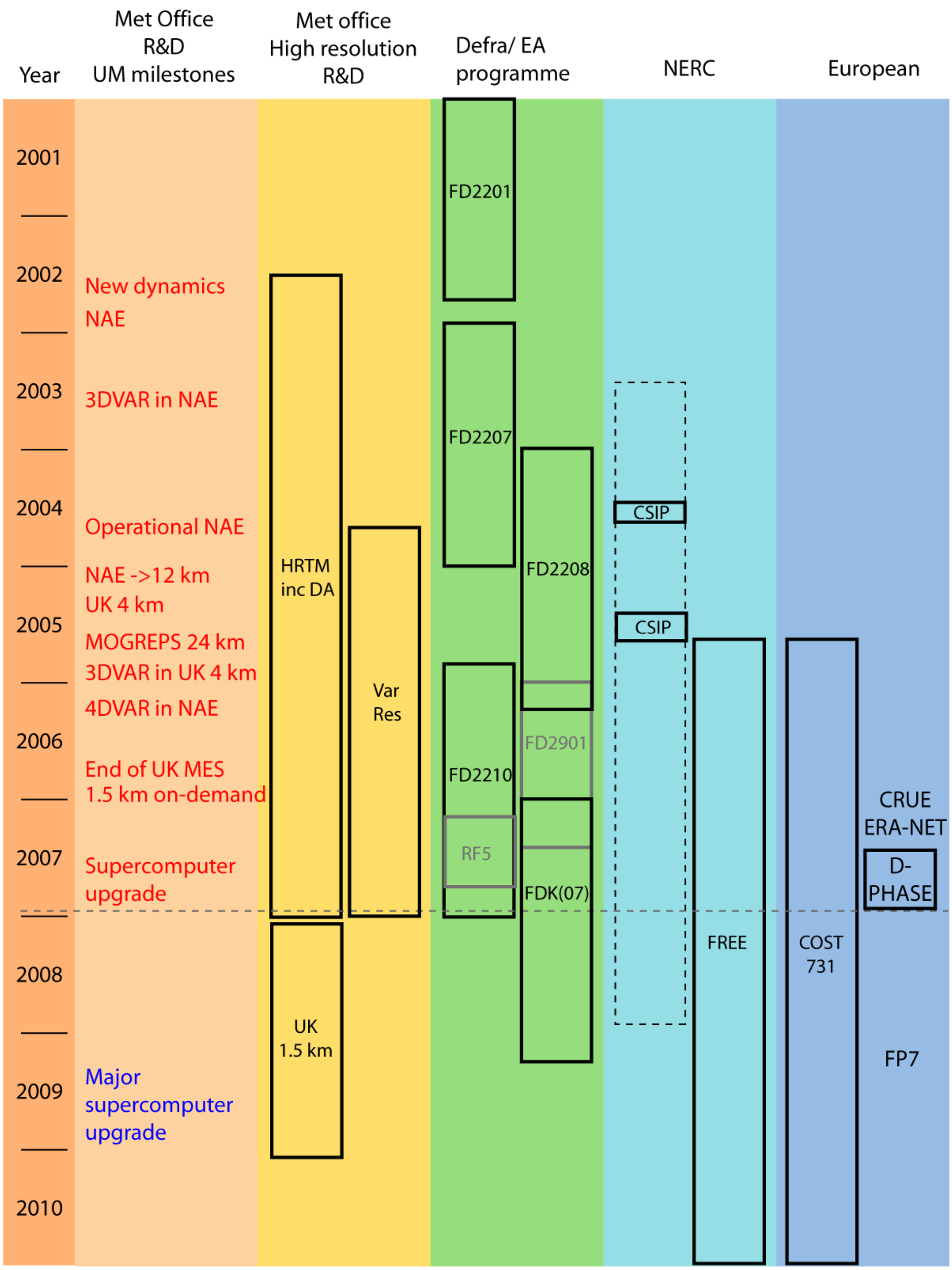
Meanwhile in Europe D-PHASE was underway between June and November 2007. D-PHASE stands for the Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events in the Alpine region and is a follow up to the Mesoscale Alpine Programme (MAP). It was aimed ‘at demonstrating some of the many achievements of MAP, in particular the ability of forecasting heavy precipitation and related flooding events in the Alpine region’. This involved setting up an end-to-end web-based forecasting system that conveyed forecast uncertainty by incorporating a large number of NWP models and hydrological models. MOGREPS was included as one of the NWP model ensembles. The system has been successful to the extent that users asked for it to remain switched on beyond November.

Beyond 2007

A major upgrade in supercomputer capability is expected in the Met Office in 2009/2010 and this will provide the capability to run an operational UK ~1.5 km model. The new model will be developed on a variable resolution grid that is anticipated to range from a grid spacing of ~12km at the corners to ~1.5km in the middle over the UK. Research will continue to develop data assimilation methods for high-resolution models and include techniques for making use of new types of observations such as Doppler radar and refractivity measurements. The data assimilation work will combine Met Office research and development with work carried out within the FREE programme and at the Data Assimilation Research Centre (DARC) at the University of Reading as well as being in contact with COST action 731 and projects arising out of the European Framework 7.

There is now a strong move towards an 'integrated' and probabilistic' approach to heavy rainfall and flood prediction both in the UK and worldwide. The projects 'Probabilistic Fluvial Forecast Modelling' (FDK(07)01) and 'Blending convective scale NWP with ensemble nowcasting' (FDK(07)06 point to the move in that direction in the UK and Netherlands. D-PHASE is a broader European example. This is expected to continue with further work required in the move towards ensemble rainfall predictions from high-resolution NWP models, probabilistic hydrological models and methods for linking the two and presenting the information.

10.2 Projects Diagram



10.2.1 Glossary of terms in projects diagram

Unified Model

- 1.5 km ‘on-demand’ – A version of the Met Office Unified Model that has a grid spacing of 1.5 km on a possible 9 different domains (each 300x300 km) that can be run on the most appropriate domain for the meteorological conditions once per day.
- 3DVAR – 3-dimensional Variational data assimilation. This is the primary methodology used to update the start of the UK 4km model forecasts with new observational information to give each forecast the most accurate start.
- 4DVAR – This is a more advanced version of 3DVAR in which both the spatial and temporal fit between the model and observations is used to create the most accurate starting point for each forecast. It is used in the global and NAE models.
- MOGREPS – The Met Office Global and Regional Ensemble Prediction System. The Met Office runs a global model ensemble and a regional model ensemble (NAE at 24 km) every day. Each ensemble has 24 members.
- New Dynamics – The new dynamical formulation of the Met Office Unified Model that allows the possibility of running on very high resolution grids and resolving individual storms.
- NAE – The North Atlantic European version of the Met Office Unified Model. The domain covers the north Atlantic and much of Europe with a current grid spacing of 12 km. Forecasts are run for 48 hours from 00, 06, 12 and 18 UTC each day.
- UK 4km – The version of the Met Office Unified Model that has a grid spacing of 4 km and a domain covering the UK. It is run four times a day at 03, 09, 15 and 21 UTC.

Projects

- HRTM – High Resolution Trial Model project.
- UK 1.5 km – UK 1.5 km model project.
- Var Res – The Variable Resolution model project.
- FD2201 – The Extreme Event Recognition project – phase 1.
- FD2207 – The Storm Scale Numerical Modelling project.
- FD2208 – The Extreme Event Recognition project – phase 2.
- FD2210 – The project Modelling Extreme Rainfall Events.
- FD2901 - Probabilistic Flood Forecasting Scoping Study.
- RF5 - Feasibility study into expanding flood warning to cover other flood risks such as groundwater
- FDK(07)01 - Probabilistic Fluvial Forecast Modelling project
- FDK(07)06 - Blending convective scale NWP with ensemble nowcasting project.
- CSIP – Convective Storms Initiation Project.
- FREE – Flood Risk from Extreme Events programme.
- COST-731 – European Cooperation in the field of Scientific and Technical research. Action 731: Propagation of Uncertainty in Advanced Meteo-Hydrological Forecast Systems.
- CRUE ERA-NET - European Flood Research co-ordination programmes.
- D-PHASE - Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events in the Alpine region.
- FP7 – European Framework Programme 7.

