The University of Reading

Met Office

Numerical Weather Prediction

The Summer 2004 Reruns with the High Resolution Trial Model

Forecasting Research Technical Report No. 466
Joint Centre for Mesoscale Meteorology Report No. 152

H.W. Lean, S.P. Ballard, P.A. Clark M.A.G. Dixon, Zhihong Li and N.M. Roberts

email: nwp_publications@metoffice.gov.uk

The Summer 2004 Reruns with the High Resolution Trial Model

Forecasting Research Technical Report No. 466

Joint Centre for Mesoscale Meteorology Report No. 152

H.W.Lean, S.P.Ballard, P.A.Clark, M.A.Dixon, Zhihong Li and N.M.Roberts

Version History

Version	Date	Approval	Changes
1.0	7/06/05		Initial version
2.0	1/07/05		Incorporate NR changes
3.0	16/08/05		Incorporate PC changes
	21/09/05	Brian Golding	Approved for release as technical report

Abstract

The High Resolution Trial Model (HRTM) suite consists of 12km, 4km and 1km models over the southern UK. This paper describes runs of these models for seven cases from summer 2004 in order to assess the high resolution models performance in short range forecasting of convective events. This work follows on from four similar case studies run on summer 2003 cases including various improvements to the model and assimilation systems.

Generally this work confirms the conclusions from the summer 2003 trial that there are potential benefits from high resolution models. Compared with 2003 there is more evidence of the over prediction of rainfall. This may, in part, be due to the fact that the convection was, generally, stronger in the 2004 cases. However in the 1km model the change may also be due to the inclusion of assimilation of cloud and rainfall data with the MOPS/Latent Heat Nudging system.

Precipitation statistics once again show that, although the results are useful, more cases are required to form definitive conclusions. The 6 hour accumulation statistics show that the 12km model is better than the 4km and 1km for absolute thresholds – this is due to the over prediction of rain. If the bias is removed before calculating the statistics so that only the spatial skill is measured, the 4km and 1km models are better showing that the structure of the precipitation fields is better. For statistics calculated for 1 hour accumulations the 4km and 1km models appear better even for absolute thresholds, except for at the first few hours of the forecasts. This shows that these models are providing valuable information on shorter timescales.

Future work on these models is discussed including ideas on how to reduce the over prediction of precipitation.

1. Introduction

Met Office is looking to introduce models into its operational NWP suite of higher resolution than the current smallest gridlength of 12km in order to improve short range forecasts of convection and other types of hazardous weather. With the introduction of the non-hydrostatic version (Davies et al 2005) of the Unified Model (UM) and the continuing increase in available computer power this is becoming more feasible. A 4km model is currently in the process of being implemented in the operational suite with the prospect of a 1km model by 2007 or 2008. This paper reports some tests of a prototype forecasting system, the High Resolution Trial Model (HRTM), which consists of a 4km and a 1km gridlength model over the southern United Kingdom.

The HRTM project is a progression from the case-study based work on high resolution (24km up to 2km) versions of the UM which was carried out in JCMM from 2000 onwards (e.g. Lean and Clark 2003) This work was on a variety of cases and showed that high resolution models have the potential to improve forecasts and the representation of various phenomena. The HRTM project started in mid 2002 and aims to develop and evaluate 4km and 1km gridlength models in a scenario closer to an operational system. The emphasis of this project has been short range (out to T+6-9) forecasts of convective rain (particularly severe convective rain). The basic concept is to run the models on fixed domains for a large number of cases in order to gather verification statistics. These statistics are then used to evaluate the potential advantages of high resolution models.

The first phase of the project consisted of a number of sensitivity studies in order to determine satisfactory values of a number of model parameters (timestep, horizontal diffusion etc). In the second phase of the project data assimilation was added to the high resolution models and trials were run on four cases from the summer 2003 period (Lean et al 2005). The conclusion from this trial was that high resolution models had the potential to provide improved precipitation forecasts. This conclusion was also reached by the Storm Scale Modelling project (Roberts 2005). This report presents the results from a second HRTM trial carried out on six cases from summer 2004. In addition to running a somewhat larger sample of cases this trial included some changes both to the model configurations and to the data assimilation systems used.

2. Model Configurations

The suite of models comprised 12km, 4km and 1km models with the same horizontal and vertical domains as used in the summer 2003 trial (Lean et al. 2005). The 4km model used the same 38 vertical levels as the 12km model and the 1km used 76 levels which were the 38 level set doubled. The 12km model was set up to be identical to the operational configuration and was run to provide both a comparison to the high resolution models and lateral boundary conditions. The 4km and 1km models both covered the southern UK and were approximately centred on the Chilbolton radar. As previously the 1km model was run without a convection scheme and the 4km model was run with a modified scheme to restrict the mass flux and encourage much of the convection to take place explicitly. One of the parameters of this scheme was changed as discussed in section 2(a) below.

The most important changes to the suite in the current work from the summer 2003 trial were in the data assimilation aspects. There were also some changes to the model configurations. The rest of this section describes these changes.

2(a) Model parameters

The model configurations used are summarised in table 1 and are largely the same as used in the summer 2003 trial (Lean et al. 2005). The main changes which were introduced into the models for this trial are briefly discussed below.

Supercomputer/Model version

Whereas the summer 2003 trial used UM version 5.5 running on the t3e supercomputer the current trial used UM version 6.0 running on the new NEC SX-6 supercomputer.

Convection Scheme in 4km model

As described in Lean et al (2005) the 4km model was originally run with the CAPE dependent CAPE closure timescale modification to the convection scheme (Roberts 2003). The aim of this is to restrict the mass flux in the convection scheme and encourage convection to take place explicitly in strongly forced situations. The parameter c, which determines how much mass flux the convection scheme is able to generate, was set to c=0.01J/Kg which is a small value which means, in practice, that virtually no rainfall is generated by the convection scheme. This has the side effect that the tendency for extreme rainfall rates in individual cells is increased. In order to reduce these problems the value was increased to c=0.5 J/Kg in runs reported here.

4km Model Timestep

The timestep for the 4km model was increased from 60s to 100s for the current work after tests showed that the change appeared to have very little impact on the output

Prognostic Rain

In the operational model the rainfall rate on each layer is diagnosed on each timestep according to the sources and sinks including a flux from the layer above and to the layer below. No account is taken of advection by vertical or horizontal wind. From typical wind speeds and rain fall velocities one would expect the horizontal advection of rain to have an effect on scales of around 10km. Models running at about this or higher resolution should therefore include prognostic rain which is properly advected by the winds. The coding to allow this facility was carried out by Richard Forbes and this modification has been added to the 4km and 1km models. For the 1km model the prognostic rain is included in the boundary conditions output from the driving 4km model. Although it is clear that prognostic rain should be introduced in high resolution models it was found during the course of this work that there are issues with the interaction of prognostic rain with assimilation. This will be discussed in section 6(a).

Ozone

In the operational 38 level 12km model the ozone fields are specified on the top 11 levels (and the value on the lowest of these levels is copied down to rest of the model). In the summer 2003 trials the 4km model had ozone specified on these same 11 levels and the 1km model (with a doubled level set) equivalently on the top 22 levels. In the course of setting up the 2004 trial it was found that with the above configuration the lowest ozone level still had a concentration of one or two orders of magnitude more than typical tropospheric values. In order to avoid the model seeing unrealistically large ozone concentrations in the troposphere it was therefore necessary to specify the ozone on more levels. In the models discussed here there were 24 levels of ozone in the 4km model and (equivalently) 48 in the 1km model.

Aerosol

While testing the models, occasional failures were seen due to negative aerosol. This was found to be caused by negative aerosol sources which arose from vertical interpolation of the aerosol source data. The problem was solved by switching to non-extrapolating interpolation in the vertical.

Additionally the facility has been added to pass aerosol data through the boundary conditions files into nested models. The 4km model now gets boundary aerosol data from the 12km model it is nested in and similarly the 1km from the 4km.

Dry Static Adjustment in 1km Model

It was discovered that the 1km model had been running using dry static adjustment switched on. This arose through some logic which erroneously switched it on when the convection scheme was switched off. This facility to remove regions of dry static instability was originally provided in the model for use in idealised runs and was not intended to be used in real cases. It was removed in the 1km model. The net effect was less rapid mixing in the unstable boundary layer.

12km Model Changes

A number of changes were made to the 12km model only to keep it up to date with changes that had been made to the operational 12km model. The main changes were:

Large Scale Precipitation scheme to 3B (was 3C previously)

Convection to 4A (formerly known as CMODS) from 3C previously

Gravity Wave Drag on with version 4A (off previously).

2(b) Assimilation.

2003 Baseline

As described in more detail in Lean et al (2005) the summer 2003 trials used 3D VAR assimilation in the 4km model. For this work the assimilation systems were largely taken over unchanged from the systems used in the operational 12km model. The 3 hour assimilation cycle length used in the operational 12km model was retained in the 4km and 1km models for the current work. Additionally the Moisture Observations Processing System (MOPS) provided cloud and surface precipitation data at 15km resolution derived from the NIMROD system. This was assimilated into the model using an Analysis Correction (AC) scheme by Latent Heat nudging (LHN) in the case of the precipitation data and humidity profiles for the cloud data. For the 1km model 3D VAR was not run but the increments from the 4km 3D VAR used. For technical reasons (the system not being set up to work with a model without a convection scheme) the MOPS/LHN systems had not been run in the 1km model in the 2003 trials.

2004 Trials

For the trials described here two main changes were made to the assimilation system:

1. 3D-Var. Introduction of scale selective 3D-Var in the 4km model. This is intended to address the problem that due to the relative smallness of the domain boundary effects are more of a problem in the 4km model than in the 12km. Analysis increments from a larger area model contain additional and useful information for the nested smaller, limited area model (LAM), which is missing from the "standard"

smaller area 3D-Var analysis. Inclusion of these increments to the smaller area 3D-Var would provide information from observations outside the LAM domain as well as data near boundary inside the domain. In addition, it may also provide information on scales longer than those that can be represented by LAM ,as there is an upper limit of the wavelength, constrained by the area. To retain the long wave information from the 12km model in the 4km model, the analysis was carried out in the following steps which allowed the 4km 3D-Var to analyse the smaller scale "short waves" (not analysed by the 12km model) only: (a) Spectrally filter the 12km increments to obtain "long waves" increments based on an appropriate cut-off wavelength in this work chosen to be 180km; (b) Add "long waves" to the 4km background; (c) Analyse the "short waves" not retained by the filtered increment added to the background in the 4km 3D-Var; (d) form the new analysis by adding in both the long wave and the short wave increments via the Incremental Analysis Update (IAU).

2. MOPS/LHN. MOPS/LHN was introduced in the 1km model. As the 1km model does not include a convection parameterisation scheme, the LHN procedure was modified appropriately to use only dynamically-resolved model rain and snow rates. All of the MOPS/LHN parameters were set to the same values as used in the 12km model component. The radar data is put into the model at 15km resolution with a time resolution of 1 hour. The cloud data is put in also with a resolution of 15km but with a time resolution of 3 hours. The LHN search radius (the radius over which the model searches for a model profile which results in the same amount of rain as seen in the radar) was kept at 72km in the 1km model (as in the 12km model). There is clearly a lot of scope for tuning the parameters of the 1km LHN/MOPS scheme. This is discussed in section 6(b).

The cases described here had to be rerun after a bug was discovered which meant that the LHN/MOPS scheme was seeing the same model background rain fields (rather than the time evolving ones) throughout the assimilation period. The runs before this bug was fixed occasionally showed large spurious areas of heavy rain appearing during the assimilation cycle.

3. Cases

In general there were more cases of heavy convective rain over summer 2004 than over summer 2003. This, combined with the increased computing power now available, allowed more cases to be run (7 as opposed to 4 in 2003). Whilst still not a large enough number to provide robust statistics it was hoped that these cases would provide confirmation of the conclusions from the 2003 study. The cases run during the course of this work are summarised in table 2. The cases listed in July are cases which were investigated as part of the preliminary field phase of the Convective Storms Initiation Project (CSIP). The 16th August event which cased notable flooding in Boscastle, Cornwall was *not* included in this study since it occurred outside the HRTM 1km domain. However due to the strong interest in this case some 12km and 4km results are included in section 4 below.

Each case consisted of four forecasts run at 3 hour intervals. The cases were run in a nominal order according to how interesting they appeared to be. The assimilation suites were run for 3 cycles (i.e. 9 hours) before the first forecast to allow the system to spin up. In addition, for comparison, a suite was run spinning up each of the four forecasts at 4km and 1km resolution from the relevant 12km T+1 analysis. The times quoted for both the assimilation and spin up runs are the analysis times so, for example a 9 UTC spin up forecast run starts at T+1 which is 10 UTC.

4. Subjective analysis of cases

(a) Overview

In this section a brief overview is given of the model performance on a case by case basis. The model results described are from the full suite including data assimilation. A comparison with the suite spinning up from 12km analyses is given in section (b) below. The cases are listed in this section in the order which they were run (the order given in table 2) and referred to in section 3.

10th July 2004

This case was the first CSIP IOP and has been analysed by Morcrette et al (2005). The point of interest in this case is a shower which initiated at around 6UTC around Newport, South Wales, and moved eastwards. A line of showers initiated downstream from this shower over the Mendips probably related to the gust front. Shortly after this several more lines of showers initiated further downstream, possibly as a result of gravity waves from the earlier showers. The end result was a number of lines of showers crossing southern England from west to east. Figure 4.1 shows the 1km domain area averaged rain rates as a function of time for all the models compared to the radar data. By this measure the 12km model very well. In contrast the 4km and 1km models, although they produce rain at the correct times, usually produce too much rain – often by around a factor of two. The 3UTC runs shows clearly the already established problem that the 4km and, to a lesser degree, 1km models have a delay in the initiation of convection. This delay is clear relative to the 12km which initiated at the correct time in this particular forecast. However in many cases (as shown later) the 12km model seems to initiate too early with the result that the delay in the 4km and 1km make them appear to be better.

Figure 4.2 shows instantaneous rain rates from the 3UTC forecasts at 10UTC when the radar image shows three distinct bands of showers. The 12km model, as would be expected, shows the convection scheme producing rain reasonably uniformly over the area of interest. The 4km model simply produces no rain in the area of the shower bands presumably as a result of the delayed initiation. The 1km model, although it does produce an area of heavy rain somewhat ahead of the most forward band in the radar, shows no sign of producing the banded structure. Overall, the 12km model produces a much more useful forecast. This behaviour may be contrasted with the behaviour of the equivalent spin up run discussed in section (b) below which does better at 1km resolution. The 6 UTC run does not do significantly better at picking up the structures although, as can be seen from figure 4.1, it does initiate the convection earlier. It shares with the 3UTC run the problem that rather than initiating bands successively downstream of existing ones the convection appears to initiate simultaneously over the whole area. The implication is that the models are missing the mesoscale structures which, in fact, caused the downstream initiation.

27th April 2004

This was a case of showers which initiated at about 15 UTC east of London. These moved west and developed rapidly to form an area of heavy rain over the Thames Valley and southwards by 18 UTC. The case was notable because there were very heavy thunderstorms with flooding over parts of London. Figure 4.3 shows the domain averaged rainfall rates compared to the radar for the four forecasts. The earlier model runs (09,12 UTC) initiated the storms too early in all three models but then they also decayed too quickly so that by 19 UTC in the 12 UTC run the area of rain had completely disappeared in the model where it was very strong by this time in reality. The fact that these errors were evident in the 12km model as well as the 4km and 1km points to them being due to an error in the large scale forcing.

At intermediate times the 4km and particularly the 1km model produced better indications of heavy rain over the London area (figure 4.4). Figure 4.5, which shows the accumulated precipitation over the first 6 hours of these forecasts, highlights that despite the better indication of heavy rain the 4km and 1km models also produce too much rain over other areas of the domain where it was not observed. The 15 UTC run handled the later stages of the storms better than the earlier runs by not decaying the storms too quickly.

3rd August

This was the most intense event in the summer 2004 period. Showers initiated over the south coast between 11-12 UTC and moved north to form three main areas of heavy rain spaced from London to Bristol at 14 UTC. These areas then merged to form a large line of rain from about 15 UTC onwards. This line appeared, from the radar imagery, to develop some cyclonic rotation before breaking into two areas of rain at around 18:30 UTC. Rain rates of over 32mm an hour were observed in the radar in some places. There was flooding in a number of places, notably in West London. The total rainfall accumulation for the period of heaviest rain (13-19 UTC) is shown in figure 4.6 from the 12UTC forecasts. The 12km model produced a good forecast of the general swathe of heavy rain extending NW from London although the area of heaviest rain is somewhat underestimated. The 4km model produced a much more fragmented rain field with areas of heavy rain scattered over most of the land area in the domain, including many areas where it was not observed. The 1km model was an improvement on this in that the general area of heavy rain, particularly west of London, was better predicted. The extent of the heaviest rain (over 32mm) was, however, seriously over predicted.

20th July

This was a case of shower initiation in a southerly flow. The showers initiated south of London and moved north and developed. The showers were followed by a more organised band of mostly light rain moving from the west later in the day. Figure 4.7 shows the models initiate the convection too early in the 6, 9 and 12 UTC runs with the initiation being earlier in the 12km model. Figure 4.8 shows that at 13 UTC when the first cell had reached a size of 20km or so across over south London all three models already had far too large areas of rain. This first cell first appears in the radar at about 12:15 in the radar mid way between London and the south coast. The models already had well developed line of convection at that time in the correct place but which had originated further south about an hour before. The implication is that the showers initiated on a correctly placed feature in the models (a convergence line is visible in the divergence field) but too early.

20th August

This case consisted of SW-NE lines of heavy rain developing and moving east. Figure 4.9 shows typical rain rate fields from the model runs compared to the radar. The 12km model arguably does the best job at suggesting the bands of rain in that it does have structure with the correct orientation (albeit with insufficient intensity). The 4km model, as would be expected, produces a small number of heavy cells which could be interpreted as being a representation of bands of rain although the cloud field also does not show bands. The 1km model fails to produce any real sign of the banded structure. The 6 hour accumulated precipitation (fig 4.10) shows the 12km model producing insufficient rain but the 4km model produces a reasonable representation with a good correspondence between the SW-NE bands in the radar data and the model. The 1km model also produces the bands although they are less clear. It is also noticeable from figure 4.9 that the 4km and 1km models produce too much rain over the sea compared to the radar. The 12km model is significantly better in this regard.

8th July

This was a case of bands of rain spiralling around a small cyclone in the Channel. An example of the instantaneous rain fields is shown in figure 4.11. Both the 4km and the 1km model suggest the bands of rain more convincingly than the 12km model (more so in the case of the 1km) although the details of the bands are wrong in both cases. The bands generally seem to be shifted to the NW in the models compared to the radar. The 6 hour accumulated precipitation (figure 4.12) shows all models producing too much rain to some degree although the 4km and 1km models are much worse in this respect. All models fail to produce the region of no rain in the northern/north eastern part of the domain which is evident in the radar data. A front, extending roughly from the Severn to the Wash in the middle of the period is correctly positioned but all the models had too much rain on it towards the NE part of the domain.

22nd July

The final case run as part of the trial involved initiation of showers in a southerly flow. The showers initiated in western Somerset and developed and moved north eventually giving a heavy thunderstorms in the midlands. Figure 4.13 shows the domain averaged rain rates against time for the four forecasts run. Once again it is clear that although the general time dependence of the rain (i.e. the initiation of rain followed by the peak rate at around 15 UTC) is captured well in some of the forecasts the overall magnitude is too large. It is also evident that there is a tendency to produce spurious and extreme maxima during the assimilation cycle – this will be discussed in section 4(b)). Figure 4.14 shows the instantaneous rain fields at 13UTC i.e. just after the showers initiated from the 6 UTC forecasts. The 4km field looks the most realistic the 12km and 1km having rain already over too great an area. The fact that the 12km initiates too early indicates that the models had, incorrectly, too much forcing too early. In this case the delay in initiation in the 4km relative to the 12km made the former appear more realistic. It is interesting to note that this scenario of convection starting too early in the models (and being offset by the delay in the 4km) seems to be quite common – it was also seen in the 10th and 20th July in this study (and has been since in some 2005 cases).

16th August

This was the Boscastle flood case and is included in this section because of its strong relevance to the severe rainfall forecasting problem. Boscastle is outside the HRTM 1km domain and so this case is not included in any of the aggregated analysis below. Figure 4.15 shows a comparison of the 12-18 UCT accumulation from the 12km and 4km models (3 UTC spin up forecast) to the radar data. The situation was a line of showers initiating around the Camel Estuary (SW along the coast from Boscastle) and developing as they moved along the coast in the SW flow. This led to a strong maximum in the rainfall accumulation around Boscastle (indicated by the circles in the figure) shown clearly by the radar data. The 12km model gave a fairly uniform low accumulated rainfall over the whole SW peninsula and no indication that high rainfall amounts might be expected. In contrast the 4km model gave a strong indication of a line of heavy showers which could have provided a useful warning.

There has been separate work on this case with 1km a model using the same configurations as those described here but on a non-standard domain (Golding et al 2005). These runs are not described here except to say that the 1km model produces less rain overall but moves the maximum in the rainfall somewhat upstream closer to the correct position.

Summary of all runs.

For a number of the above cases the 1km domain averaged rain rates against time were plotted. Figure 4.16 (solid curves) shows the same curves averaged over all forecasts of all cases (i.e. 28 forecasts in all). The 12km model produces approximately the correct amount of rain in this sense although the rates are too high by almost a factor of 2 at the start of the forecasts. The 4km and 1km models in contrast produce much more rain – more than a factor of 3 at the start of the forecast although this decreases with forecast time. By T+12 both the 4km and 12km models are producing roughly the correct amount of rain.

Similar curves plotted for the summer 2003 trials (figure 4.17) show similar results for the 12km and 4km model although the 4km model had a smaller excess and reduced faster with time. In the 2003 trial the 1km model had average rates similar to the 12km model implying that this model has become much worse in the 2004 trials. Part of the reason for this could be that there were more extreme convective cases in 2004, it is clear that the radar rain rates in figure 4.16 are lower than in figure 4.16. However is thought that the change is mostly due to the introduction of MOPS and LHN into the 1km model. This is discussed further in section 4(b).

The fact that the 4km model tends to give roughly the right average precipitation rates by the time T+12 is reached should be treated with caution. Figure 4.18 shows a similar average rain rate against time plot for the 4km model averaged over a period in April and May 2005. This model was run for 24 hours each day by spinning up from the T+1 12km analysis from the 0 UTC 12km run. The 12km model and radar are also shown for comparison. Once again the 12km model seems to produce somewhat too much rain on average and the 4km produces even more (nearly a factor of two too much at the time of peak rain rate). Now there is no sign of the rain rates becoming the same as the radar after about 12 hours although they do after about 17 hours. It is noticeable that the rain decays too quickly in the 4km and 12km models after the second peak. This is clearly very different from the 2004 rerun data. The overall rain rates are much lower and the rain was presumably not as dominated by convection. However this does show that the simplistic conclusion that the over predication of rain is likely to be due to the data assimilation rather than the model is unlikely to be the whole story.

Conclusions from Subjective analysis.

From the subjective analyses of the 7 cases presented above a number of overall conclusions can be drawn (and compared to the similar conclusions from the summer 2003 trial).

- The 12km model often underestimates the peak intensity in localised rainfall events. However it usually produces about the right amount when averaged over the whole 1km domain area. This conclusion is the same as from the 2003 trial.
- In contrast the 4km and 1km models tend to produce overestimates for the peak rain rates and similarly the domain averaged rain rates are too high

(sometimes by up to a factor 3 or so). These overestimates were worse at the start of the runs and during the assimilation cycle.

- As was concluded from the 2003 trial, the structures of the rainfall fields often look more realistic in the 4km and 1km models although the details are often wrong. The representation of convection explicitly rather than by the convection scheme is expected to lead to these improvements. However as will be mentioned in section 4(b) there is evidence that having assimilation in the high resolution models can disrupt the generation of realistic structure.
- If the accumulated rainfall is the main quantity of interest it is often the case that the 4km and 1km models give a better indication of the potential for localised high accumulations. On some occasions, however, they may also be more prone to giving false indications of extreme events.

(b) Discussion of assimilation issues and comparison to spin up runs.

Since the assimilation in the 4km and 1km models is still highly experimental a key benchmark is to compare their performance against 4km and 1km models simply spinning up from the corresponding 12km analysis. It is accepted that spin up forecasts will contain no high resolution detail for the first few hours (which is why ultimately it will be essential to have assimilation in these models for short range forecasts) but it might be expected that after that period they would provide a useful point of comparison.

The dotted lines in figure 4.16 show the averaged rain rates for the spin up versions of the 4km and 1km models. As was observed during the 2003 trials (Lean et al 2005) these curves start off very low (the rain in the 12km model is produced by the convection scheme whereas the 4km and 1km produce most rain explicitly). As the forecast progresses the rain rate increases rapidly as the convection spins up and overshoots reaching a peak at about T+6 before gradually decaying to reach the correct value at around T+11. Bearing in mind that the cases contributing to these curves are all convective it was shown from the 2003 trials that this behaviour resulted from CAPE building up unrealistically at the start of the forecast and then producing too much rain later once the convection had spun up.

In a number of cases it is evident that runs including assimilation are less likely to produce correct structures in the rainfall fields – particularly when these structures are on a fairly small scale. An example is shown in figure 4.19 which shows some fields from the 10th July case from the 3 UTC spin up run. This should be compared with figure 4.2 which shows the corresponding fields from the assimilation suite. The lines of showers advancing east can be seen clearly in the radar images but are not produced by either the 12km or 4km model. The 1km model with assimilation fails to produce the lines but the spin up forecast does show several separate lines. This latter forecast is not perfect, however – there are signs that the lines would prefer to have a NE-SW orientation. The evolution of this forecast (starting with a single shower over Newport in South Wales) is also closer to the radar data than the assimilation forecast. Figure 4.20 illustrates this showing the evolution of the initial shower into two in the Bristol area. Although the model looks superficially quite similar to the radar it appears that the shower may split in two in the model rather than initiate a new one down stream. In contrast in the same forecast from the assimilation suite (not illustrated) the initial shower has completely died out by this time so there is no rain at all in the vicinity of Bristol.

A second example is shown in figure 4.21 from the 3rd August case. The fields shown are 3 hours after the analysis time. In the 4km model in particular shows that the line of storms has been broken up into a number of areas of rain in different parts of the domain. The effect is less marked in the 1km model but there is still the appearance of spurious rain to the south.

A number of examples have been seen where the assimilation appears to introduce too much rain into the model. It is clear that this is true on average from figure 4.16. It could simply be a result of the fact that the extra noise introduced tends to initiate more convection. It is, however, thought that this problem often originates from the MOPS/LHN scheme. An example is shown in Figure 4.22 which compares T+4 forecast 4km and 1km forecasts with the corresponding T+1 forecasts. The assimilation has had the effect of adding an area of very heavy rain extending SE into the Channel from the Isle of Wight area. The radar data shows that this rain is not spurious – there is an area of rain in this position which was not in the background for that assimilation cycle. It would seem that the LHN is correctly putting in the rain but giving it a much too high intensity. Work is ongoing to address this problem in the MOPS/LHN scheme. A second characteristic of this excess rain introduced during the assimilation cycle is that it decays rapidly. This can be seen in figure 4.16 and also in figure 4.23 which shows the same models as figure 4.22 two house later.

In summary it can be concluded that assimilation is essential for short range high resolution forecasts because of the unrealistic rain rates seen for at least six hours after starting the forecast from a 12km analysis. There are also problems with assimilation at high resolution with too much rain being generated and some high resolution structures being disrupted. These are the subject of ongoing work.

5. Statistical analysis

In this section the statistical analysis of the results of the 2004 trial runs are presented. The statistics presented are the same as those for the 2003 trial (Lean et al 2005) and the reader is referred to this for detail of the statistics and detailed discussion of their pros and cons. The statistics allow convenient summaries to be produced taking into account all runs (rather than just selected runs/times shown in the previous section). Both the assimilation and the spin up runs are included.

(a) Surface/single level RMS stats

These were obtained using the Met Office Area based Verification System which was used to generate statistics comparing surface observations to the model. The data discussed here were calculated by taking data on the model grids and interpolating the nearest four data points to the observation. This was done for all available observations within the 1km domain area. The temperature observations were not corrected for height. The statistics were calculated every hour for each forecast and then all the forecasts averaged together as a function of forecast time (i.e. time after the analysis time). Due to technical problems only data from the first 6 of the 7 cases are included – the 22nd July case is not included.

Fig 5.1 shows the RMS and bias data for the Mean Sea Level (MSL) Pressure and for Temperature. The MSL pressure bias data shows both the 1km models doing better than the 12km one. The 4km models are broadly similar to the 12km. The RMS errors are similar for all the models although the 4km and 1km models tend to be somewhat higher presumably simply as a result of having more short range structure. It is noticeable that the assimilation suites have slightly higher errors at the start of the forecasts. The spin up runs would be expected to have a lower RMS error at the start due to the relatively smoother fields from the lower resolution model. For the temperatures the bias curves show that the high resolution models generally have a lower bias than the 12km model although the error bars are large in this case. The spin up models have a larger bias at the start of the forecast and the 1km model with assimilation has a negative bias, the opposite sign to the other models, but of about the same magnitude. The RMS temperature errors are similar to but slightly larger than the 12km errors.

Fig 5.2 shows the same verification data for cloud cover and visibility. The cloud cover bias shows that the 4km assimilation models is similar to the 12km although the 1km appears significantly better. The spin up models appear to have too little cloud at the starts of the forecasts but recover later – this is probably what would be expected from spinning up from a 12km analysis if there was a lot of convective cloud in the 12km model. The RMS errors for cloud show that, once again, the 4km and 1km models are close to but slightly worse than the 12km. The visibility stats show a wide spread in the bias between the models – the 4km assimilation model is the worst. It is expected in convective situations that the visibility signal will be dominated by the reduction of visibility in rain. The RMS errors in the visibility show that the 4km and 1km assimilation models are again somewhat worse than the 12km particularly at the start of the forecast. The spin up 1km and 4km runs have RMS errors very close to the 12km model.

In conclusion, the RMS errors tend to be similar to but slightly worse in the 4km and 1km than those seen in the 12km model but this is to be expected from the high resolution models having more short range structure. The bias results are more complicated and require further study.

(b) Precipitation Statistics.

The precipitation statistics presented here are the same as those described in some detail in Lean et al (2005). In summary these statistics use a technique developed by Roberts (2004) to take into account varying degrees of spatial errors in the forecasts. Spatial errors may be expected in high resolution forecasts both as a result of large scale errors in the driving model (for example fronts etc being in the wrong place) and as a result of the lack of predictability of small scale features (individual shower cells being wrongly positioned).

In order to calculate these statistics the rainfall accumulation fields from the models are interpolated or aggregated onto the 5km grid on which the radar data is available. To enable comparison between all three models this can only be done over the area of the 1km model. For a given rainfall threshold the model data is used to generate probabilities by counting the number of (5km) gridpoints within a square of a given size around the gridpoint in question. The size of the square is given by the sampling radius and varies from 1 (which represents a point by point comparison) up to the point where a significant proportion of the model domain is covered. The radar data is converted into probabilities in the same way. A skill score is calculated which is the root mean square error in the probabilities normalised by the uncorrelated value. The value ranges from 0 for a forecast with no skill to 1 for a perfect one.

Figures 5.3 and 5.4 show this skill score for each of the cases considered here as a function of sampling radius. These scores were calculated for an accumulated rainfall threshold of 4mm in the 6 hour period of each forecast. The value on the left hand side of the graph gives the skill score for point by point verification and has the lowest value. Moving to the right allows for progressively larger spatial errors and the score rises. At the extreme right of the graph there is very little spatial information and the score levels off at a value which largely represents the difference in area averaged accumulation (i.e. the bias). Each of the curves shown is the result of aggregating together the four forecasts of the case in question.

The curves are very different between the seven cases highlighting once again that it would be necessary to run a large number of cases to get a statistically significant sample. This is despite the fact that these cases all being convective are broadly of the same type. The cases in which the high resolution models do worst relative to the 12km (the 10th 20th and 22nd July) are all cases in which small scale convection initiates from nothing during the period covered by the model runs. The case with least spread between the skill scores from the different models was the 8th July in which the precipitation pattern was dominated by bands of rain surrounding a small cyclone. Although this is the sort of situation where the high resolution models might be expected to add little the high resolution models did have slightly higher scores.

In general the high resolution models are often producing rather poor scores compared to the 12km model and the assimilation runs seem to be systematically worse. This conclusion is confirmed in figure 5.5 which shows the same scores aggregated over all runs of all cases. The spin up 4km run is only very slightly worse than the 12km (indeed out to a sampling radius of about 80km the curves are indistinguishable). The 1km spin up is somewhat worse and the assimilation 4km is slightly worse again. The 1km with assimilation gives the worse score of all by this measure. Figure 5.6 shows the scores again aggregated over all the runs for a more extreme threshold of 16mm in 6 hours. The 12km is much better than the high resolution models by this measure with longer sampling radii although the spin up models are roughly as good out to around 50km. The assimilation models are worse at all sampling radii.

Taking these results in combination with the domain averaged rain rate results (fig 4.14) it is likely that these skill scores are dominated by the bias in overall rain rate. In order to separate this effect from the spatial accuracy the curves have been recalculated using relative thresholds (this replaces but is equivalent to the use of Briar skill scores with relative thresholds used in the 2003 trial). Rather than using a fixed threshold to calculate the scores (e.g. 4mm in 6 hours) the threshold is adjusted for the radar and each model separately so that in each case the same proportion of points are above the threshold. In this way any bias in the overall rainfall amount is removed and the spatial quality of the forecasts are compared.

Figures 5.7 and 5.8 show the skill scores for each case using thresholds set such that only the top 1% of gridpoints are above. Figure 5.9 shows this score aggregated over all the cases. The high resolution models now are doing better relative to the 12km model. Although at longer sampling radii the spin up models are still marginally better than the assimilation ones it is noticeable that the assimilation 1km has the best score for radii between around 40 and 125km. Both the 1km models do marginally better than their corresponding 4km models in the range of sampling radii out to around 100km. Figure 5.10 shows the same curves (again aggregated over all cases curves) for a relative threshold of the top 10%. Since there are more points above the threshold the curves tend to be smoother and in fact very nearly all lie on top of each other although the 12km is marginally the worst for all but radii below 30km. These results show that for a 6 hour accumulation the high resolution models do better for relative thresholds i.e. when the effect of the overall overproduction of rain in the models is removed.

It is also informative to consider the statistics calculated for hourly accumulations and these are shown for thresholds of 1 and 4mm in an hour in figure 5.11. Since the accumulation period is shorter the statistics tend to vary a good deal from hour to hour due to the fact there are often only a few model points over the threshold in question.

This means that the statistics are less significant. For this reason only the aggregated stats over all cases are shown. The fraction skill score is now plotted as a function of forecast time for a particular sampling radius. A sampling radius of 50km has been chosen since this is approximately mid way between 0 and the point where the curves plotted against radius level off.

The 1mm in an hour skill scores show that after T+2 the high resolution models are better than the 12km. The assimilation models tend to do slightly better than the spin up ones. The high resolution spin up models are expected to do worse at the start of the forecast. The models including assimilation often have spurious rain at the start of the run (as discussed in section 4(b)) and the spin up runs are still spinning up. It is particularly noticeable that the high resolution models have scores which are generally increasing with time compared to the 12km which falls off. This may be because with the four forecasts being run at intervals through the period of convection in each case the end of the forecasts contain more of the decaying phase of convection where the 12km model fails to pick up organisation correctly.

It is reasonable to ask why the high resolution models appear to do well in these hourly scores whereas in the roughly equivalent 4mm in 6 hours scores (figure 5.5) the 12km model does best. It is likely that this is largely due to the extra time dependent information incorporated in the hourly scores. The 12km model tends to miss the organisation of showers and produce moderate amounts of convective rain fairly uniformly over the whole area of showers. When the rainfall is accumulated over 6 hours the movement of the individual showers or bands of rain tends to produce a more uniform field (which becomes even more uniform when long sampling radii are considered). An example of this, from the 20th August case, is shown in figure 4.9. In contrast if only hourly accumulations are considered the rain is likely to be in much more localised areas in the radar which will compare better to the high resolution models than the relatively uniform 12km field.

The second frame in figure 5.11 shows the hourly skill scores for the higher threshold of 4mm in 1 hour. In this case the 12km remains better for longer and becomes worse than the assimilation high resolution models at around T+3. The spin up models are worse and only become better than the 12km at T+5.

As with the 6 hour scores it is likely that these scores are influenced by the overall over prediction of rain in the high resolution models. Figure 5.12 shows the hourly skill scores recalculated for relative thresholds. These statistics also show a clear signal that the 4km and 1km models are doing better than the 12km model with the exception of the start of the forecasts. By these measures there is very little to choose between the spin up and assimilation forecasts.

The sampling radius of 50km chosen for the hourly skill score plots is rather large compared to the gridlength of the high resolution models. The same curves have also been calculated for a smaller radius of 30km (not shown). The general trends with forecast time and the relative position of the models is unchanged. The main difference is that the overall values of the skill scores are lower as would be expected from the plots against radius. If we equate spatial error to temporal error via a typical velocity of 10m/s then 50km corresponds to a time error of between 1-2 hours which is approximately the predictability time of small showers. This implies that a sampling radius of 50km is appropriate for looking at showery situations.

The overall conclusions from the precipitation statistics are listed below:

- There is a great deal of variability between the seven cases run which implies that many more cases would need to be considered to obtain statistically valid results. The following conclusions on aggregated scores need to be treated with caution as a result.

- The 6 hour skill scores with fixed thresholds show the 12km model is the best followed by the spin up 4km and 1km followed by the assimilation 4km and 1km. However the relatively low scores for the high resolution models is expected to be mostly due to the overall excess of rain.
- The 6 hour skill scores with relative thresholds (i.e. taking account only of the spatial accuracy) imply that the 4km and 1km models are better than the 12km model. In this case there is no clear signal between the assimilation and spin up models or between the 1km and 4km.
- The skill scores calculated from hourly accumulations imply that the 4km and 1km models (both spin up and assimilation) are better than the 12km model with the exception of the start of the forecasts.

6 Discussion and suggestions for future work.

(a) Comparison to 2003 trial.