10.3 Tabulated description of projects

Extreme Events Recognition Project phase 1 FD2201 2001 – August 2002	
Description	'The aim of the research was to investigate the nature of very extreme rainfall events and the meteorological situations leading to their occurrence, and also the susceptibility of river catchments to their spatial and temporal rainfall patterns.'
Funded Contractors	Met Office and Salford University.
Funding	Commissioned within the Flood Forecasting and Warning theme of the joint Defra/Environment Agency Research Programme.
Status	Completed August 2002
Reports	Extreme Rainfall and Flood Event Recognition. R&D technical report: 2201
Key Findings	The project came up with classifications of the different meteorological situations that can lead to extreme rainfall events. The following recommendations were made: 'New events should be routinely analysed and tested to see how they fit into the archetypal situation proposed;(2) The Met Office Mesoscale Model can be used to provide details of the synoptic evolution, expected rainfall intensity, accumulation and distribution of rainfall as related to the archetypal situation;(3) A joint Defra/Met Office/EA Project was proposed to establish a prototype 24-hour early warning system;(4) Recent work at the University of Salford on the use of Doppler radar and NWP data to identify extreme convective events should be considered in the context of recommendation (3);(5) A decision support methodology based upon a question and answer scoring scheme seems to offer useful guidance on assessing the susceptibility of river catchments to flooding arising from extreme rainfall, but further assessment and testing are necessary;(6) The rainfall training data sets offer a means of testing operational procedures and models under extreme conditions.'
Web link	Final reports at Defra link http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=ProjectList&Completed=0&FOSID=12 Search for FD2201

Storm Scale Numerical Modelling project FD2207 December 2001 – December 2004	
Description	The aim of this research was to investigate whether the use of a very fine resolution NWP model can provide more accurate predictions of rainfall amounts up to 12 hours ahead, with the emphasis on convective events.
Funded Contractors	Met Office
Funding	Commissioned within the Flood Forecasting and Warning theme of the joint Defra/Environment Agency Research Programme for half the funding the other half came the Public Weather Service Customer Group's research programme.
Status	Completed. A follow on project 'Modelling Extreme Rainfall Events' was initiated.
Reports	Stage 1 report 'Results from high-resolution modelling of convective events', 2003 Stage 2 report 'The impact of a change to the use of the convection scheme to high-resolution simulations of convective events', 2003 Stage 3 report 'Precipitation diagnostics for a high-resolution forecasting system', 2003 Stage 4 report 'Measuring the fit of rainfall analyses and forecasts to radar', 2004 Stage 5 report 'Verification of the fit of rainfall analyses and forecasts to radar', 2004 Stage 6 report 'Final scientific report', 2004 Stage 7 report 'Final project report', 2004
Key Findings	The project provided evidence that a storm scale (~1km grid spacing) configuration of the Unified Model does indeed have the potential to deliver a significant improvement in our ability to predict high-impact convective rainfall events. It was recommended that the development of a storm scale model should continue and move towards operational implementation in conjunction with a post-processing system for the delivery of precipitation products (mostly probabilistic) for use in flood prediction. A considerable amount of research is still needed to meet this potential especially in the area of data assimilation. New projects proposed 1. To assess and optimize the ability of a storm-resolving model to predict extreme rainfall events. 2. Blending of ensemble nowcasting with high-resolution NWP forecasts 3. Incorporate high-resolution NWP forecasts into hydrological models
Web link	Reports are available from the Met Office web page http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/ Also from Defra (search for project FD2207) http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=ProjectList&Completed=0&FOSID=12

High Resolution Trial Model (HRTM) project June 2002 – December 2007	
Description	Following the development of the non-hydrostatic dynamics in the Unified Model which makes operational high-resolution modelling (grid spacing <5km) a feasible possibility, the HRTM project was initiated to determine the optimal configuration for trial models at 4km and 1km, to gain experience at running at these resolutions, assess the ability of these models and make recommendations for future implementation.
Funded Contractors	Met Office
Funding	Public Weather Service Customer Group's research programme
Status	Nearly Completed. New project to be initiated 'UK 1.5km implementation project'
Reports	The Summer 2003 Reruns with the High Resolution Trial Model by H Lean et al The Summer 2004 Reruns with the High Resolution Trial Model by H Lean, S Ballard, P Clark, M Dixon, Z Li and N Roberts
Key Findings	The HRTM project has developed suites to run 4km and 1km (and 1.5km) models in various configurations. Trials of convective events over several summers have shown the statistical benefit of 1/1.5km. This project has formed a platform for and benefited from the Storm Scale Modelling project and the Modelling Extreme Rainfall Events project. There have also been strong interactions with the Convective Storms Initiation Project (CSIP). An operational 4km model has been implemented based on the HRTM configuration and results. More recently, more advanced data assimilation (DA) has been introduced (alongside DA advances) and a variable resolution grid (from the variable resolution code development project). The quasi-operational 'on-demand' 1.5km model has been implemented in accordance with recommendations from this project.
Web link	Reports are available from the Met Office web page http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/

Extreme Events Recognition Project phase 2 FD2208 January 2004 – January 2006	
Description	Following on from phase 1 which focused on achieving a better understanding of Extreme events and their characteristics. Phase 2 focused on the understanding of extreme events but also developing and trialling possible new ways of forecasting them. The ultimate aim of the work was to enable a better forecasting service for these types of events.
Funded Contractors	Met Office, Salford University and the Centre for Ecology and Hydrology (CEH).
Funding	Commissioned within the Flood Forecasting and Warning theme of the joint Defra/Environment Agency Research Programme.
Status	Completed: January 2006
Reports	Report on analysis of less-extreme events and recent extreme event Report on evaluation of extreme-event forecast system Evaluation of a vorticity indicator for extreme events Spatio-temporal rainfall datasets and their use in evaluating extreme event performance of hydrological models The Extremes Dataset Report on EA requirement for decision-support scheme Final report
Key Findings	The work was carried out in 5 work packages: • Extending the historical analysis of extreme events in Phase 1 to more recent and less extreme events • Developing and evaluating an extreme event prediction system as a service to forecasters. • Evaluating an indicator for extreme convective (thunderstorm) events based on vorticity. • Develop rainfall datasets from historical heavy rainfall events, enhanced to represent extreme events, and use them to evaluate the extreme event performance of flood forecasting models. The work indicated that improvements to NWP models and observing systems over the next five years should result in better resolution and forecasting of the type of situations that result in extreme events (e.g. including using model grids finer then the current $4km^2$) 2. Extreme events are not in a distinct distribution from less extreme ones. 3. Environment Agency flood forecasting and warning systems can now be tested to see how they perform under extreme rainfall event conditions and flood forecasters trained using the datasets produced from this work. Both of these elements should result in improvements in recognition of extreme events and mitigating their effects by provision of better warnings. See final report via web link below for more detailed conclusions.
Web link	http://www.metoffice.gov.uk/extremeevents/

Convective Storms Initiation Project (CSIP) Main field campaign 6th June -25th August 2005	
Description	The goal of CSIP is to understand the mechanisms responsible for the initiation of precipitating convection in southern England. The very-high-resolution mesoscale model under continuing development in the Met Office has had some success in predicting the location and intensity of thunderstorms a few hours ahead, although this remains a difficult challenge. One of the largest uncertainties is in modelling the initiation of the convection. This is a joint project between UK universities, the Met Office and the Institute for Meteorology and Climate Research, Karlsruhe, Germany. A major aim of CSIP is to compare the results of the Met Office model with detailed observations of the early stages of convective clouds and to use newly gained understanding to improve the predictions of the model.
Funded Contractors	Aberystwyth, Bath, Leeds, Manchester, Reading and Salford Universities
Funding	Natural Environment Research Council (NERC) through the joint universities National Centre for Atmospheric Science (NCAS) and the Universities' Facility for Atmospheric Measurement (UFAM) for the universities. Public Weather Service Customer Group's research programme for the Met Office work.
Status	Field campaign completed: Research ongoing.
Reports Papers	'Summary of the Convective Storm Initiation Project Intensive Observation Periods', K Browning et al, 2006 Met Office Tech report 474. 'The Convective Storms Initiation Project', 2007/2008 K Browning et al. Accepted for the Bulletin of the American Meteorological Society (BAMS) 'Combination of mechanisms for triggering an isolated thunderstorm: observational case study of CSIP IOP 1', 2007/2008 C Morcrette et al. Accepted for publication in Monthly Weather Review.
Key Findings	The field campaign was very successful and as a result there is a large volume of very useful observational data. There were 18 Intensive Observation Periods (IOPs) during the main field campaign. Of these, seven have been designated as having most meteorological interest in conjunction with good data coverage. Many of the others are also very useful datasets. Research since the field campaign has lead to papers in peer reviewed literature and insights into NWP model performance. The CSIP IOPs are being used in the NERC funded project Flood Risk from Extreme Events (FREE).
Web links	http://www.see.leeds.ac.uk/research/ias/clouds/current/csip/ http://www.cas.manchester.ac.uk/research/projects/csip/ Met Office technical reports available from http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/

Flood Risk from Extreme Events (FREE) research programme 2005 - 2010	
Description	'FREE is designed to deliver fundamental environmental science advances compatible with, and supporting, the developing applied-research programme Flood Risk Management Research Consortium (FRMRC) led by EPSRC, with Defra, the Environment Agency, UKWIR and NERC. Many areas in FREE are complementary with themes in the FRMRC and close interaction and collaboration will occur. Also alignment will be maintained with the Joint Defra/EA Flood and Coastal Erosion Risk Management, the Met Office R&D programmes, the DTI Floods Foresight project, and EU FP5 and FP6 initiatives. A central feature of FREE is to bridge the interfaces between the various water environments and users of flood forecasts and to properly quantify uncertainty in conditions where non-linearity is often dominant. The programme will focus on precipitation, river catchments and coastal flooding processes within two time domains. These time domains are firstly flood forecasting and warning over time-ranges from minutes to weeks and secondly quantification of long-term flood risk over periods extending from seasons to multi-decades. FREE will bring together, for the first time in an integrated research programme, researchers in the hydrological, meteorological, terrestrial and coastal oceanography communities. A central feature is to bridge the interfaces between the various water environments and users of flood forecasts and to properly quantify uncertainty in conditions where non-linearity is often dominant.'
Funded Contractors	Various UK universities, Centre for Ecology and Hydrology (CEH)
Funding	Natural Environmental Research Council (NERC) The programme owner is NERC.
Status	Ongoing
Reports	Science and implementation plans (see web link below)
Key findings	-
Web links	http://www.nerc.ac.uk/research/programmes/free/

European COoperation in the field of Scientific and Technical research (COST) Action 731: Propagation of Uncertainty in Advanced Meteo-Hydrological Forecast Systems 2005 - 2010	
Description	'The main objective of the Action is to address issues associated with the quality and uncertainty of meteorological observations from remote sensing and other potentially valuable instrumentation. It will also consider their impacts on hydro-meteorological outputs from advanced forecasting systems.' There are 3 working groups. 1. WG-1 Propagation of uncertainty from observing systems (radars) into NWP; 2. WG-2 Propagation of uncertainty from observing systems and NWP into hydrological models; 3. WG-3 Use of uncertainty in warnings and decision making.
Funded Contractors	20 European Met Services (including the Met Office). Funding is for networking only – not the work
Funding	COST
Status	Ongoing
Reports	Progress report available via web link (see below)
Key findings	-
Web links	http://www.cost.esf.org/index.php?id=205&action_number=731 http://cost731.bafg.de/servlet/is/Entry.9691.Display/

Probabilistic Flood Forecasting Scoping Study FD2901 Jul 2006 – May 2007	
Description	To report on 'sources of uncertainty in the flood forecasting process, current approaches within the Environment Agency to assessing uncertainty in flood forecasts, international research on ensemble flood forecasting techniques, possible applications of decision support systems, and risk based forecasting techniques in other (non-water) sectors. The outputs from this project will help to inform Defra and the Environment Agency in developing a plan for bringing this important development into operational use over the next few years.'
Funded Contractors	Atkins Environment
Funding	Joint Defra/EA Flood and coastal erosion risk management R&D Programme
Status	Completed
Reports	Technical report – see web link
Key findings	See technical report
Web links	Information, including documents, can be found at http://www.defra.gov.uk/environ/fcd/research/RandDProgCon/imc.htm

Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events in the Alpine region (D-PHASE) June – November 2007	
Description	'The main objective of MAP D-PHASE, the MAP Forecast Demonstration Project (MAP FDP), is to demonstrate the benefits in forecasting heavy precipitation and related (flash) flood events, as gained from the improved understanding, refined atmospheric and hydrological modelling, and advanced technological abilities acquired through research work during the Mesoscale Alpine Programme (MAP). Specifically, an end-to-end forecasting system for Alpine flood events will be set up to demonstrate state-of-the-art forecasting of precipitation- related high-impact weather. This system will include probabilistic forecasting based on atmospheric and hydrological ensemble prediction systems with a lead time of a few days, followed by short-range forecasts based on high-resolution deterministic atmospheric and hydrological models for selected regions or catchments, and will be completed with real-time nowcasting tools. Throughout the forecasting chain, warnings will be issued and re-evaluated as the potential flooding event approaches, allowing forecasters and end users to alert and make decisions in due time.' It uses ensemble and deterministic NWP models from several countries and several hydrological models.
Funded Contractors	The list of contributers is long and I don't know how each is funded.. The Met Office is involved by contributing forecasts form the Met Office Global and Regional Ensemble System (MOGREPS.
Funding	Forecast Demonstration Project (FDP) of the WWRP (World Weather Research Programme of WMO) – although I believe the visualisation platform wasn't directly funded!
Status	Ending, although may be extended.
Reports	See web site
Key findings	See web site
Web links	Excellent web link including access to the visualisation platform and the Mesoscale Alpine Project (MAP): http://www.map.meteoswiss.ch/map-doc/dphase/dphase_info.htm

Use of probabilities May 2006 –2008	
Description	To develop experimental probabilistic output for trial use towards operational implementation. Phase 1: Establish an initial user requirement for probabilistic precipitation forecasts in the Environment Agency. Phase 2 (2007/2008) develop a basic web visualisation capability within the Met Office to support a trial of selected probabilistic products in the Environment Agency. Operational implementation is planned for 2008.
Funded Contractors	Met Office
Funding	Environment Agency
Status	Ongoing
Reports	
Key findings	
Web links	Trial page http://www.metoffice.gov.uk/weather/uk/probability/

Feasibility study into expanding flood warning to cover other flood risks such as groundwater RF5 February – October 2007	
Description	'This project will provide clear recommendations on the feasibility of providing a more comprehensive flood warning service. The objective of the project is to produce a report and recommendations covering the following key objectives: a) to recommend any science required to support an expanded service; b) to study the feasibility of expanding the flood forecasting and warning service to cover the whole range of flood risk; c) to recommend mechanisms for integrating provision of warnings for these sources of flooding into the existing flood warning system; d) to identify the timescales, and resources required to develop and operate an expanded service; and e) to recommend any legislative or policy changes required.' Looking at the feasibility of giving warnings for pluvial as well as fluvial flooding (drains, sewers, flash floods as well as river flooding) given new developments in rainfall prediction (e.g. storm-resolving NWP models)
Funded Contractors	Jacobs, Met Office
Funding	Defra under the Making Space for Water programme
Status	-
Reports	-
Key findings	-
Web links	Search for 'RF5' in Defra home page http://www.defra.gov.uk