The conclusions from the runs presented here are similar to those obtained from the 2003 trial (Lean et al 2005) in the sense that the high resolution models can lead to more realistic structures in the precipitation fields and more realistic peak rainfall rates. In the 2004 trial it is more clear that there is a problem in the 4km and 1km models with over prediction of rain. The 4km model had too much rain in a domain averaged sense by a factor of less than 2 at the start of the forecasts but was down to about the right value by T+4 in the averaged 2003 results. In the current trial the 4km model has around a factor of 3 too much rain at the start and only becomes close to the correct value after about 12 hours. This increase in an already existing bias may simply be a result of the fact that there were more cases of extreme convection included this year. The 12km model also has a larger positive bias in the average rain rate in 2004 compared to 2003. A second possible factor is the fact that in the present trial the 4km model is running with prognostic rain (diagnostic rain was used in the 2003 trials). It has been found by DWD that prognostic rain and ice interacted unfavourably with LHN resulting in excessive rain rates. However a test carried out on the UM at 4km directly comparing diagnostic and prognostic rain version implied that running with diagnostic rain gave only a small decrease in the rain rates.

In contrast the 1km model appears to have changed from having approximately the same amount of rain as the 12km (except for the very start of the forecast) in 2003 to having even more rain than the 4km model in the 2004 trial. Despite the limited number of cases this seems a large change, certainly larger than the change in the 4km model. The points made above about the 4km model are also true here but, in addition there are a two other important differences in the 1km model from the 2003 runs. Probably the most important difference is the incorporation of MOPS/LHN in the 1km model in the 2004 trial. Figure 4.14 shows clearly that the over prediction is worse at the start of the forecast so it seems likely that problems with MOPS/LHN is at least partly responsible for this problem. Secondly the dry static adjustment has been switched off in the 1km model. In the (spin up) tests which were carried out when this change was made it was found that the removal of the dry static adjustment appeared to result in less rain overall rather than more. Despite this it is still possible that due to interactions with the assimilation that this change is relevant to the observed change in the model.

(b) Future Work

It is clear that one focus of the work in the immediate future must be towards reducing the over prediction of rain in the 4km and 1km models. A second must be to move towards a configuration which is closer to that required for a nowcasting system.

The over prediction problem must be addressed both with changes to the model physics and with the assimilation system (primarily LHN/MOPS). Preliminary work by Richard Forbes has shown that the rain rates can be reduced significantly if changes are made to various microphysical parameters such as the ice fall speeds and deposition rates. These experiments are valid because they are within the range of uncertainty of the parameters in question. Another aspect is that the version of the model used in this trial made use of specific moisture quantities in the dynamics and this neglects the water loading term (the change in buoyancy due to water vapour). Preliminary trials both with this term added and using an alternative version of the model with the mixing ratio version of the dynamics (which automatically includes water loading) imply that that this term can significantly reduce the vertical velocities and rain rates.

It is also possible to reduce the overall rain rates by changes to the diffusion and convection schemes and these will also need to be investigated. It is expected that we will test the use of the targeted diffusion scheme. In this context it should be noted that the 4km model in particular currently sometimes suffers from very high values of rainfall rates (e.g. up to 200 mm/hr) and vertical velocities (up to 20 m/s) on individual gridpoints. Changes in diffusion are required to reduce the values on these points.

In the light of the evidence that has been presented that the over prediction of rain is related to the LHN/MOPS scheme continuing work will be carried out to tune this scheme, particularly for the 1km model. In the near future tests will be made of more frequent data input (every 15 minutes for the rain data and every hour for the cloud). It is expected that this, on its own, may help with the problem since currently a features will be nudged into the model for too long in the same place which may lead to high rain rates. Experiments may also be carried out in the 1km model with data with spatial resolution data greater than the current 15km (the source data is at 5km resolution).

The 4km model is currently being implemented in the operational suite. A simple 36 hour spin up run running once a day from the T+1 12km analysis from the 0 UTC run is currently running. This model uses a larger domain than that used in the models described in the present report which covers the whole UK area. Later in 2005 a full assimilation system at 4km will be implemented in order to provide a number of 4km forecasts each day out to T+36.

The focus of the HRTM follow on project will change to work on the 1km model and, in particular, setting up the 1km model for use as a nowcasting system – i.e. to provide forecasts out to about T+6. Most of the substantial work to bring this about is likely to be on the data assimilation system. It will be necessary to move to hourly cycling (rather than the current three hour cycles) at least for assimilation of rain and cloud.

7. Conclusions

A second set of summer cases has been run with the HRTM suite using cases from 2004. Compared to the 2003 trials a number of changes have been made to both the model and assimilation systems. The most important of these were the inclusion of LHN/MOPS in the 1km model and the inclusion of prognostic rain in the 4km and 1km models.

Subjective analysis of the precipitation fields leads to similar conclusions to previously, namely that the 4km and 1km models can give more realistic precipitation structures largely as a result of running without a convection scheme. In this trial it is very obvious that the 4km and 1km models tend to produce too much rain overall – sometimes by more than a factor of two. This is thought to be a combination of problems with the data assimilation (particularly the LHN/MOPS) and the model physics.

The objective precipitation scores for 6 hour accumulations show that the 12km model generally appears to do better for fixed accumulation thresholds. This result is dominated by the over prediction of rain. If a relative threshold is used instead so only the spatial distribution is tested the 4km and 1km models generally produce better scores than the 12km. The scores for 1 hour accumulations have also been considered as a function of forecast time for fixed sampling radii of either 30 or 50km. The 4km and 1km models come out best for both the absolute and relative thresholds with the exception of short forecast times (where the high resolution forecasts are dominated by either the assimilation or the spin up). The cross over time is at T+1 to T+2 (although it is later at T+3 for the higher threshold of 4mm in an hour).

The scores imply that the 4km and 1km models do better than the 12km for spatial distribution of rainfall and for verification over short periods. In this trial there is not much statistical evidence that the 1km model is better than the 4km but this is likely due to the general over prediction of rainfall.

The basis for carrying out this work was the expectation that the 4km and 1km models should be able to provide improvements to the forecast compared to the 12km model and that the 1km should also be significantly better than the 4km. Despite the over prediction of rain we have evidence that the high resolution models do provide these benefits for short time periods (1 hour accumulations) and spatial distribution of rainfall. In addition the models clearly have the potential to provide improvements for forecasts of longer period rainfall accumulations. There is every reason to expect that future work to resolve the over prediction issue will realise this potential.

References

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

Golding et al (2005) “Boscastle and North Cornwall Post Flood Event Study - Meteorological Analysis of the Conditions Leading to Flooding on 16 August 2004” edited by Brian Golding FR tech report 459

Lean H.W. et. al. 2005 “The Summer 2003 Reruns with the High Resolution Trial Model” FR Tech Report 457/JCMM Report 151.

Morecrette, C.J., Browning, K.A., Blyth, A.M., Bozier K.E., Clark, P.A., Ladd, D., Norton E.G., Pavelin E. 2005 “Initiation of multiple bands of convective showers over southern Britain”. To be submitted to Q J R. Meterol. Soc.

Roberts, N.M., 2003, “The impact of a change to the use of the convection scheme to high resolution simulations of convective events”. FR tech report 407/JCMM report 142.

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 modelling project. FR tech report 432/JCMM report 146.

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 FR Tech Report 455/JCMM Report 150

	12km	**4km**	**1km**
Horiz Domain			
Approximate Gridlength (km)	12	4	1
Gridlength (deg)	0.11	0.036	0.009
lat BLC		-5.45	-2.630
lon BLC		356.6	359.3
grid size	146x182	190x190	300x300
pole lat	37.5	37.5	37.5
pole lon	177.5	177.5	177.5
Vertical Levels			
No levels	38	38	76
Top of model (m)	approx. 40000	approx. 40000	approx. 40000
Other definition	G3	G3	2G3
BL levels	13	13	26
ozone levels	11	24	48
LBC issues			
Driving model	Global	12km	4km
Rimwidth	8	8	8
Time frequency	60 min	30 min	15 min (should be 5 min)
Aerosol Boundary values from	UKmes boundary model	12km	4km
Timings			
Run length	8hr (spin up) 3 hr cycle + 6 hr f/c (assimilation)	8hr (spin up) 3 hr cycle + 6 hr f/c (assimilation)	8hr (spin up) 3 hr cycle + 6 hr f/c (assimilation)
Timestep	5 mins	100s	30s
Radiation timestep	60 min	15 min	5 min

Table 1 Summary of settings in models (continued on next page)

Parametrisations			
Convection	4A scheme (previously known as CMODS). CAPE closure timescale 1800s	3C scheme with CAPE dependent CAPE closure settings 1 for 4km i.e. function with t=1200s, c=0.5	No convection scheme
Microphysics	3B dual phase including iterative melting.	3C dual phase No iterative melting but with prognostic rain.	3C dual phase. No iterative melting but with prognostic rain.
Radiation			
Gravity Wave Drag	On	Off	Off
Boundary Layer	13 levels	13 levels	26 levels
Other			
(Max del-4 diffn for stability		3.3e4	4.2e3)
Horizontal Diffusion	None	del-4, 8 min i.e. 1.14e4 (5.1/6.1) 8.53e3 (5.1.1/6.1.1)	del-4, 8 tsteps i.e. 1.43e3
Solver tolerance (absolute)		1.0e-3	1.0e-3
alphas		(off centre to 0.7 as prev)	
RHcrit	oper values, mostly 0.85	oper values, mostly 0.85	oper values, mostly 0.85
Boundary Layer weights		All 2.0	All 2.0
Targetted Diffusion	On	Off	Off

Table 1 (continued).

Case No	**Date**	**Model Runs**	**Description**
1	10th July	3,6,9,12	Gust fronts initiating showers downstream from initial development over S Wales at 06 UTC.
2	27th April	9,12,15,18	Heavy storms initiating over London at about 15:30 UTC and subsequently moving west.
3	3rd August	6,9,12,15	Showers initiating along S coast at around 12:30 UTC moved N and developed into line of V heavy rain with lightning and hail by around 15UTC.
4	20th July	6,9,12,15	Showers initiated at around 13 UTC in sourtherly flow.
5	20th August	3,6,9,12	Bands of heavy showers moving east.
6	8th July	3,6,9,12	Bands of rain around a cyclone in the Channel.
7	22nd July	6,9,12,15	Showers initiate around 13 UTC over Somerset subsequently move north and develop.

Table 2. Cases run in summer 2004 trial. The case number was a nominal priority and represents the order in which the cases were run.

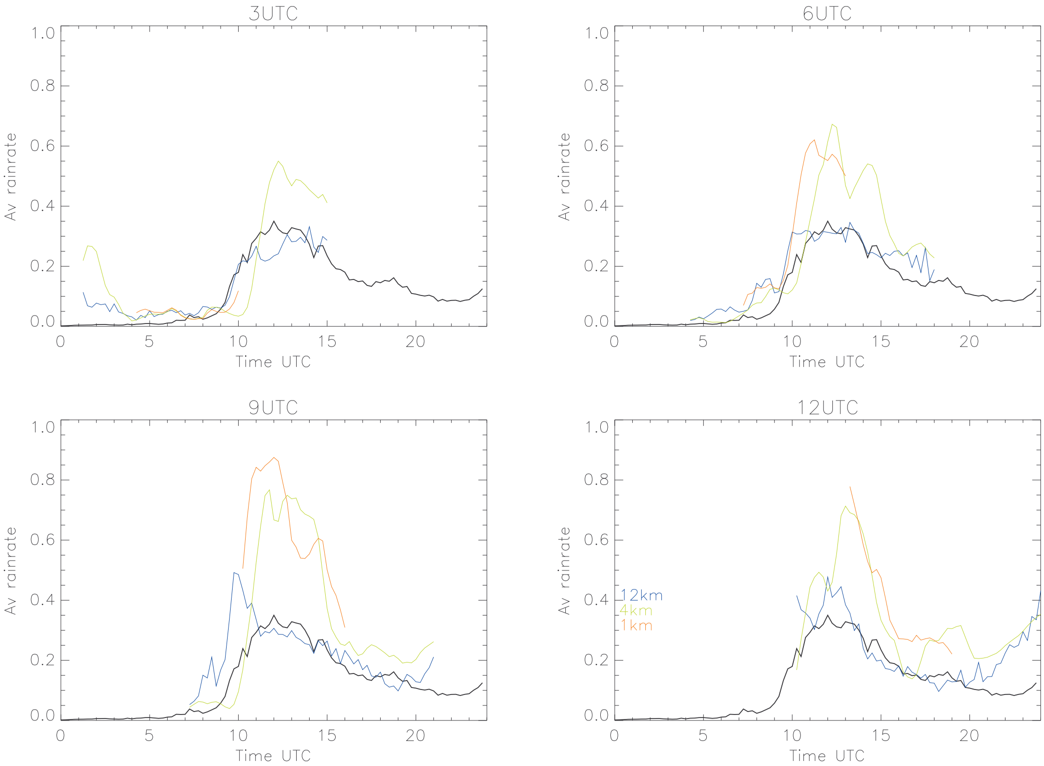


Figure 4.1 1km model area averaged rain rates for 10th July 2004 showing radar data (thick black line), 12km model (blue), 4km (green) and 1km (red) from each of the four forecasts. The 1km data runs only from T+1 to T+7 in each case whereas the 12km and 4km data runs from T-2 to T+12.

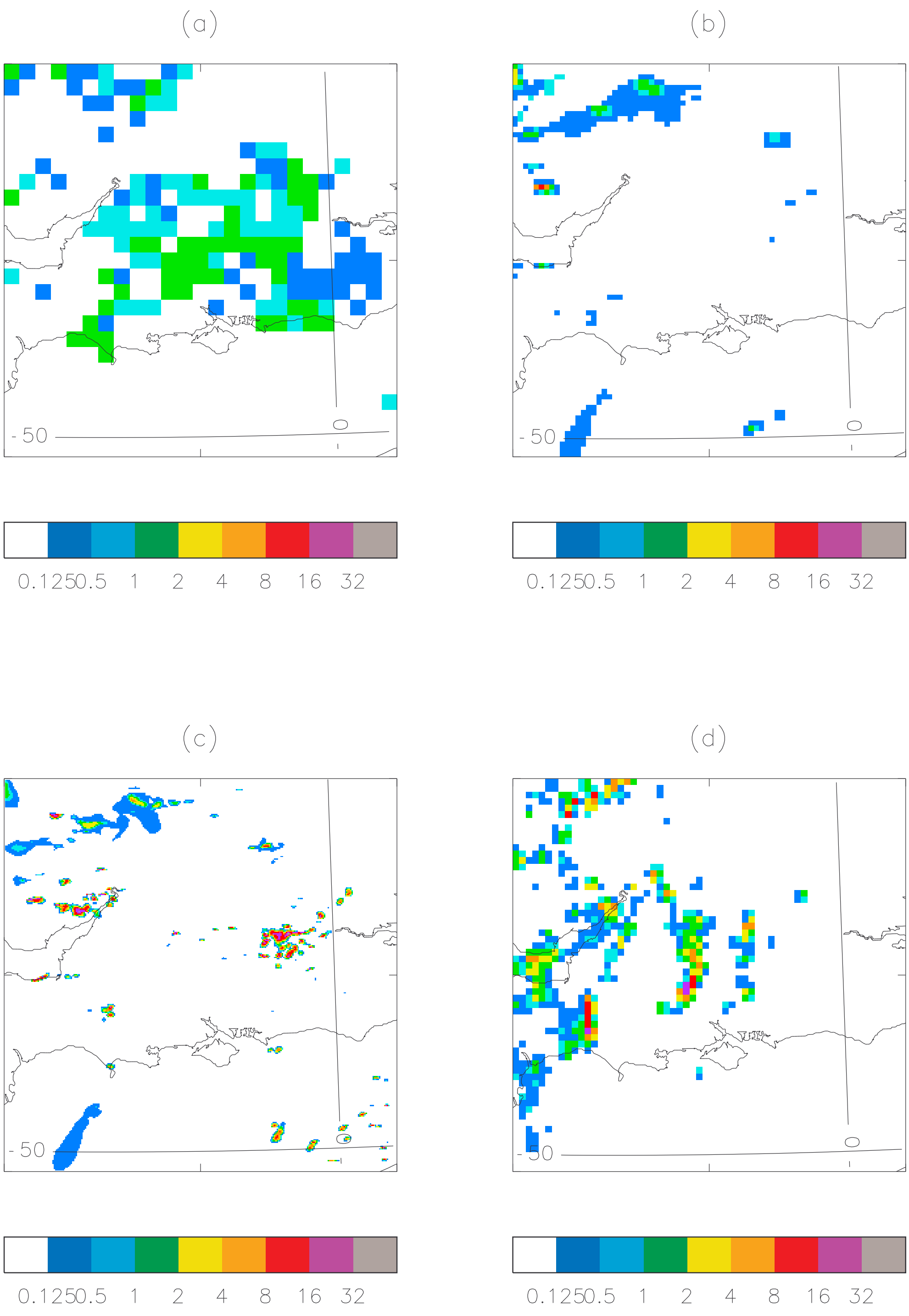


Figure 4.2. Instantaneous rain rates at 10 UTC on 10th July 2004 from 3 UTC model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data.

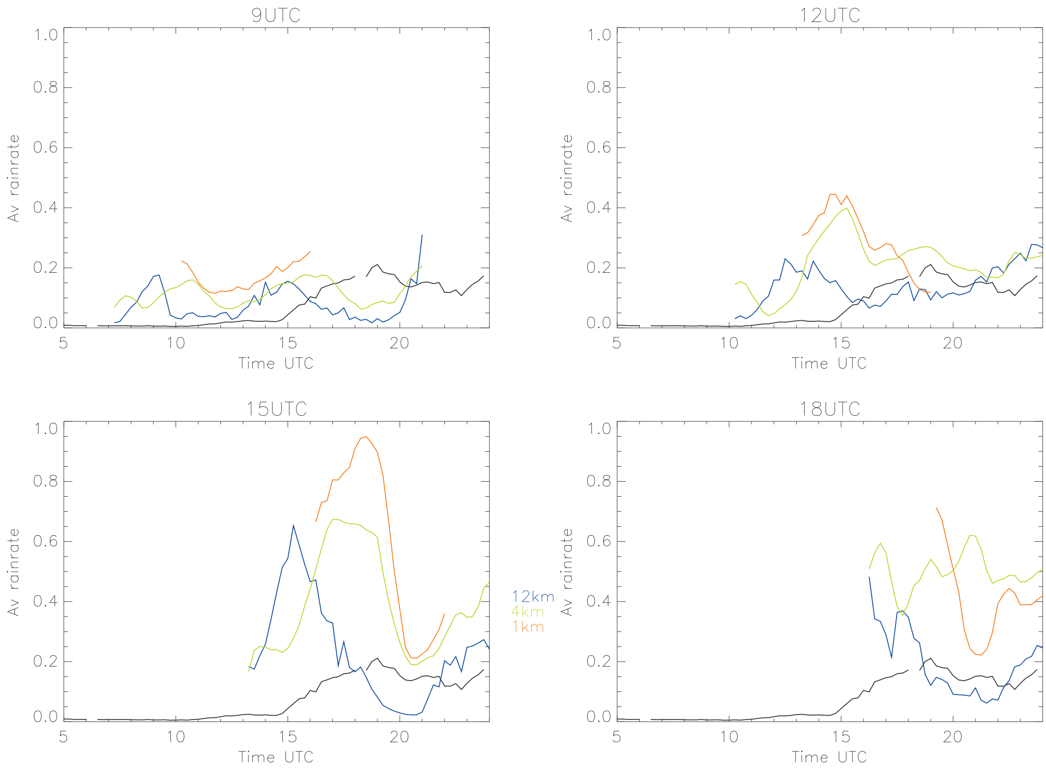


Figure 4.3 1km model area averaged rain rates for 27th April 2004 showing radar data (thick black line), 12km model (blue), 4km (green) and 1km (red) from each of the four forecasts. The 1km data runs only from T+1 to T+7 in each case whereas the 12km and 4km data runs from T-2 to T+12.

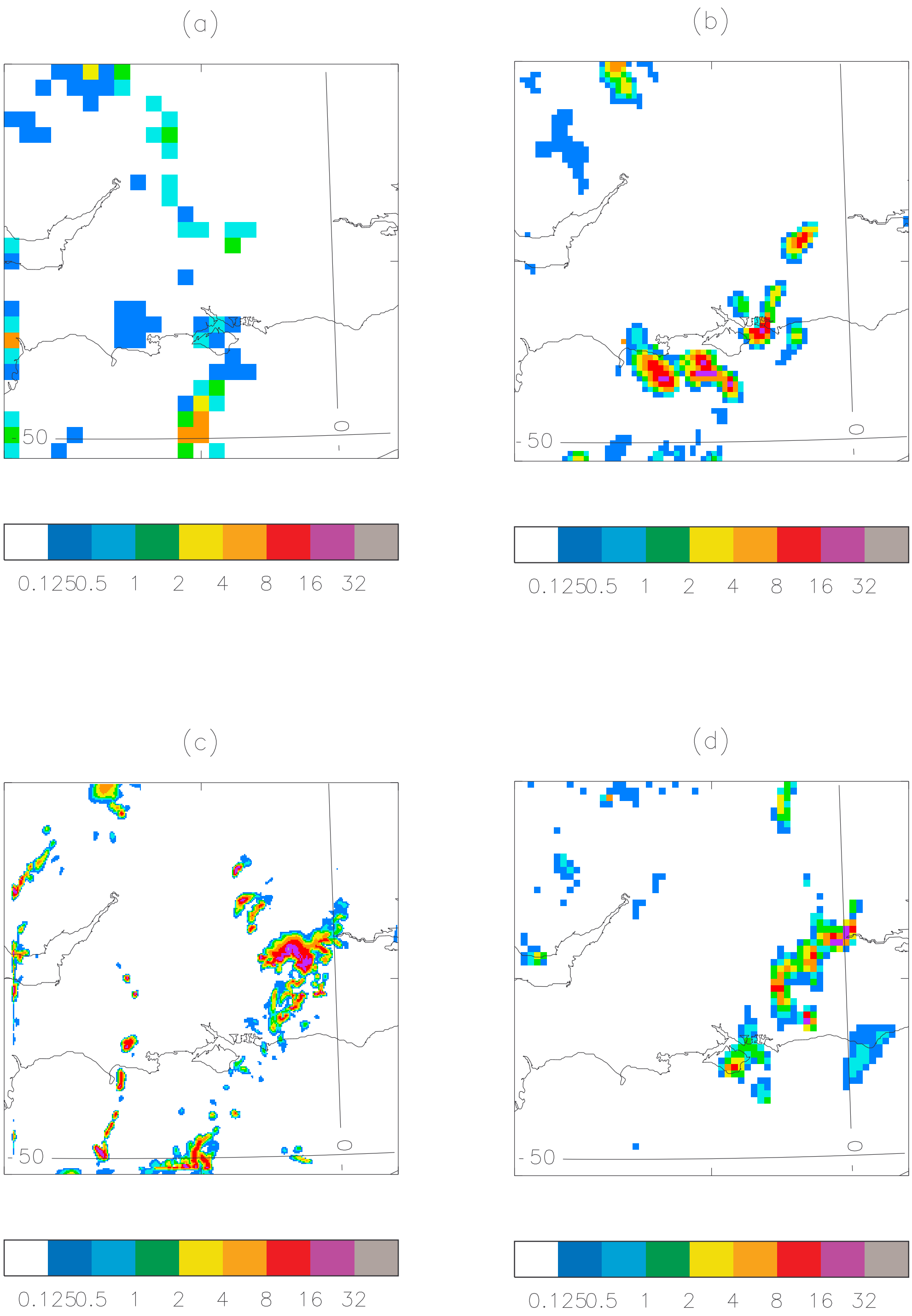


Figure 4.2. Instantaneous rain rates at 17 UTC on 27th April 2004 from 12 UTC model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data.

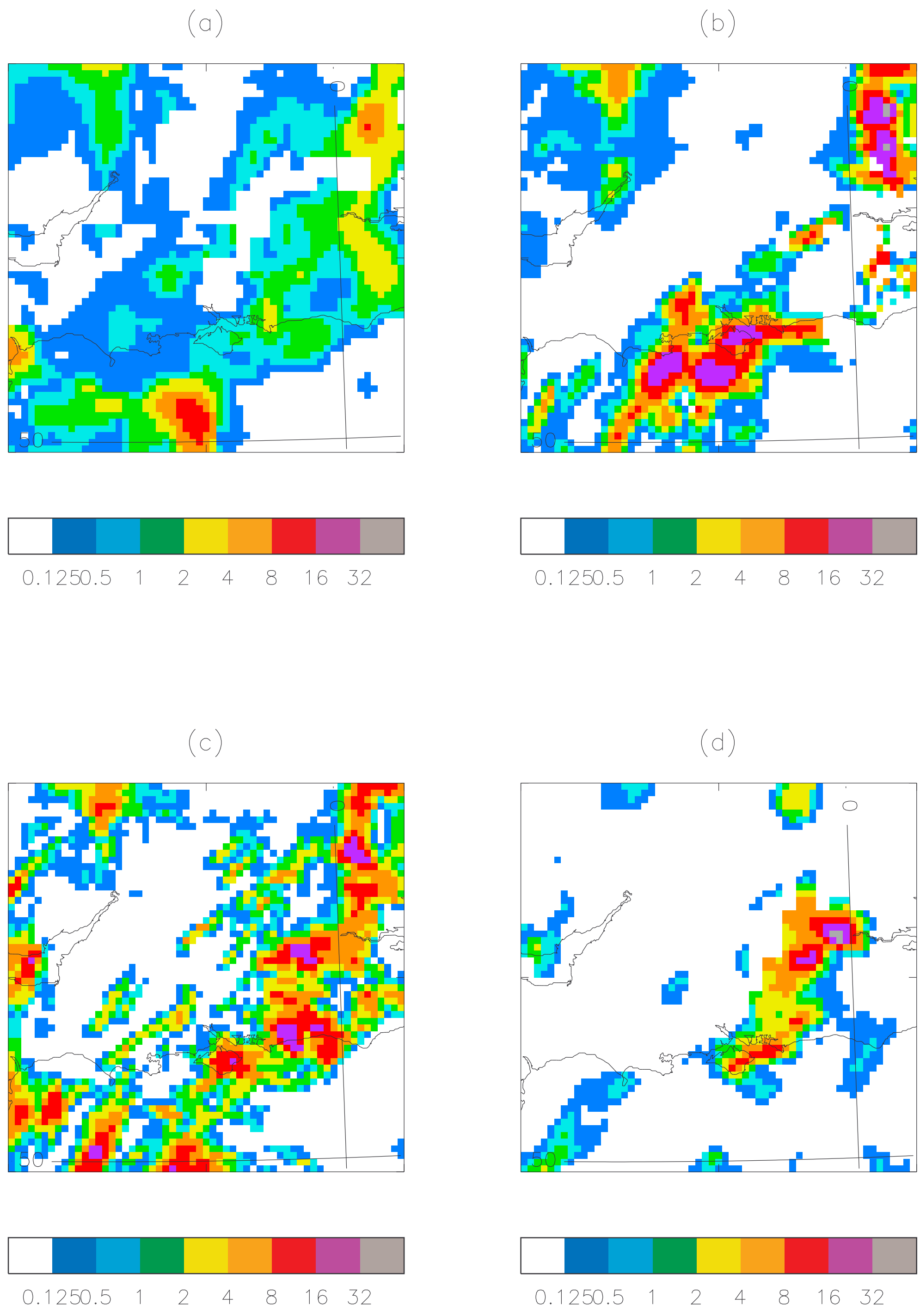


Figure 4.5 27th April 2004 13-19 UTC precipitation accumulation from QM12 model runs. (a) is 12km, (b) 4km and (c) 1km. (d) is equivalent radar data. For clean comparison all fields have been interpolated/aggregated onto a 5km grid.

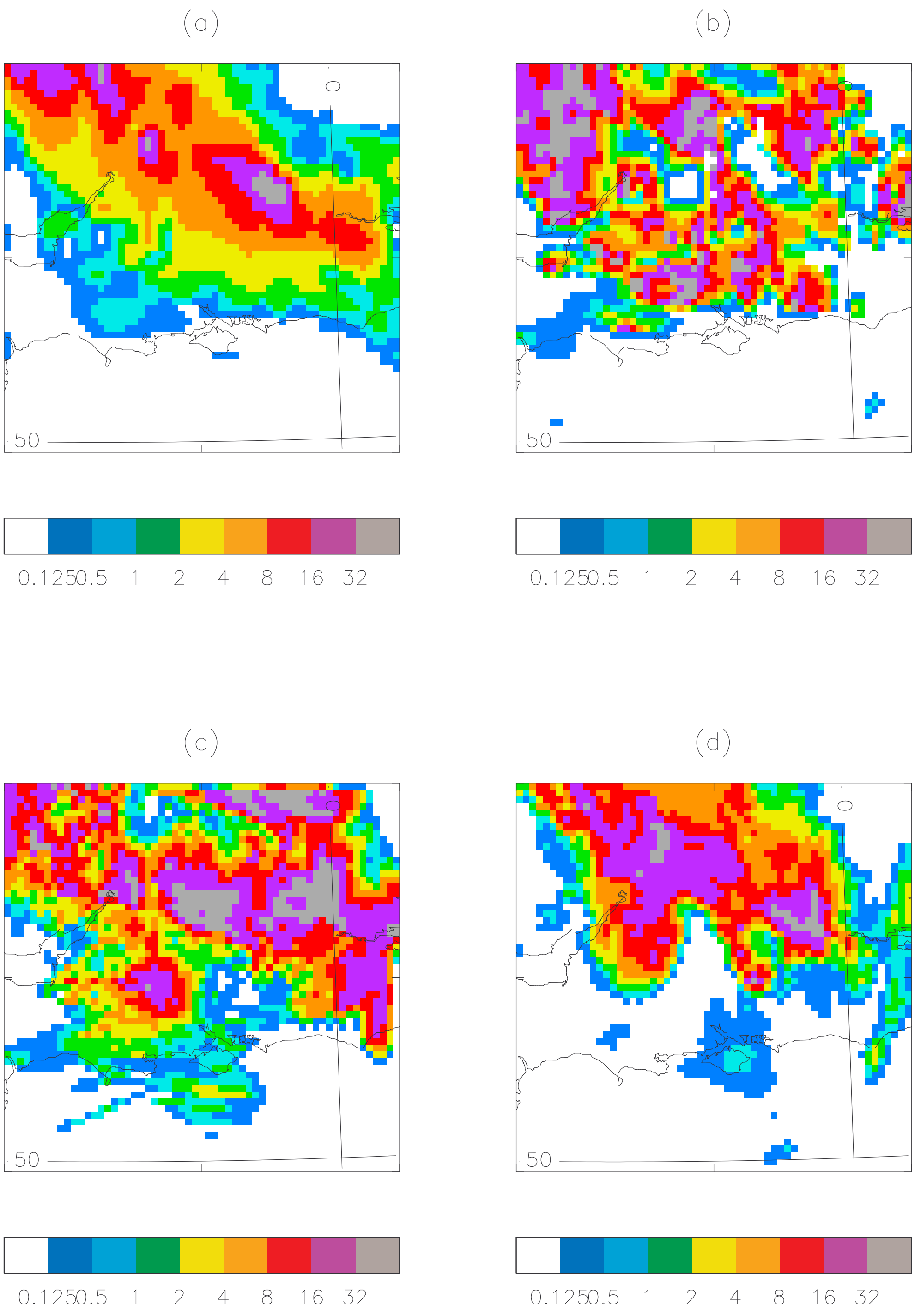


Figure 4.6 3rd August 2004 13-19 UTC precipitation accumulation from 12 UTC model runs. (a) is 12km, (b) 4km and (c) 1km. (d) is equivalent radar data. For clean comparison all fields have been interpolated/aggregated onto a 5km grid.

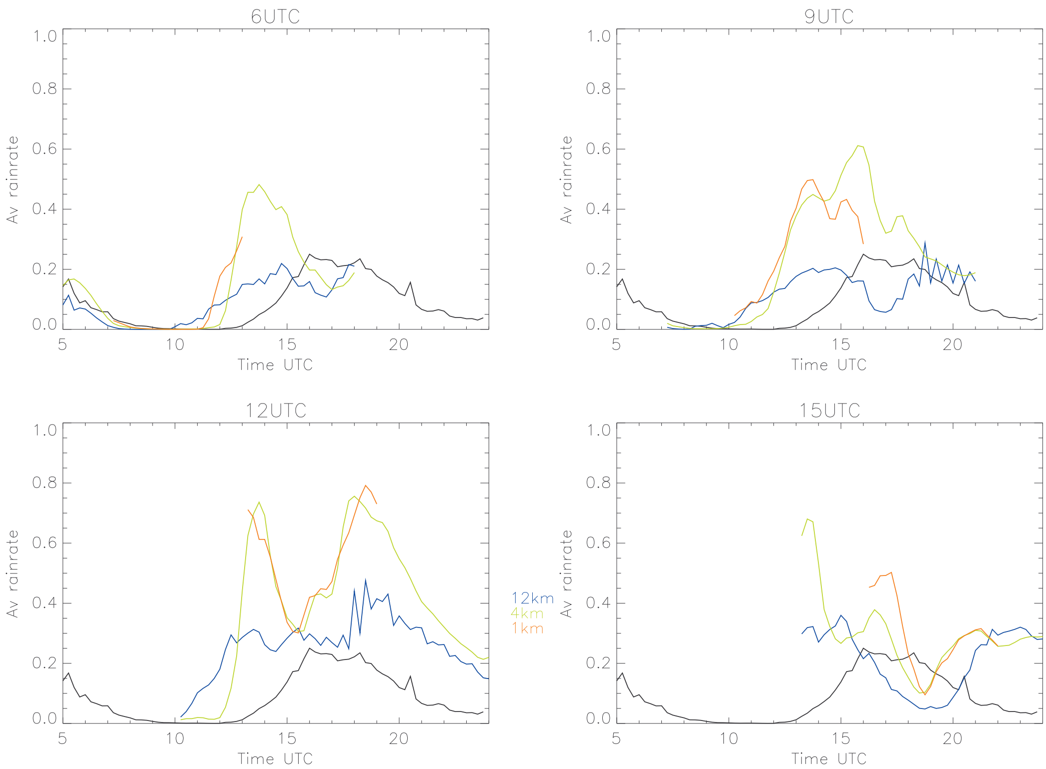


Figure 4.7 1km model area averaged rain rates for 20th July 2004 showing radar data (thick black line), 12km model (blue), 4km (green) and 1km (red) from each of the four forecasts. The 1km data runs only from T+1 to T+7 in each case whereas the 12km and 4km data runs from T-2 to T+12 (in the latter the period T-2 to T+1 is the assimilation cycle).

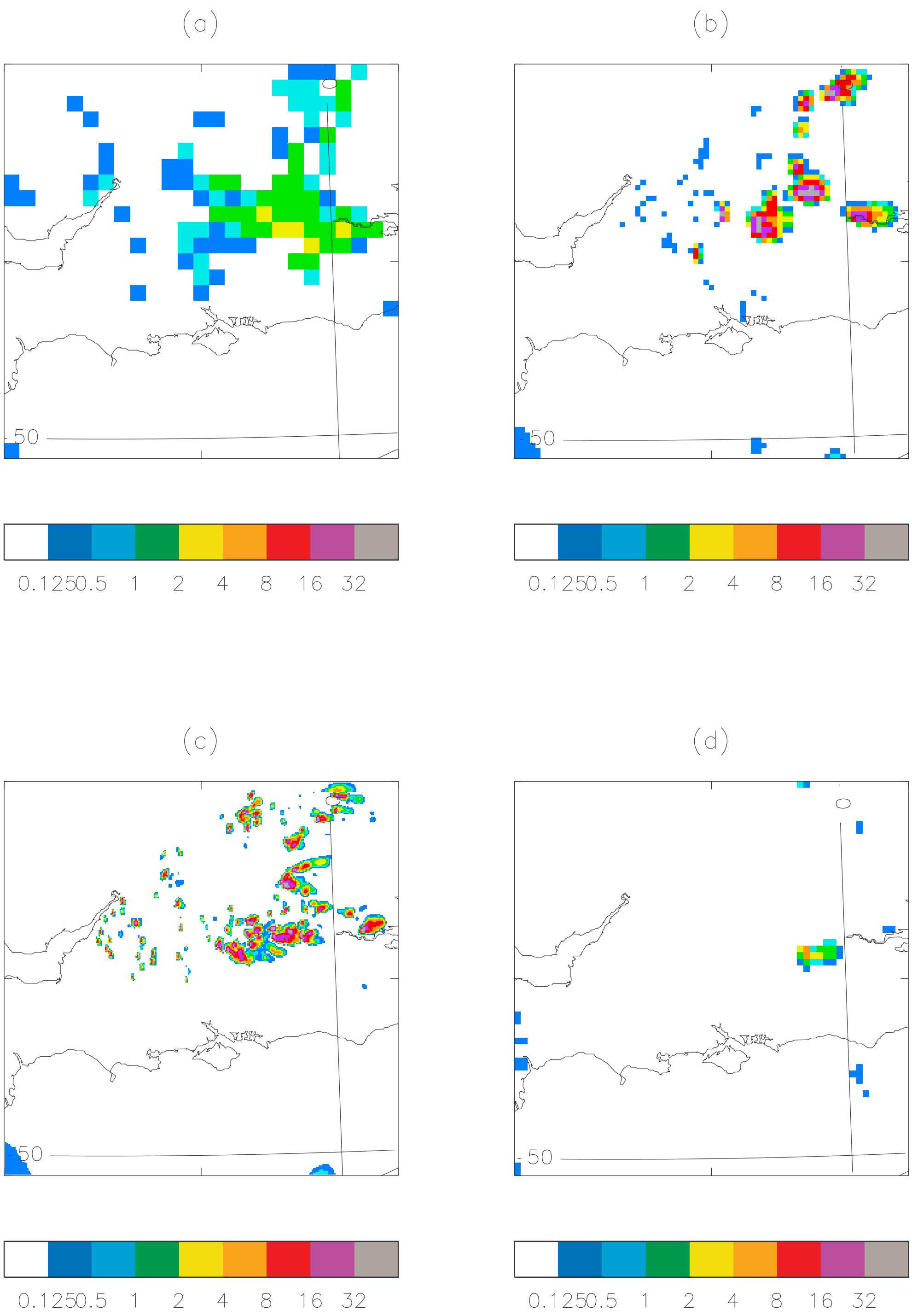


Figure 4.8. Instantaneous rain rates at 13 UTC on 20th July 2004 from 6 UTC model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data.