Probabalistic Fluvial Forecast Modelling FDK(07)01 2007 – 2008	
Description	To evaluate the impact of adopting probabilistic flood forecasting operationally for hydrological and hydraulic fluvial flood forecasting models, and to recommend practical ways of reducing model runtime and an operational statistical framework for processing multiple model runs with probabilistic inputs and interpretation of the hydrological model components.
Funded Contractors	Delft Hydraulics (Netherlands), Centre for Ecology and Hydrology (CEH)
Funding	Commissioned within the Joint Defra/EA Flood and coastal erosion risk management R&D Programme. Incident Management and Community Engagement (IMC) Theme.
Status	Ongoing
Reports	-
Key findings	-
Web links	http://www.defra.gov.uk/environ/fcd/research/RandDProgCon/imc.htm

Blending convective scale NWP with ensemble nowcasting FDK(07)06 July 2007 – March 2009	
Description	To explore the possibility of generating high resolution ensemble precipitation forecasts with a maximum range of two days by blending stochastic precipitation nowcasts, generated using techniques similar to those embodied in the Short Term Ensemble Prediction System (STEPS) and Spectral Prognosis models, with high resolution, deterministic Unified Model forecasts and ensemble forecasts of precipitation produced by MOGREPS (Met Office Global and Regional Ensemble Prediction System).
Funded Contractors	Met Office
Funding	Commissioned within the Joint Defra/EA Flood and coastal erosion risk management R&D Programme. Incident Management and Community Engagement (IMC) Theme and half funded by the Public Weather Service Customer Group's research programme
Status	Ongoing
Reports	-
Key findings	-
Web links	http://www.defra.gov.uk/environ/fcd/research/RandDProgCon/imc.htm

Note that at the time of writing the web links worked, but there is not guarantee they will in future.

11. Discussion

11.1 Extreme events

11.1.1 Meeting the criterion

The extreme events studied in this project should have had a 1:100 year return period according to the Flood Estimation Handbook (FEH) method. Four of the five cases examined in stage 1 met that criterion. The only one that was less clear was the High Wycombe / London case on the 3rd August 2004. The return period for High Wycombe was 1:64 years, but according to radar considerably more rain fell in parts of northwest London where there was serious flooding and so it was very likely that it was a 1:100 year event in that area.

Of the other cases, the third to be examined in detail (Case F – the Albrighton flood) was very definitely an extreme event, as were the two major events in summer 2007 and the storm on the 13th August 2006 (case H). The other case (G) that was briefly discussed was probably not a 1:100 year event, but nevertheless caused flash flooding and disruption.

11.1.2 Classification of extreme events

In the extreme event recognition project (Hand et al 2004), extreme events over the UK were split into three types; orographic, frontal and convective. From the 50 cases they examined, 30 were classified as predominantly convective, 15 frontal (of which 9 had embedded convection) and 5 orographic, although some were a combination. They also found that most extreme events occurred in the summer months, which is not surprising given the higher moisture content of warm air. Most of the summer events are convective in nature because the strong heating over land and because strong frontal systems (associated with deep cyclones) do not occur so much in summer. It is therefore very important to be able to predict the convective events if we are going to predict most extreme events. That is why the capability of a storm-resolving NWP model is being examined.

11.1.3 What makes an extreme event

None of the extreme events examined in this project exhibited any meteorological behaviour that could not be represented in an NWP model. The dynamics that lead to the development of storms and the physical process within the storms can be simulated (with differing levels of accuracy). The extreme nature of the events was a result of the coincidence and interaction of a few key meteorological factors (e.g. hot-humid air, large instability, local ascent), each of which may not have been particularly unusual in itself, but the combination was rare for any single location.

This same occurs when throwing dice. There is a 1 in 6 chance of throwing a 6 with a single die, but to throw five sixes with five dice is a 1 in 7776 chance. This is despite there still only being a few dice and without breaking the rules – there has been no tampering with the dice. With the dice example the chance of getting a 6 is the same for all dice. In the meteorological context, the key factors are not equally likely; some are rarer than others. For example, a particularly warm and humid air-mass with a large amount of potential instability is quite unusual, and a low-level convergence line is common, but both are necessary for the thunderstorm. It's a bit like an accumulator bet where the pay-out requires all the results to be correct. They all have different odds (some long, some short), but the chance of them all being correct is small and the pay-out is large. Whether an NWP model can predict an extreme event depends on whether it can predict the individual meteorological factors and bring them together in the correct location with any interactions and feedbacks.

11.2 Is a storm-resolving model capable of predicting an extreme event

The case studies in stage 1 showed that it is possible for a storm-resolving NWP model to predict an extreme rainfall event. This does not mean that all the forecasts were extremely accurate. Some were better than others, but it was possible to generate the high rainfall totals and the forecasts were realistic (i.e. produced the correct type of thunderstorms). This view was reinforced by the results in stages 2 and 3 and the additional cases.

This is an important result. It means that a storm-resolving model has the potential to greatly improve on our current ability to predict such events. Especially given that the results have also shown that a 12 km model is not capable of an accurate representation of the extreme events that were convective in nature and the ~1 km forecast were generally (though not always) better than the 4 km forecasts. It will be important to make sure that the output from such a model includes information about forecast uncertainty (e.g. probabilities), and this will be discussed later.

11.3 Meteorological components – how to get an accurate forecast

In three cases studied in detail in stages 2 and 3, it became evident that the accuracy of the forecasts depended on getting all the relevant meteorological factors (or components) sufficiently correct. These varied from case to case and are listed below.

Case B 3rd July 2004

1. Larger-scale rotation in the upper-level flow
2. Very warm air at low levels
3. Considerable potential instability.
4. A stable layer (or ‘lid’).
5. Low-level convergence caused associated with the upper-level rotation.
6. Low-level convergence line caused by the outflow from an earlier shower.
7. Local re-generation of storms.

Early forecasts at all resolutions miss-placed the upper-level rotation and the forecasts were poor. Later forecasts captured that rotation well. In addition the 1 km model was able to capture the low-level convergence lines (and the earlier shower) and the forecasts were extremely good.

Case E 19th June 2005

1. Larger-scale rotation in the upper-level flow.
2. Very warm air at low levels.
3. Considerable potential instability.
4. A stable layer (or ‘lid’).

5. Enhanced low-level convergence due to local deflection of the flow around the North York Moors.
6. Low-level frontal zone.

The first component in the list (larger-scale rotation) was not correct (or even close to being correct) in any of the forecasts at any of the resolutions. Some of the 1 km forecasts were able to capture all the rest of the components, but despite that, all the forecasts were poor because the first component was missing.

Case F 4th July 2006

1. Very warm air at low levels
2. Considerable potential instability
3. A stable layer (or 'lid')
4. Storm location probably determined by variations in low-level temperature and humidity
5. Flow varies with height in a way that is conducive to the generation of new storm cells in the same geographical location as previous cells.
6. Local storm dynamics important.

The 12-km forecasts did not have sufficient resolution to develop such localised storms. The 4-km forecasts were also generally poor because of insufficient resolution. The 1 km model could produce storms but was not particularly accurate spatially because the low-level variations in the temperature and humidity were not particularly accurate in the first place. The 1 / 1.5 km forecasts were able to produce high rainfall totals because they could generate new storm cells in the same geographical location (behind the original cell in a system relative frame of reference). The difficulty in predicting this event was highlighted by the sensitivity studies. Changes to the number of vertical levels (76 to 38 levels) or horizontal resolution (1 to 1.5 km) or more diffusion changed the way the storm cells were represented, which in turn had an impact on the development of new storm cells, which in turn changed the rainfall totals significantly.

Some of the components showed up on all three occasions. They were the potential instability, the low-level warm air and the lid, and are typical pre-cursors in severe thunderstorm situations. Others like the upper-level rotation or low-level convergence act to focus the convective activity into a particular region or specific location. The local storm dynamics may determine whether rain falls over an area for an extended period or moves through more quickly.