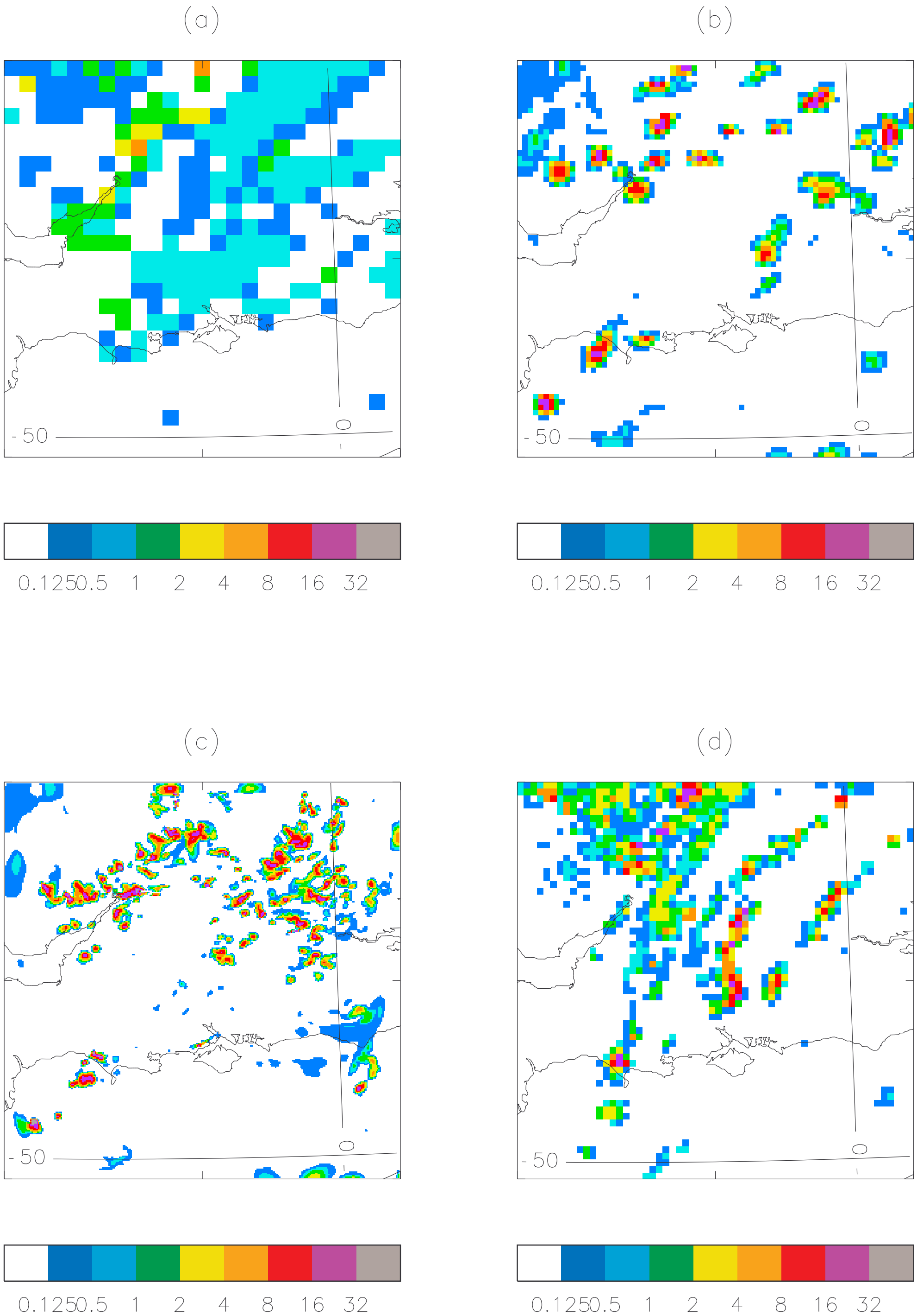


Figure 4.9. Instantaneous rain rates at 14 UTC on 20th August 2004 from 9 UTC model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data.

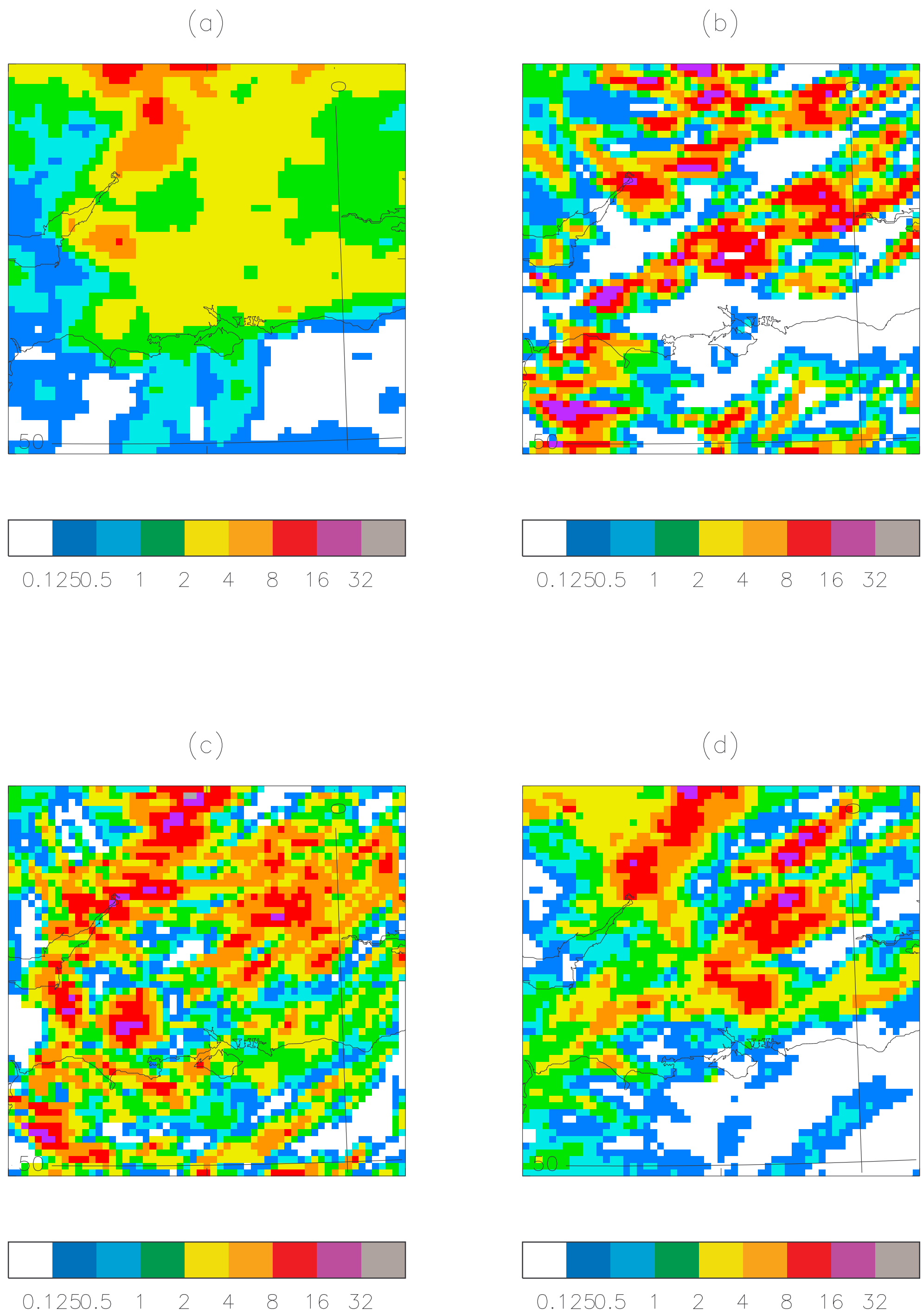


Figure 4.10 20th August 2004 10-16 UTC precipitation accumulation from 9 UTC model runs. (a) is 12km, (b) 4km and (c) 1km. (d) is equivalent radar data. For clean comparison all fields have been interpolated/aggregated onto a 5km grid.

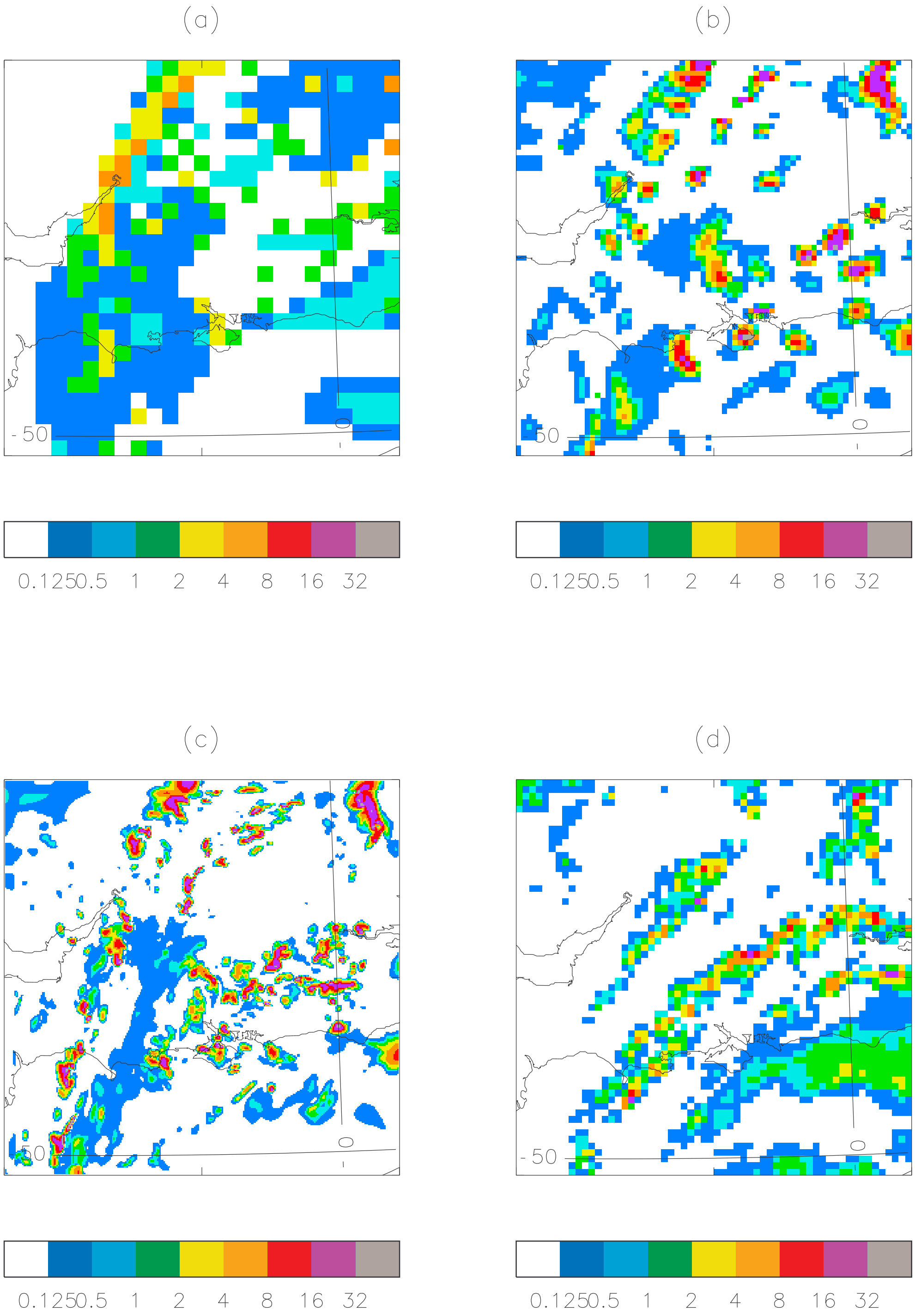


Figure 4.11. Instantaneous rain rates at 13 UTC on 8th July 2004 from 6 UTC model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data.

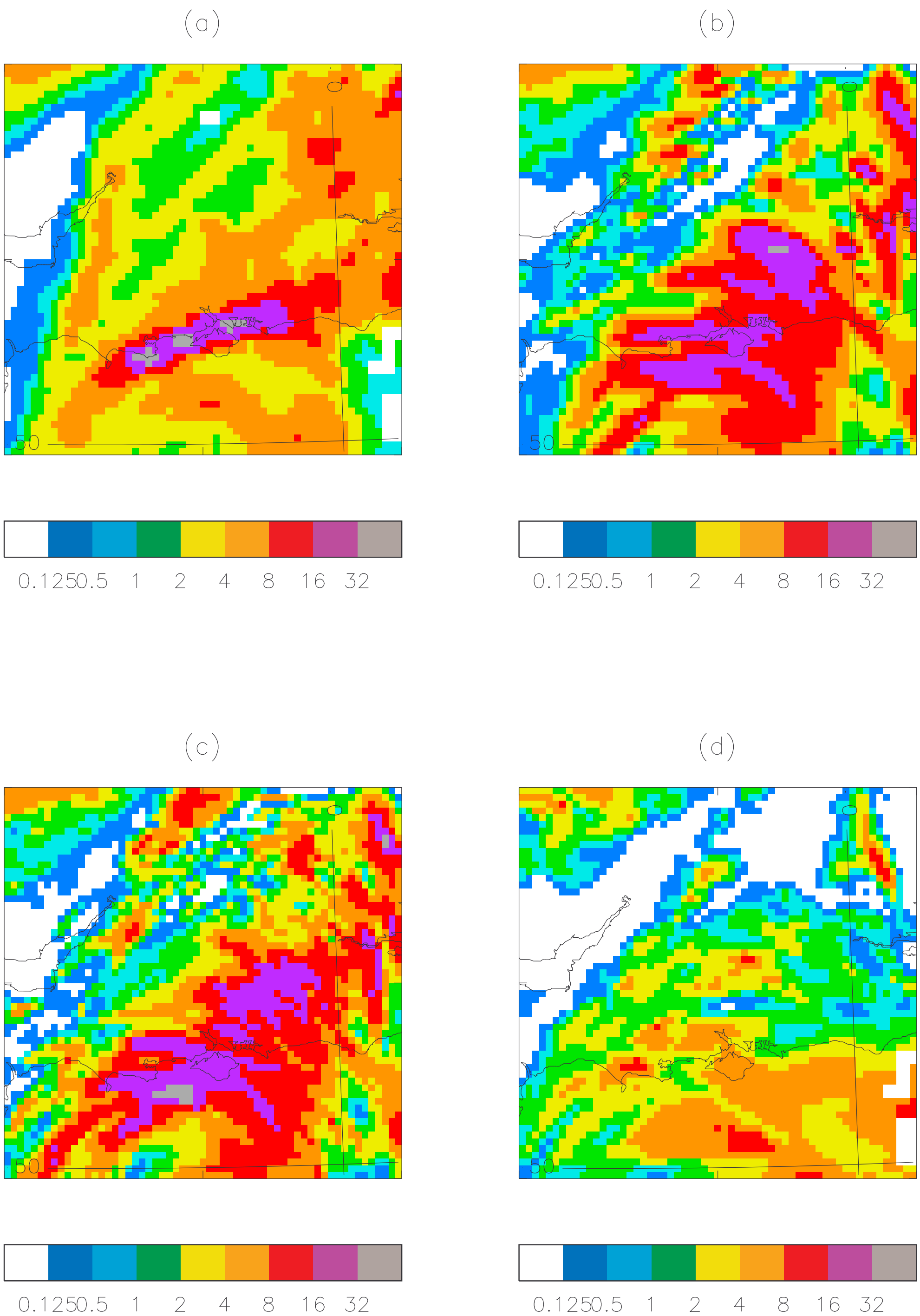


Figure 4.12 8th July 2004 7-13 UTC precipitation accumulation from 9 UTC model runs. (a) is 12km, (b) 4km and (c) 1km. (d) is equivalent radar data. For clean comparison all fields have been interpolated/aggregated onto a 5km grid.

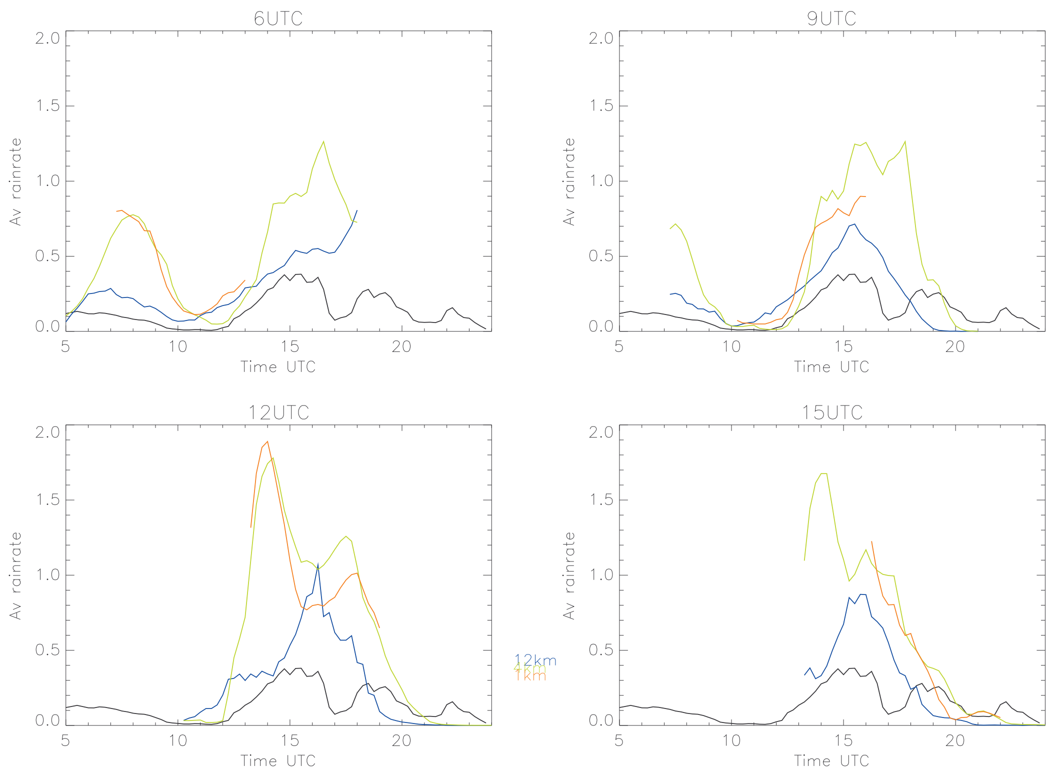


Figure 4.13 1km model area averaged rain rates for 22nd July 2004 showing radar data (thick black line), 12km model (blue), 4km (green) and 1km (red) from each of the four forecasts. The 1km data runs only from T+1 to T+7 in each case whereas the 12km and 4km data runs from T-2 to T+12 (in the latter the period T-2 to T+1 is the assimilation cycle).

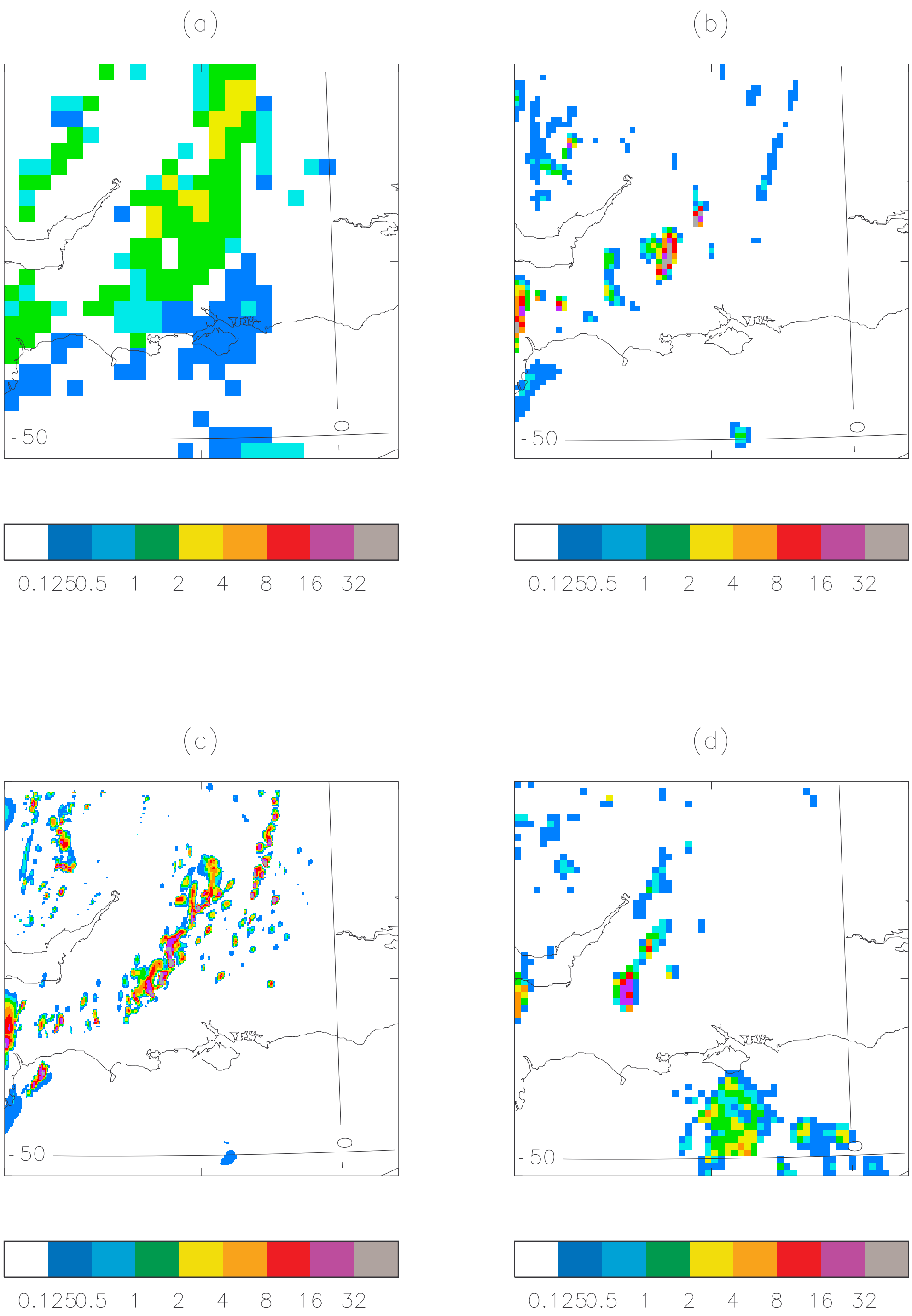


Figure 4.14. Instantaneous rain rates at 13 UTC on 22nd July 2004 from 6 UTC model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data.

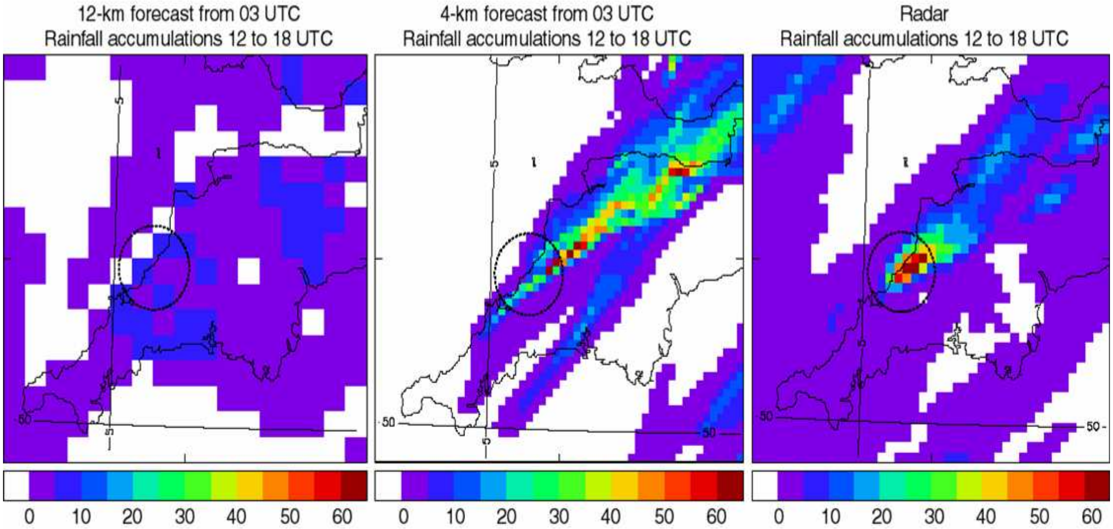


Figure 4.15. 12-18 UTC rain accumulations for 12km and 4km models (3 UTC spin up run) and radar. The circle show a 20km radius around Boscastle.

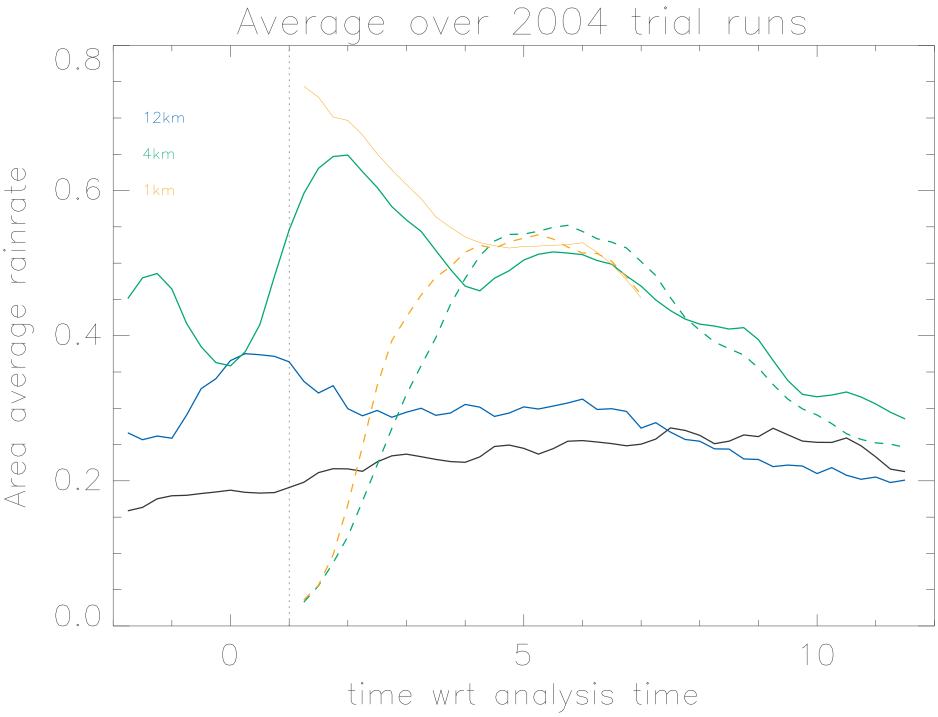


Figure 4.16. 1km domain area averaged rainfall rates averaged over all four cycles of all the cases. Black is the radar data, blue, 12km, green 4km and red 1km. The dotted lines are the spin up runs and the solid are the ones with assimilation. Note that the 1km run with assimilation is shown only from T+1 onwards since the portion T-2 to T+1 was run as a separate model run. The vertical dotted line serves to emphasise the end of the assimilation cycle and the start of the forecast.

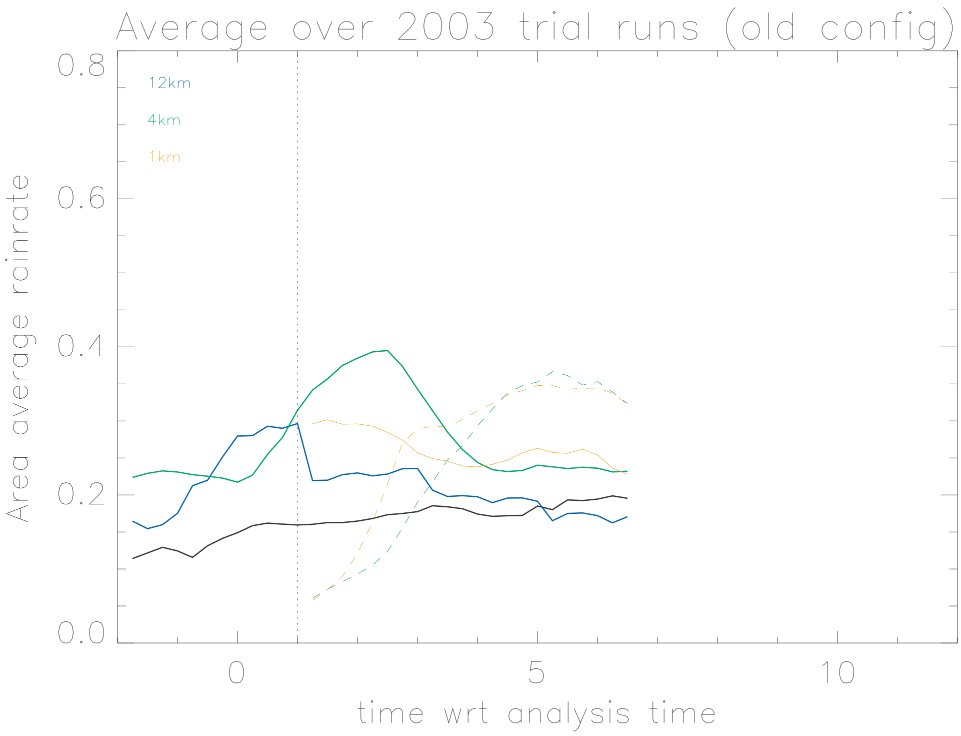


Figure 4.17. As figure 4.15 but for 2003 reruns. The models were only run out to T+7 in these cases.

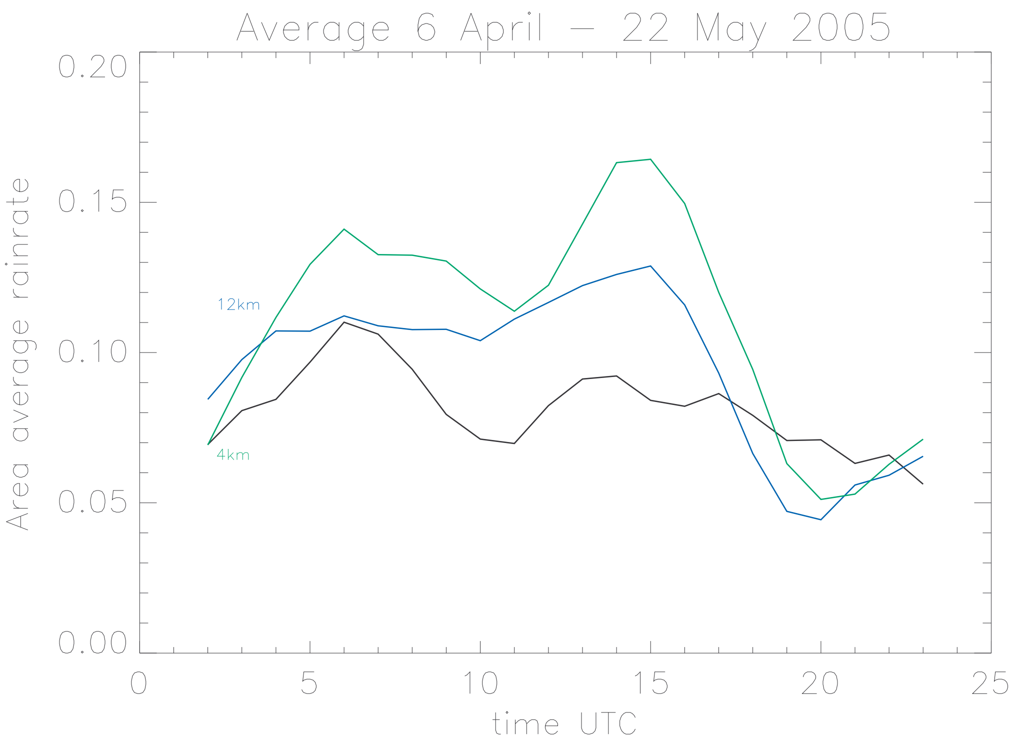


Figure 4.18 1km domain averaged rain rates for daily model runs from 6th April-22nd May 2005. The black curve is the radar data, blue is the 12km model and green is the 4km.

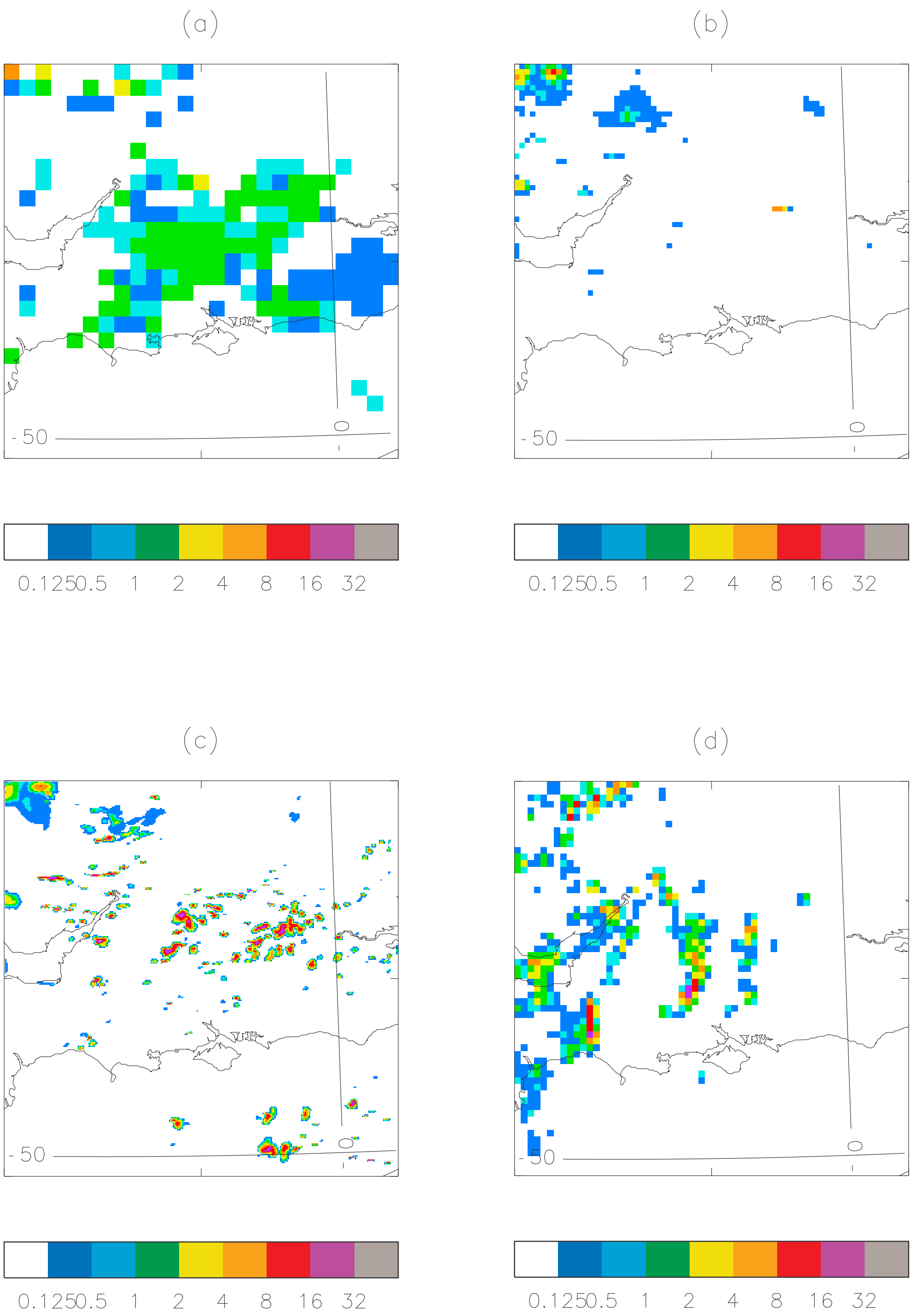


Fig 4.19 Instantaneous rain rates at 10 UTC on 10th July 2004 from 3 UTC spin up model runs. (a) is 12km model, (b) is 4km and (c) is 1km. (d) is the radar data. (i.e. as fig 4.2 except from Spin up runs).

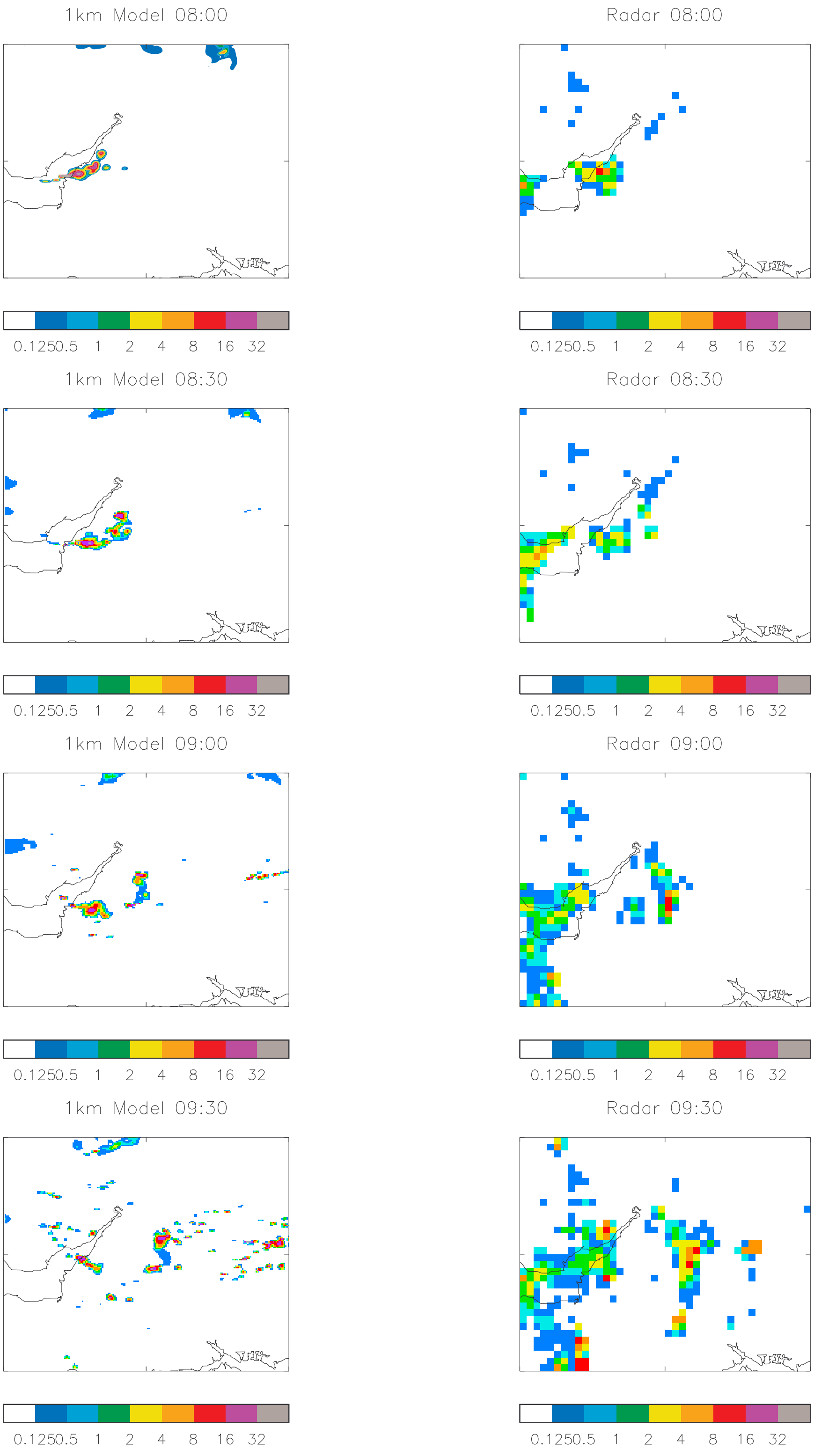


Figure 4.20 Comparison of 1km model 03 UTC run to radar at 30 minute intervals from 08:00 to 09:30 UTC on 10th July 2004 showing initial evolution of showers.