11.4 The importance of the larger-scale dynamics

This project has examined the performance of a kilometre-scale model, but that does not mean all the concerns about forecast skill are at that scale. It is important to recognise that when trying to predict a severe convective storm with a storm-resolving model, the larger scale dynamics must be sufficiently correct or the forecast will go wrong.

In case B (3rd August 2004); as the forecasts got closer and closer to the event, the upper-level vortex became more accurately represented and this made a big improvement to the forecast skill. In case E (19th June 2005) the upper-level rotation was wrong in all the forecasts and none of the forecasts were able to produce the rain that was observed.

It was found in the Extreme Events Recognition Project (Hand et al 2004) that half of the extreme convective events were related to larger-scale forcing. The study by Roberts (2000) showed that nearly all storm clusters that were large and characterised by a high density of lightning discharges were associated with a mesoscale upper-level vortex that could be identified in satellite water vapour imagery.

The use of a storm-resolving model will not correct deficiencies in the representation of the larger-scale flow that may come into the domain from the coarser-resolution model in which the high-resolution model is embedded. Such errors may have a significant impact on the prediction of extreme rainfall events. The way to reduce those errors is to improve the data assimilation in the coarser-resolution model or embed the high resolution model within a coarser-resolution ensemble that can sample the uncertainty in the larger-scale dynamics.

The above statements may seem intuitively obvious, but it is important to emphasise because it contradicts a view often held about the nature of convection over the UK. It is typical for scientists and forecasters to focus on just the local effects as if that is all we need to get right and assume that there is no variability on larger scales or that it is automatically correct or of little consequence. That is not an appropriate view to take.

11.5 The importance of the high-resolution grid

A high-resolution grid (~1 km grid spacing) is absolutely essential to be able to represent the dynamics of more localised thunderstorms and many of the important local pre-cursors to the triggering of storms.

For Case B (3rd August 2004) the 1 km model gave a good forecast once the larger-scale upper-level vortex was sufficiently correct, but having a good forecast also depended on the model being able to represent the low-level convergence lines along which the showers first developed. These were not represented in the 12 km model and poorly at 4 km. One of the convergence lines developed at the edge of a cold outflow from a previous shower. Only the 1 km model had sufficient resolution to generate the earlier shower and then through the cloud microphysics produce the cold outflow that lead to a sharp convergence line. Once the storms formed the outflows produced new convergence lines along which new storms developed. This too, could only be well reproduced by the 1 km model. The 12 km model couldn’t resolve the storms in the first place and the 4 km model wasn’t able to reproduce the backward development of new storm cells.

In Case F (4th July 2006) the high rainfall totals occurred because of the backward propagation of localised storms. The mechanism involved new cells forming where the warm moist inflow met the cold outflow from the storm and the local convergence and ascent was sufficient to allow new convection to break through the lid. The relative strength of the inflow and outflow meant that the new cells formed behind the moving storm but relative to the ground the new cells had developed in the same location as previous cells and this is what gave the localised area of very high rainfall. The 1 km model was able to reproduce the same behaviour because it had sufficient resolution to represent the storm cells and the storm-scale dynamics that lead to the formation of new cells. Not only that but new cells, although too small, formed where the old cells had been and did give high rainfall accumulations (although still too low). The 12 km forecasts could not resolve any storms and the 4 km model had a long delay in initiating any storms because of insufficient resolution.

11.6 Predictability

The cases have shown that forecast accuracy will vary considerably from case to case and depends crucially on getting all the necessary meteorological components correct. Some events will be inherently more predictable than others and this will depend on the predictability of the meteorological components that are relevant on that particular day. An examination of the predictability of the components can help to distinguish between more predictable and less predictable events. This sort of information could be used to help with forecast decision making if the components are known in advance.

It's worth considering the expected predictability of the components in the cases examined. All of the three main cases had very warm air at low levels; something which might only happen a few days a year, but we have to be careful here to make the distinction between uncommon and difficult to predict. A very hot air-mass may not be very common, but may nevertheless be very predictable in a short-range NWP model forecast because it covers a large area and doesn't change quickly. In the same way the potential instability and the presence of a lid are reasonably predictable. What's less predictable is the exact nature of the lid and whether the air below is hot enough to allow convection to get past that lid. In situations when the larger-scale dynamical forcing dominates, the strength of a lid may be less significant and the major source of forecast uncertainty will be the characteristics of the dynamical forcing.

11.6.1 Classification

From an examination of the meteorological components in the cases studied, severe convective storms can be split into three categories according to the meteorology that leads to their formation and hence their predictability in NWP models.

1. Storms that are strongly forced by the local topography.
2. Storms where the larger-scale dynamics play a major role.
3. Storms that depend on subtle variability in the local environment.

Of course this is not meant to imply that all severe storms will neatly fall into one of these categories without any overlap. Most will be a combination, but they will usually settle into one category more than the others.

Storms that are strongly forced by the local orography

The event that fits into this category best is case C, 16th August 2004, the Boscastle flood. The location of the storm was very strongly tied to the shape of the coastline and hills. This sort of event is the most predictable. We should expect that a storm-resolving model will be able to generate convection in the right place on these occasions if the orography and coastline is sufficiently well represented, which it should be on such a fine grid. The nature of the coastline of the UK means that local convergence zones and sea breeze fronts can play an important role in storm initiation. Contrast the situation in the UK with the American mid-west where there is no coastline and variations in the orography are small.

Storms where the larger-scale dynamics play a major role

Three of the events clearly fit into this category; case B (3rd August 2004), case E (19th June 2005), case J (20th July 2007). The predictability of the storms will be largely determined by the predictability of the larger-scale dynamics. In general, the larger-scale dynamical features (e.g. upper-level vortex, frontal zone) will be reasonably predictable over a 12 to 24-hour forecast period if they are well represented at the start of the forecast. Of course, they may well be outside the high-resolution domain at the start and it is not necessarily the case that they will be well represented if they come from data-sparse areas over the sea. Case E was an example of a very poor representation of an upper-level vortex, but to be that much in error is unusual, especially now that the data assimilation system in the 12 km NAE model has been significantly improved (inclusion of 4DVAR (Rawlins et al 2007)). The loss of predictability will usually be not so much do with a larger-scale dynamical feature being missed, but that it is somewhat misplaced and miss-shaped and so the area of storm triggering becomes more uncertain. The smaller the dynamical feature of concern the more of an issue this becomes.

Storms that depend on subtle variability in the local environment

The event that fits best into this category is case F, 4th July 2006. The triggering depended on local variations in the warmth and humidity of the air near the ground and variations in the strength of the lid. These situations may be marginal; convection will not be severe unless the lid is strong enough to prevent weaker showers developing more widely instead, but if the lid is too strong convection may not develop at all. Once a storm has developed it will not usually persist for long in an environment with little dynamical or topographical forcing unless the local storm dynamics and background flow can interact to generate new cells. If the new cells can form in the same location then very high rainfall totals are possible. On some occasions the interactions between new cells can lead to a much larger storm complex that can maintain itself for several hours. This occurred in case H (10th May 2006).

This category is the least predictable for two reasons: (1) the initial triggering depends on local variations in the atmosphere that are difficult to detect with conventional observations, and may therefore not be represented well in the model, and (2) once convection has developed, what happens next is critically dependent on the first storms being modelled correctly, both in location and structure.

One of the challenges in future is to somehow be able to predict in advance how predictable a particular situation will be. It may be possible to use these categories as a guide. If we know in advance that the model is predicting a category 1 storm then we can be more certain about the forecast than if it is a category 3 storm and the uncertainty we convey with the forecast can be adjusted appropriately. There will also be other ways of assessing the predictability, for example, using ensembles or forecast consistency or particular key diagnostics. More work is needed in this area.