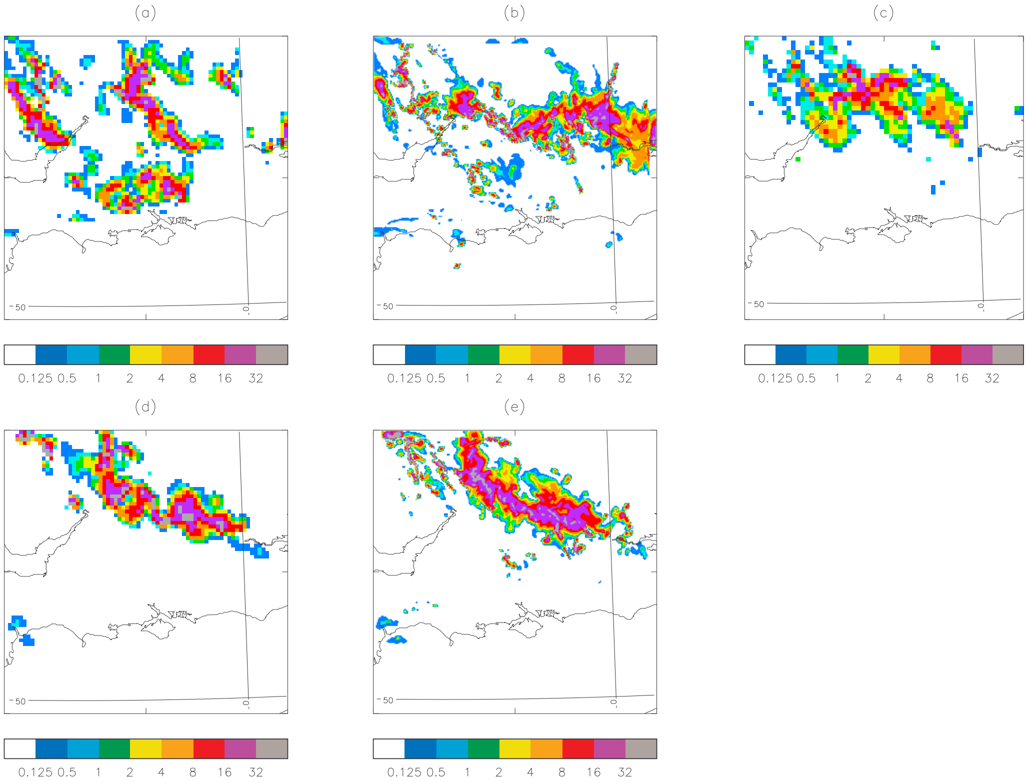


Fig 4.21. Instantaneous rain rates at 16 UTC from the 12 UTC model runs. (a) and (b) are the rain rates from the 4km and 1km models respectively from the assimilation suite and (c) is the radar data. (d) and (e) are the 12km and 4km rain rates from the spin up models.

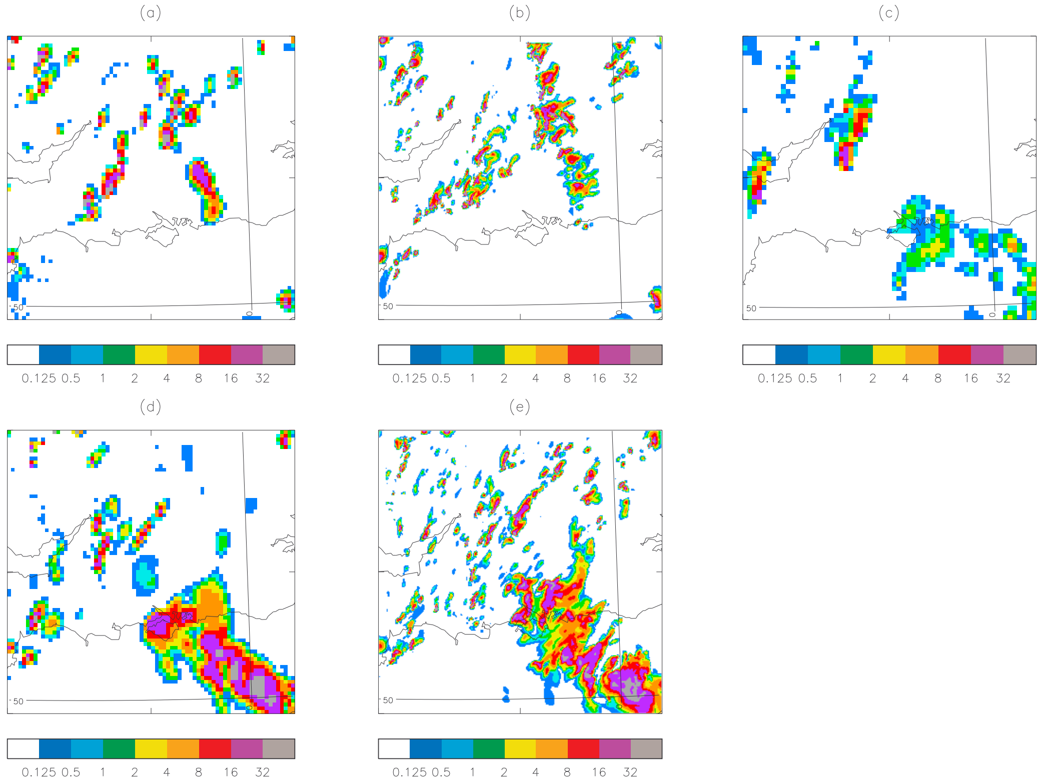


Figure 4.22 Instantaneous rain rates for 14 UTC 22nd July 2004. (a) and (b) are 9 UTC 4km and 1km model runs. (c) is radar data. (d) and (e) are 12UTC 4km and 1km model runs.

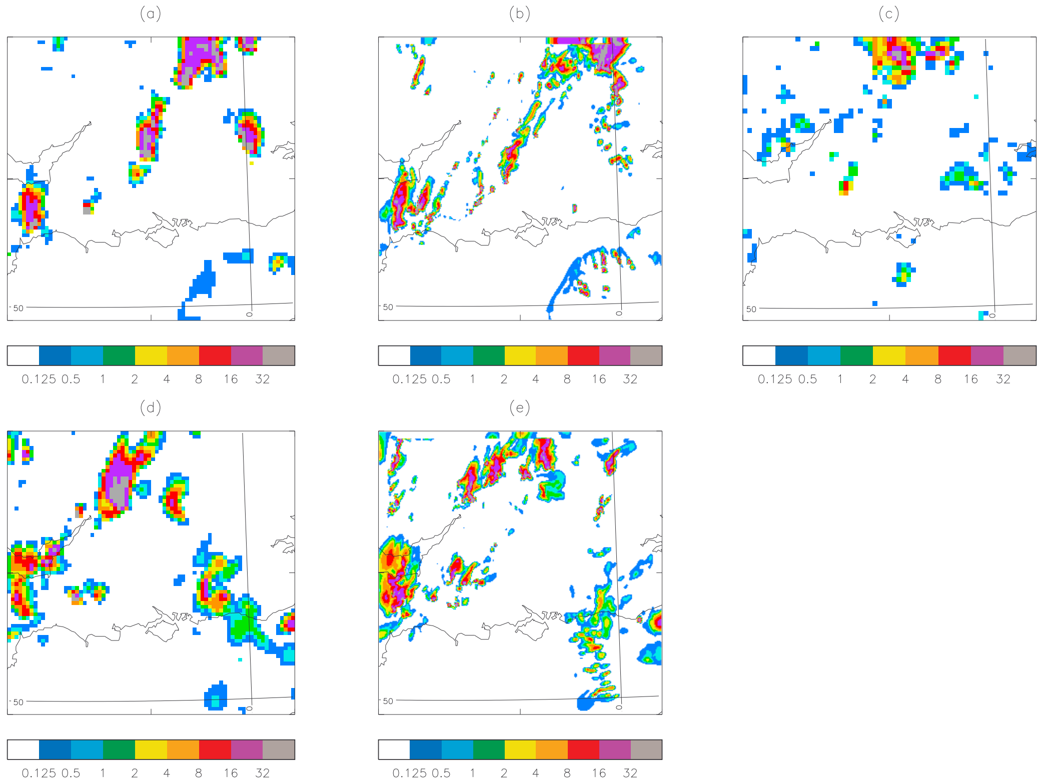


Figure 4.23. As figure 4.22 but two hours later (16 UTC 22nd July 2004).

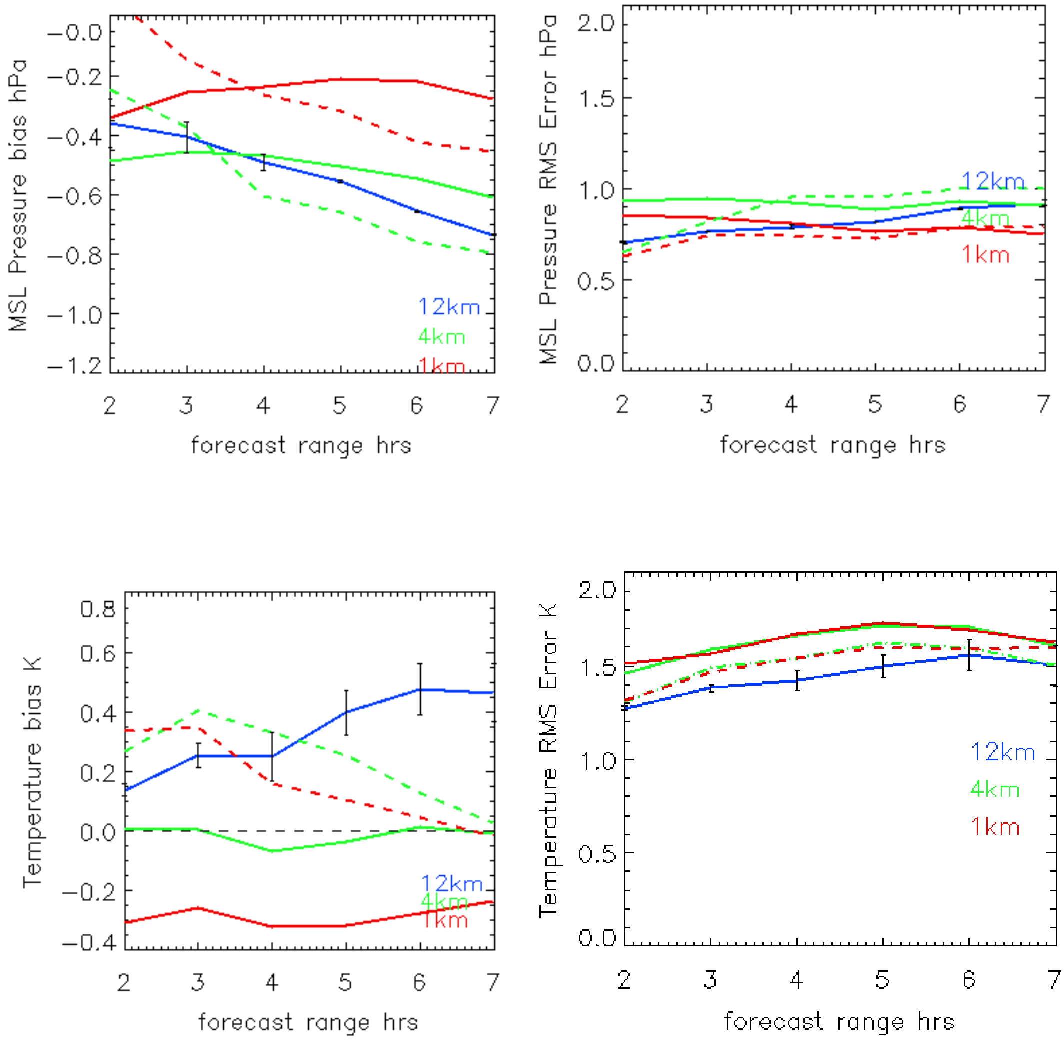


Figure 5.1. Bias and RMS errors between models and surface observations averaged over all forecasts of the first 6 cases in table 2. The coloured lines represent the various models as in the key. The solid lines are the runs with assimilation and the dotted are spin up. The error bars represent 1 standard deviation for the 12km model – for clarity this is shown for the 12km model only but is representative of the spread in the other models also.

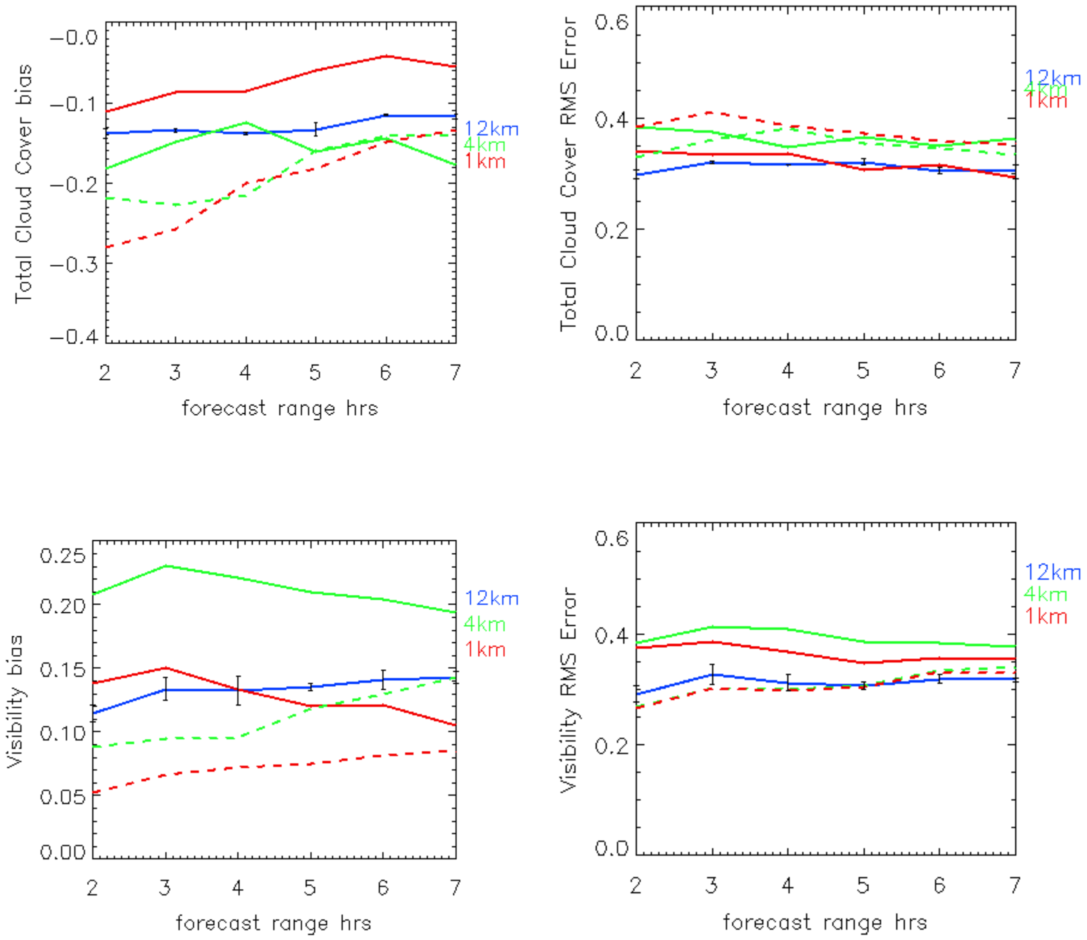


Figure 5.2. As fig 5.1 but for fractional total cloud cover and visibility. The visibility stats were calculated by taking the $\log_{10}$ before doing the calculations.

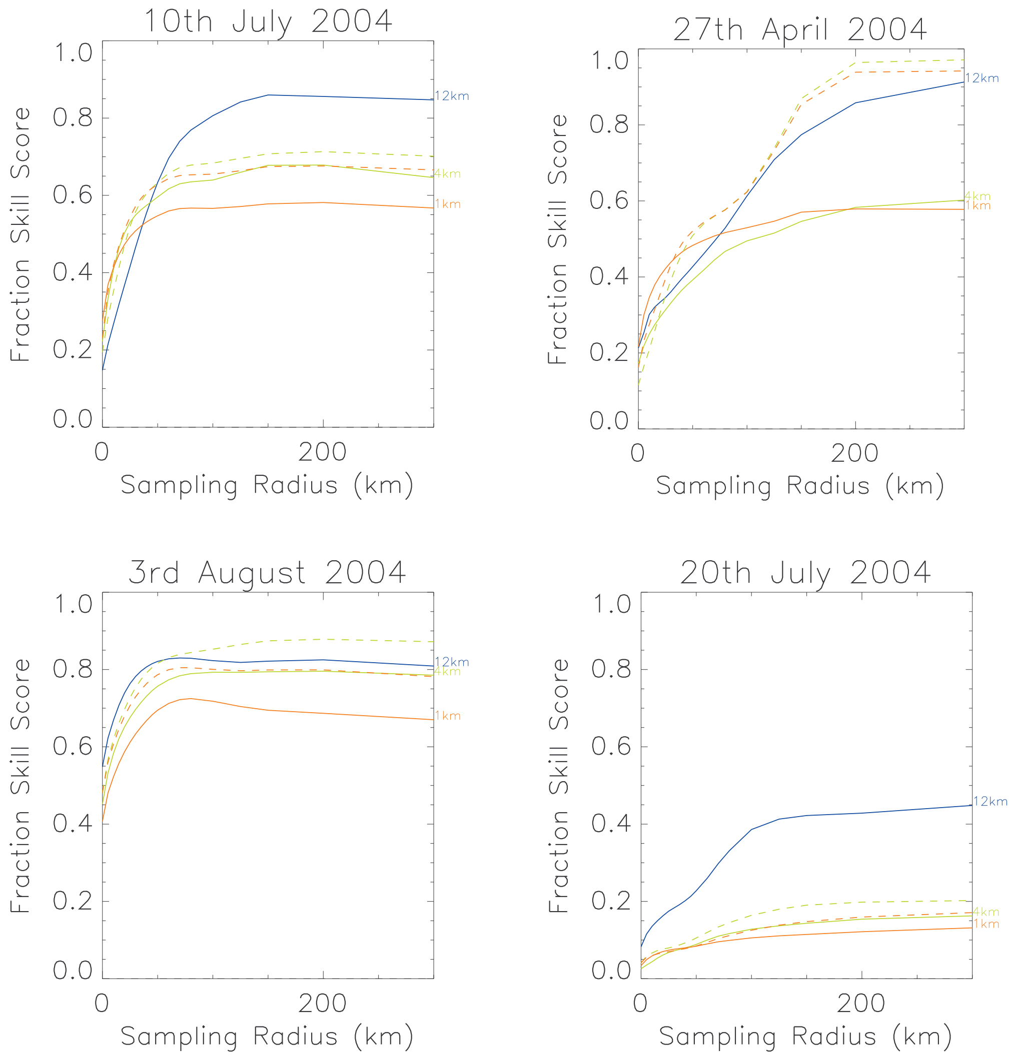


Figure 5.3 Fraction skill scores against radius for threshold of 4mm in 6 hours for the first four cases. In each case the scores are aggregated over the four forecasts run. Blue line is 12km model, green is 4km and red is 1km. Solid lines are assimilation suites, broken are spin up.