It is worth saying that for storms or showers in general, i.e. not just severe, there is another category; ‘scattered showers’. With scattered showers the predictability of each shower is low, but the showery regime is usually very predictable. Widespread scattered showers do not become extreme events in terms of rainfall amounts and so they are not of concern here.

11.7 Presentation and probabilities

There will always be uncertainty in forecasts of extreme rainfall events and the more precision we expect the more uncertainty there will be. It is essential that probabilistic outputs for users are developed and that users understand what they mean.

Some possible probabilistic products have been presented in this report. There are a myriad of other possibilities (Roberts 2004). The products shown in section 9 were mostly constructed from single forecasts using an estimate of the spatial uncertainty of the forecast. Products in the near future should also incorporate information from older forecasts (a 'time-lag ensemble'). In the short term, that is the way forward. Further into the future probabilistic output should come from a high-resolution ensemble forecast system.

Education into the use of probabilities is vital. If we don't include uncertainty in the output and don't properly explain what it means, we will miss out on the full benefit of having a sophisticated high-resolution modelling system.

The scale-selective verification approach used to obtain an objective measure of forecast skill (the Fractions skill Score) has proved to be useful for determining the scales over which a forecast is deemed to have sufficient accuracy and therefore the estimate of uncertainty that should be used for forecast products (Roberts and Lean 2008). It is now being adopted for evaluating the UK 4-km model and should also be used for higher-resolution models both in a research context and then operationally. This method has also proved useful for determining the scales at which changes to model parameters have an impact and could be used to examine differences between ensemble members. It has also shown by Roberts and Lean that improved resolution does give more accurate forecasts on average on scales that are useful for flood warning even if there is only small improvement at the scale of the grid.

11.8 Orographic rainfall

Most of the cases were convective and it is for the convective situations that we expect the greatest impact from a storm-resolving model.

Two of the cases, though, were frontal and the rain was significantly enhanced in places by the local orography. The first was case D (8th January 2005 – the Carlisle flood), the second was case I (25th June 2007 – the Hull/Sheffield floods) and both were very significant.

The study of case D showed that the 4 km and 1 km models gave much better rainfall predictions than the 12 km model. The benefit was retained when the rainfall predictions were fed into a hydrological model (Roberts et al 2008, submitted to Met Apps). Similarly, for case I, both the 4 km and 1 km models gave an enhancement to the rainfall over the hills to the west of Sheffield, which is what was observed. The 1km model was probably more skilful than the 4 km model in each case, but verification is difficult in these situations. Either 1 km or 4 km is considerably better than a 12 km model and we can expect more accurate forecasts of this type of event in future.

11.9 Hydrological modelling

High-resolution model output should now be starting to be used in hydrological models for flood warning, particularly for orographically enhanced rainfall situations such as case C (the Carlisle floods in 2005). The general area of rain in these situations should be reasonably predictable over the 3 to 18 hour forecast period (Roberts 2008) and the high-resolution model is then able to simulate the orographic enhancement. As mentioned above, this has been tried for the Carlisle event and the results were very promising (Roberts et al 2008).

In convective situations the forecast uncertainty is greater, but the results from this project still suggest that it would be worthwhile to feed storm-resolving rainfall forecasts into hydrological models for flood warning, with the proviso that a more probabilistic approach is taken.

11.10 Modelling issues

11.10.1 Ensembles

When trying to predict severe convective storms with a 'storm-resolving' model (grid-length ~1 km), the larger scale dynamics must be sufficiently correct or the forecast is very likely to be poor. This is particularly true of type-2 storms, which cover a large proportion of severe storms (Roberts (2000), Hand (2004)). However, even if storms are not directly triggered by a larger-scale dynamical feature they still need to be in the right environment for a good forecast to be possible (e.g. the correct side of a frontal zone).

Case B (3rd August 2004, London) showed that as the forecasts got closer and closer to the event, the upper-level vortex became more accurately represented and this had a big impact on forecast skill. We were, in a sense, looking at an ensemble of forecasts in which the larger-scale dynamics played a prominent role. This demonstrates the value an ensemble forecast system might have. It may well have been the case that with an ensemble some of the forecasts with longer lead times would have had the correct upper-level vortex, or even some of the forecasts in case E (19th June 2005, Hawnby) might have had the correct upper-level flow. Also, confidence in the 1-km solution was increased because of the run-to-run consistency. In the same way consistency amongst members of an ensemble can raise confidence in the solution (provided the ensemble has a history of representing uncertainty well). The Met Office now has an ensemble forecast system (MOGREPS) with 24 members and a grid spacing of 24 km running for 54 hours (Bowler et al 2007). The problem when predicting extreme convective events is that a grid spacing of 24 km is not sufficient to represent the storms themselves or local dynamical precursors, which is precisely why a ~1 km model has been developed.

A useful way forward would now be to combine the two and nest the ~1 km model inside ensemble members. At present, computer resources are insufficient to run a large ensemble of high resolution forecasts (even in research mode), so there will need to be a methodology for selecting appropriate members from the ensemble. In terms of predicting the possibility of intense convection this will have to involve an automatic way of differentiating coarser-resolution forecasts that are more likely to produce severe convection from those that are not and just running from a select few of the members. This needs to be an area of research. It may be reasonably easy in situations when convection is strongly forced from the larger-scales (type 2), but in other situations when variations in the larger-scale environment are less important and the storms are triggered and maintained by local effects (e.g. the Boscastle storm, Golding 2005) (types 1 and 3) it may not be trivial. The aim would be to have a system that could provide plausible alternative scenarios for convective storms and use a 'storm-resolving' model to represent the convection under those different scenarios. A further area of research would then be needed to examine appropriate ways of presenting risk and uncertainty from such a system.

Looking further ahead, there will be benefit from perturbing the high-resolution model in a suitable way to generate a short-range high-resolution ensemble that is based on uncertainties in the more local meteorology as well as the larger-scale dynamics. Research is necessary in this area.

11.10.2 Data assimilation

Data assimilation (the process of making the start of a new forecast fit new observations better) can fall into two categories, both of which are equally important, but fulfil a different role. (1) Data assimilation in the coarser resolution model that provides the boundary information to the storm-resolving model. (2) Data assimilation on the storm-resolving model grid.

Coarser resolution data assimilation

Much of the discussion above has been about the relevance of larger-scale dynamical features such as disturbances in the upper-tropospheric flow that may come from outside the high-resolution domain. It is therefore important to continually improve data assimilation systems in the coarser resolution model that provides the boundary information, (as it always was before storm-resolving models were developed). This is particularly true for the UK where dynamical features on scales of ~100 to 500 km can spin up in data sparse areas over the Atlantic. It will be most important for forecasts beyond ~6 hours, and may be less of an issue for nowcasting systems (0-6 hours), but even then it will matter. The main point here is simply that data assimilation at coarser resolution is very important even when moving to a high-resolution grid.

Data assimilation on a high-resolution grid

In cases B and E we saw the importance of low-level convergence lines for the initiation of convection. If those lines had not been present or were wrongly located, the initiation of convection would have been wrong. Since these features can only be properly represented in the ~1-km model (not at 12 km) there is a need for a data assimilation system that can adjust the model dynamics on those scales. It will require the introduction of advanced data assimilation methods using high-spatial-density observations (e.g. Doppler winds from radar).

Not only is it important to capture the local winds and any convergence lines at the start of a forecast, it is also important to represent the causes if the correct behaviour of the flow is to be maintained into the forecast. When the cause is to do with orography or coastlines it is not so problematic because they don't move and are represented in sufficient detail in a ~1-km model (provided things like sea-surface temperatures, surface roughness and soil moisture are accurate enough). However, in cases B and E, there were convergence lines that existed because of the outflow from earlier showers. This is more problematic because the correct thermal, moisture, cloud and rainfall structures are also needed. In addition, the earlier showers were themselves dependent on having the correct 3-D moisture and temperature patterns. Other factors like cloud shadowing require an accurate cloud distribution. Frontal convergence requires an accurate representation of the frontal structure.

The location of storm initiation in case F was critically dependent on the temperature and humidity pattern near the surface and the strength of the lid. Again, data assimilation on the high-resolution grid will be required to get this right.

New developments in high-resolution data assimilation could be instrumental in improving short-range forecasts of severe storms in high-resolution models at local scales. This will have to include the use of new observation types such as Doppler winds from radar and humidity measurements based on refractivity measurements from radar. Just sticking to traditional observation types will not provide the coverage that is required. High resolution data assimilation research is vital, especially for improving nowcasts (0-6 hours), but for longer period too. It is challenging, but the potential benefits are great.

11.10.3 Cloud physics and turbulence

The importance of getting the environment correct prior to the initiation of the storms has already been discussed. This has neglected the representation of the showers themselves. If the rainfall predictions are to be accurate, then the development of the showers from non-precipitating clouds through to full-blown thunderstorms needs to be modelled correctly. The timing of initiation, growth of the showers, intensity of the showers, development of secondary cells and decay of the showers are all important. For example, a forecast of storms that are in the correct location but the wrong size intensity and motion will not be as useful as it could be.

One of the problems we have seen in the two cases (and others from stage 1, Roberts 2006 and from the Storm Scale Modelling Project, Roberts 2005) is the ‘measles' effect. This was seen again in cases B and F. Showers appear as a rash of small intense cells (perhaps to early) that are too extensive and do not readily upscale into larger storms. It is thought that this behaviour is partly due to insufficient sub-grid-scale mixing within clouds because the only mixing above the boundary layer comes from numerical diffusion (the convection parametrization scheme provides additional mixing in the 12-km model but is not used at ~1km). This is about to be addressed by the introduction of a turbulence parametrization that essentially mixes the air between grid points where the vertical velocity is high – i.e within the clouds. Experiments with types of sub-grid-scale mixing for cases B and F have shown that both were sensitive and for case F (without larger-scale forcing), the forecasts were extremely sensitive to both turbulent mixing and small changes in resolution It meant the difference between a few weaker storm cells or a large storm complex and highlights the need to have an appropriate sub-grid-scale mixing for the particular grid spacing being used. Research and development is required and ongoing.

The other factor to consider is the way cloud and precipitation is represented in the model. This too will have a bearing on the structure of the showers and the amount of rain produced and will clearly interact with the new turbulence parametrization. The representation of the cloud and precipitation will play an important role in the generation of cold-pool outflows because the downdraughts are driven by cooling from evaporating rain. In case B, this was crucial for the formation of the convergence line that lead to the initial triggering of the deep convection and it seems that an error in the location of the earlier showers was offset by an over-prediction of their intensity. It is likely that changes to microphysics could make a significant difference to the forecasts of case F. They would affect the strength of the updrafts and downdrafts and then the development of new cells. This case could be a good test-bed for new developments in cloud microphysics.

12. Conclusions

12.1 Key Findings

1. There are no differences between the physical and dynamical processes that lead to ‘extreme’ events compared to other heavy rainfall events. It is the coincidence and interaction of those processes that leads to extreme events and the difficulty in their prediction.

2. A storm-resolving NWP model is capable of providing useful forecasts of ‘extreme’ rainfall events. The use of a storm-resolving model has the potential to improve greatly our current ability to predict such events provided that it is understood that the output must include information about forecast uncertainty (e.g. probabilities).

3. The accuracy of the forecasts may vary considerably from case to case and depends crucially on getting all the necessary meteorological components correct.

4. For many events is vital to accurately represent larger-scale disturbances in the flow that may originate from outside the high-resolution domain; as well as any local effects.

5. A high-resolution grid (~1 km grid spacing) is absolutely essential to be able to represent the dynamics of more localised thunderstorms and many of the important local pre-cursors to the triggering of storms.

6. Some convective situations are inherently more predictable than others, and that predictability is strongly linked to the meteorology of the situation. A classification into three types of storm has been made. Knowledge of the likely type of storm can provide information about its predictability.

7. This work has lead to a greater understanding of both the meteorological processes that lead to extreme rainfall events and the strengths and weaknesses in the model in predicting them.

12.2 Recommendations

1. The development and operational implementation of a storm-resolving model (grid spacing ~ 1 - 2 km) should continue. It is likely to be the best route towards a significant improvement in our ability to provide warnings of severe convective storms.

2. There will always be uncertainty in forecasts of extreme rainfall events and the more precision we expect the more uncertainty there will be. It is essential that probabilistic outputs for users are developed and that users understand what they mean.

3. The model is capable of representing extreme storms, but getting the positioning correct presents a big challenge. Research and development into new data assimilation methods (for improving the initial state of forecasts) on

the high resolution grid and the use of new types of observations in data assimilation is, and must continue to, play a vital role in the development of a storm-resolving model.

4. The uncertainty in a single 'deterministic' forecast is difficult to quantify and this can limit the usefulness of probabilistic outputs based on one forecast. We should now move towards developing a storm-resolving ensemble prediction system and ensemble-based probabilistic products. Initially, such a system would embed the storm-resolving model in a coarser-resolution ensemble. The reason for taking this approach is that, over Western Europe, the dominant source of spatial uncertainty will often come from misplaced larger-scale dynamical features.

5. Further research is required to determine whether an estimate of the predictability can be deduced in advance from the meteorological conditions. Research into the use of 'time-lag' ensembles (combining the most recent forecast with older forecasts) as a possible means of determining predictability should also be undertaken.

6. High-resolution model output should be used in hydrological models for flood warning, particularly for orographically enhanced rainfall situations such as the Carlisle floods in 2005. For convective situations, a more probabilistic approach is necessary and needs to be developed.

7. There are still some outstanding issues to do with the way a storm-resolving model represents convective cells. Further work is needed in the areas of cloud microphysics and sub-grid-scale turbulence.

Appendix 1

The generation of fractions and computation of the Fractions Skill Score.

Figure A1 gives a schematic picture of how fractions are computed over different sized squares. In this example the threshold has been exceeded where the grid squares are shaded and not reached where the grid squares are white. If we focus on the central grid square, then at the grid square itself (i.e. the grid scale) the forecast fraction is 0/1 = 0, but the radar fraction is 1/1 = 1 (the forecast is wrong). Over a 3x3 square the forecast fraction is 4/9 = 0.44 and the radar fraction is 3/9 = 0.33. Over the whole 5x5 domain the forecast fraction is 6/25 = 0.24 and the radar fraction is also 0.24 (the forecast is correct for that specific central grid square over that scale).

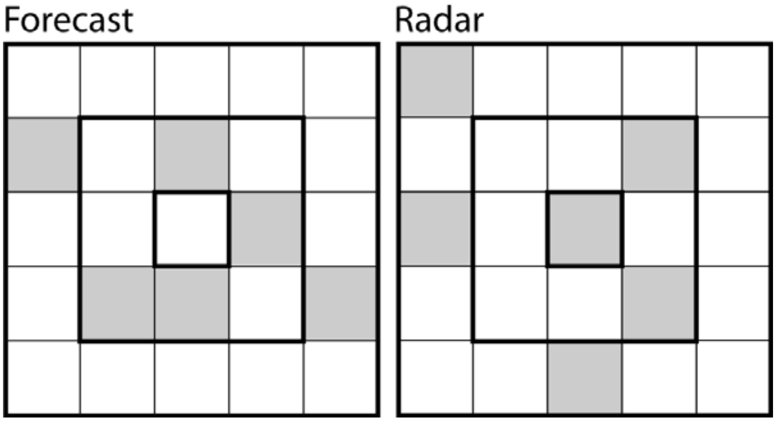


Figure A1. A schematic comparison between forecast and radar (see text)

Computing the FSS

The FSS is a variation on the Brier Skill Score.

$$FSS = 1 - \frac{FBS}{FBS_{worst}}$$

Where *FBS* is the Fractions Brier Score and is a variation on the Brier Score (Brier, 1950) in which both the forecast and observed probabilities (fractions) can have any value between 0 and 1. It is given by:

$$FBS = \frac{1}{N}\sum_{j=1}^{N}(O_j - M_j)^2$$

M_j and O_j are the forecast and radar fractions respectively at each point, with values between 0 and 1. FBS_{worst} is given by:

$$FBS_{worst} = \frac{1}{N}\left[\sum_{j=1}^{N} O_j^2 + \sum_{j=1}^{N} M_j^2\right]$$

It is the largest *FBS* that could be obtained from the forecast and observed fractions when there is no collocation of non-zero fractions and therefore the worst possible *FBS*.

The FSS has the following characteristics:

1. It has a range of 0 to 1; 0 for a complete forecast mismatch, 1 for a perfect forecast.
2. If either there are no forecast grid squares which exceed the threshold and some occur, or some are forecast and none occur, the score is always 0.
3. As the size of the squares used to compute the fractions gets larger, the score will asymptote to a value that depends on the ratio between the forecast and observed frequencies of the event. I.e. the closer the asymptotic value is to 1, the smaller the forecast bias. The use of percentile thresholds ensures that the FSS tends to 1 as the neighbourhood size approaches the size of the verification area.
4. The score is most sensitive to rare events (or for small rain areas).

A more complete discussion of the FSS is given in Roberts and Lean, including a discussion of the link between the verification method and probabilistic post-processing of precipitation forecasts.

13. References

Browning et al. (2006). A summary of the Convective Storm Initiation Project Intensive Observation Periods. Met Office NWP technical report 474.

Browning, K.A. and N.M. Roberts (1994) Use of satellite imagery to diagnose events leading to frontal thunderstorms. Part I of a case study. Meteorol.Appl. 1, 303-310.

Bowler, N. E., A. Arribas, K. R. Mylne and K. B. Robertson, (2007): The MOGREPS short-range ensemble prediction system Part I: system description, Met Office NWP Technical Report No 497

Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, S. E. John and T. P. Legg (2007): The MOGREPS short-range ensemble prediction system Part II: Performance and Verification, Met Office NWP Technical Report No 498

Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., and Wood, N. (2005). A new dynamical core for the Met Office's global and regional modelling of the atmosphere. Q. J. R. Meteorol. Soc., 131, 1759–1782.

Deng A., and D. R. Stauffer. (2006): On improving 4-km mesoscale model simulations. J. Appl. Meteor. Climatol., 45, 361-381

Essery, R., M. Best, and P. Cox, 2001. MOSES 2.2 Technical Documentation, Hadley Centre Technical Report No. 30, Met Office Hadley Centre

Golding, B. (1998). Nimrod: a system for generating automated very short range forecasts. Meteorol. Appl., 5, 1–16.

Golding, B. W (editor)., 2005: Boscastle and North Cornwall Post Flood Event Study - Meteorological Analysis of the Conditions Leading to Flooding on 16 August 2004. Met Office NWP Technical Report., 459

Gregory, D. and P. R. Rowntree, 1990. A Mass Flux Convection Scheme with Representation of Cloud Ensemble Characteristics and Stability-Dependent Closure. Mon. Wea. Rev., 118 1483-1506

Halliwell, C. (2006). Implementation of a subgrid turbulence scheme in the UM: Version 1.3. UM Documentation paper 28, Met O±ce, UK., pages 1-17.

Harrison, D.L., S. J. Driscoll, and M. Kitchen, 2000. Improving precipitation estimates from weather radar using quality control and correction techniques. Met. Appl., 7, 135-144.

Hand, W. H., and Fox, N. I. and Collier, C. G. (2004), A study of twentieth-centuary extreme rainfall events in the United Kingdom with implications for forecasting. Meteor. Apps., 11, 15-31.

Jones, C.D., and B. Macpherson, 1997. A latent heat nudging scheme for the assimilation of precipitation data into an operational mesoscale model. Met Appl., 4, 269-277

Lean, H.W., and P. A. Clark, 2003. The effects of changing resolution on mesoscale modelling of line convection and slantwise convection. Quart. J. Roy. Meteor. Soc., 129, 2255-2278

Lean H. W. S.P. Ballard, P.A. Clark M.A.G. Dixon, Zhihong Li and N. M. Roberts (2005). The Summer 2004 Reruns with the High Resolution Trial Model. Met Office NWP technical report 466

Lean, H.W., and P. A. Clark, 2003. The effects of changing resolution on mesoscale modelling of line convection and slantwise convection. Quart. J. Roy. Meteor. Soc., 129, 2255-2278

Lilly, D. K. (1967). The representation of small-scale turbulence in numerical simulation experiments. Proc., IBM Scienti¯c Computing Symposium on Environmental Sciences, pages 195-210.

Lock, A.P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000. A New Boundary Layer Mixing Scheme. Part 1: Scheme Description and Single-Column Model Tests. Mon. Wea. Rev., 128 3187-3199

Lorenc, A.C., S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. D. Dalby, D. Li, T. J. Payne, and F. W. Saunders, 2000. The Met. Office global three-dimensional variational data assimilation scheme. Quart. J. Roy. Meteor. Soc., 126, 2991-3012

Moore RJ. (2007). The PDM rainfall–runoff model. Hydrology and Earth System Sciences, 11(1): 483-499.

Rawlins, F., Ballard, S.P., Bovis, K.J., Clayton, A.M., Li, D., Inverarity, G.W., Lorenc, A.C., Payne, T.J. (2007) The Met Office global four-dimensional variational data assimilation scheme. Q. J. R. Meteorol. Soc., 133, 347-362

Roberts, N. M. (2000). A guide to aspects of water vapour imagery interpretation: the significance of dry regions. Met Office NWP technical report 294

Roberts, N. M. (2000). The relationship between water vapour imagery and thunderstorms. Met Office NWP technical report 300

Roberts, N. M., (2003). 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

Roberts, N.M., (2004). Measuring the fit of rainfall analyses and forecasts to radar: Strategy for stage 5 evaluation. Stage 4 report from the Storm-Scale numerical. Met Office NWP technical report 432.

Roberts, N.M., (2005) An Investigation of the ability of a storm scale configuration of the Met Office NWP model to predict flood producing rainfall. Final scientific report from the storm scale modelling project. Met Office NWP technical report 455.

Roberts, N. M. (2006). Simulations of extreme rainfall events using the Unified Model with a grid spacing of 12, 4 and 1 km. Met Office NWP technical report 486.

Roberts, N. M. (2007) Meteorological components in forecasts of extreme convective rainfall using 12-km and 1-km NWP models: A tale of two storms. Met Office NWP technical report 500.

Roberts, N. M. and H. W. Lean (2008). Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Submitted to Monthly Weather Review. Accepted following minor revisions.

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. Part 1:the basic experiment. Mon. Wea. Rev., 91, 99-164.

Smith, R.N.B., E. M. Blyth, J. W. Finch, S. Goodchild, R. L. Hall, and S. Madry, (2006). Soil state and surface hydrology diagnosis based on MOSES in the Met Office Nimrod nowcasting system. Met Appl., 13, 89-109

Theis, S.E., A. Hense, and U. Damrath, (2005): Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. *Met. Appl*., **12**, 257-268

Wilson, D. R., and S. P. Ballard, (1999). A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. Quart. J. Roy. Meteor. Soc., 125, 1607-1636

Met Office technical reports are available from the Met Office web site.

Acknowledgements

The work was half funded by Defra/EA under the Flood and Coastal Erosion Risk Management R&D Programme. I would like to thank Brian Golding for his reviews of this and earlier reports and Carol Halliwell for her comments on the part about sub-grid-scale turbulence. I would also like to thank the project board comprised of Melanie Andrews, Peter Clark, Tim Harrison, Bob Hatton, Tim Harrison, Tim Wood and Adrian Wynn for their valuable contribution.

Ergon House
Horseferry Road
London SW1P 2AL

www.defra.gov.uk

defra
Department for Environment
Food and Rural Affairs