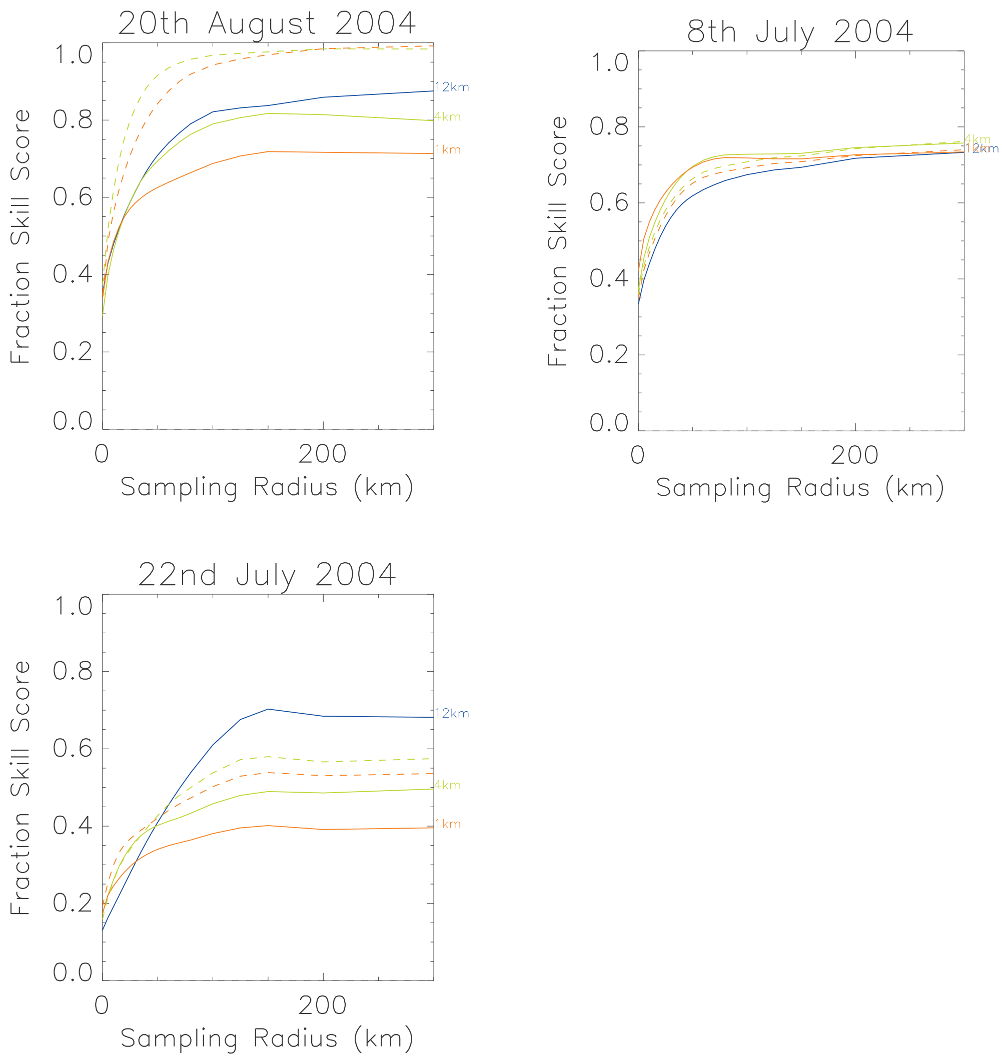


Figure 5.4 As figure 5.3 for remaining three cases.

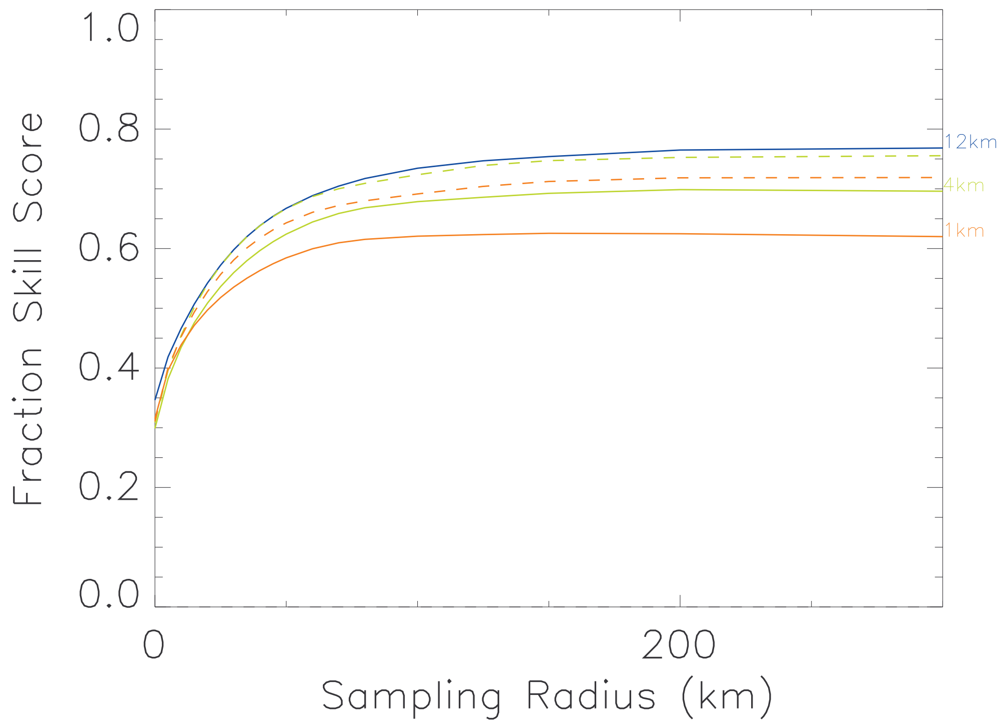


Figure 5.5 As figure 5.3 but aggregated skill score (for 4mm in 6 hours) over all forecasts of all cases.

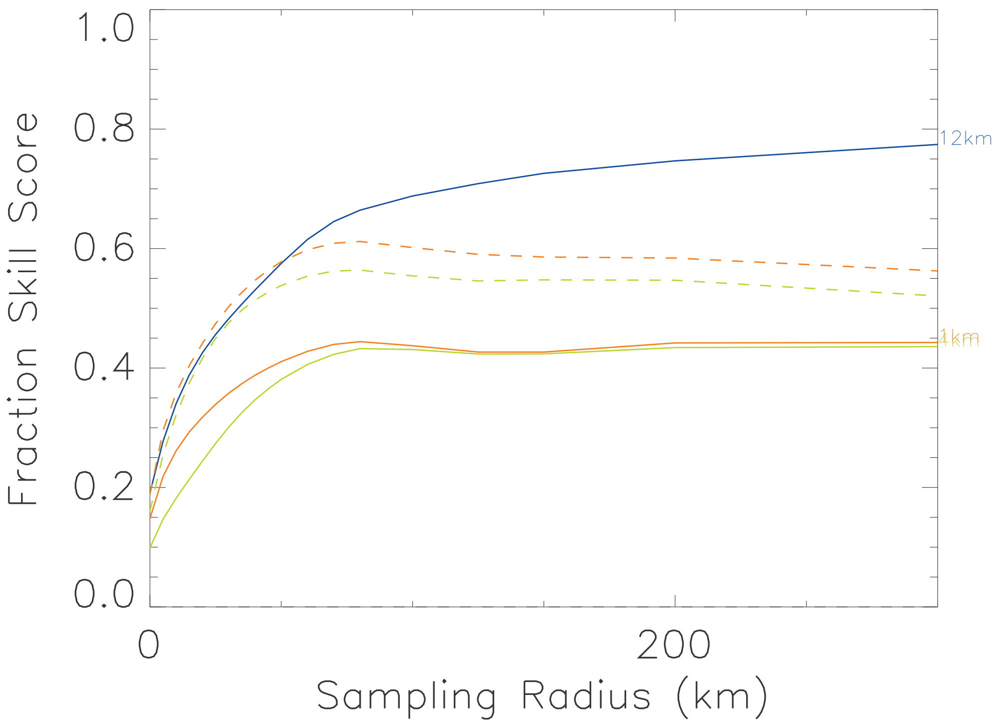


Figure 5.6 As figure 5.5 but for a threshold of 16mm in 6 hours.

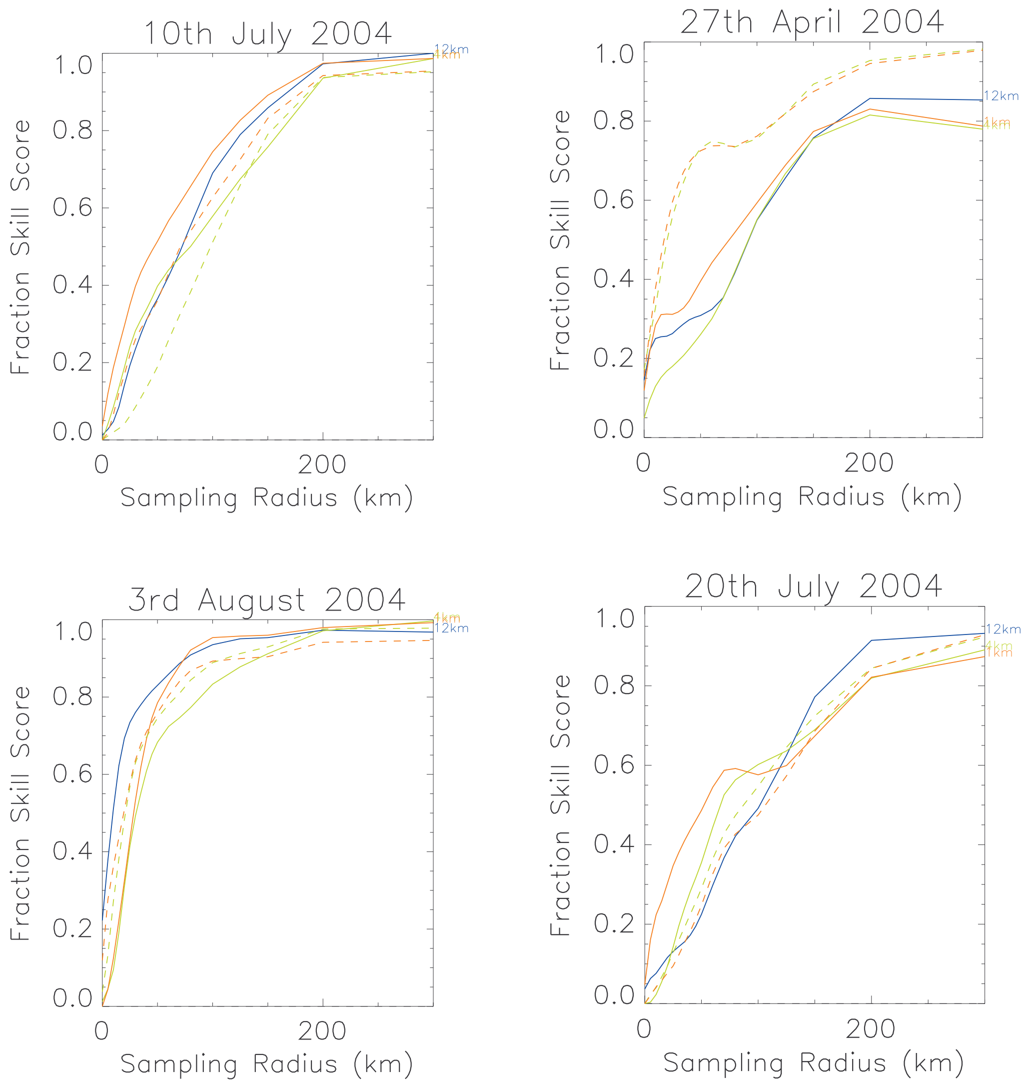


Figure 5.7. Fraction skill scores against radius for relative threshold of top 1% of points for 6 hour periods for the first four cases. Blue line is 12km model, green is 4km and red is 1km. Solid lines are assimilation suites, broken are spin up.

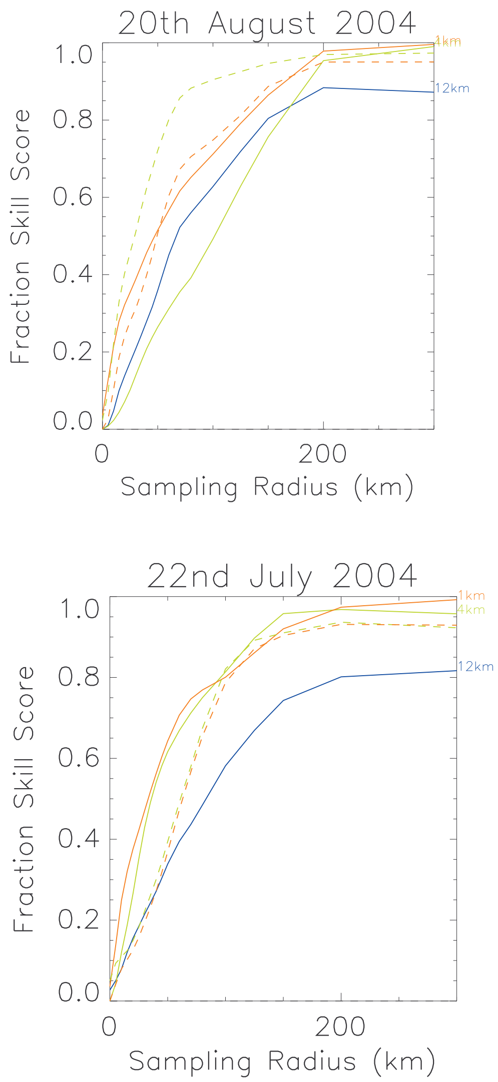


Figure 5.8 As figure 5.7 for remaining three cases.

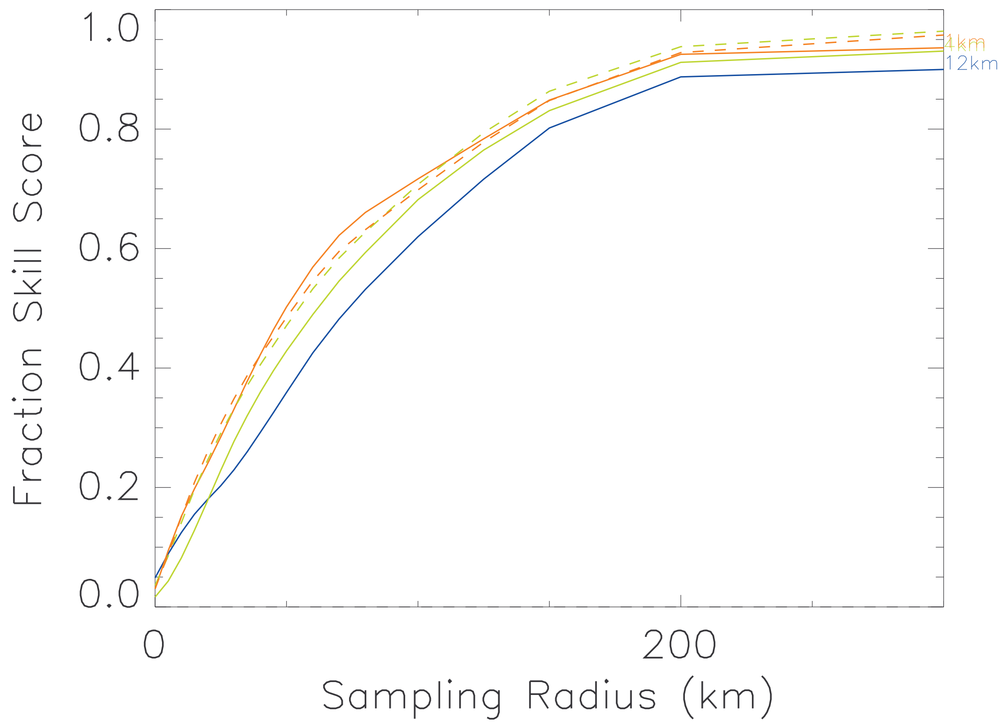


Figure 5.9 As figure 5.7 but aggregated skill score (for top 1% of points) over all forecasts of all cases.

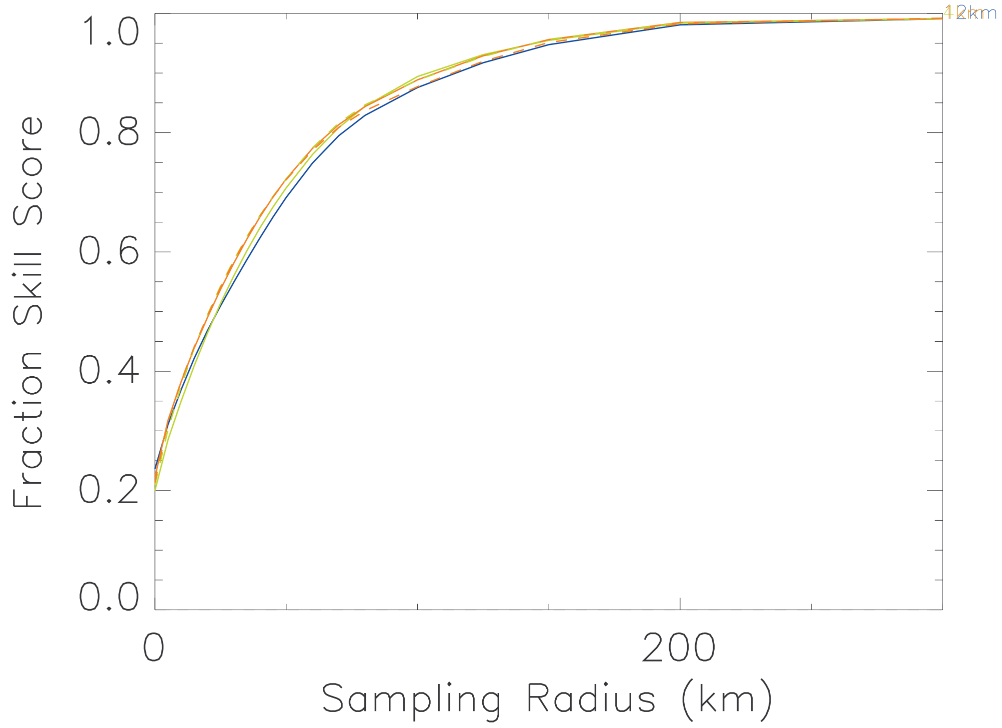


Figure 5.10. As figure 5.9 but for a relative threshold of top 10% of points.

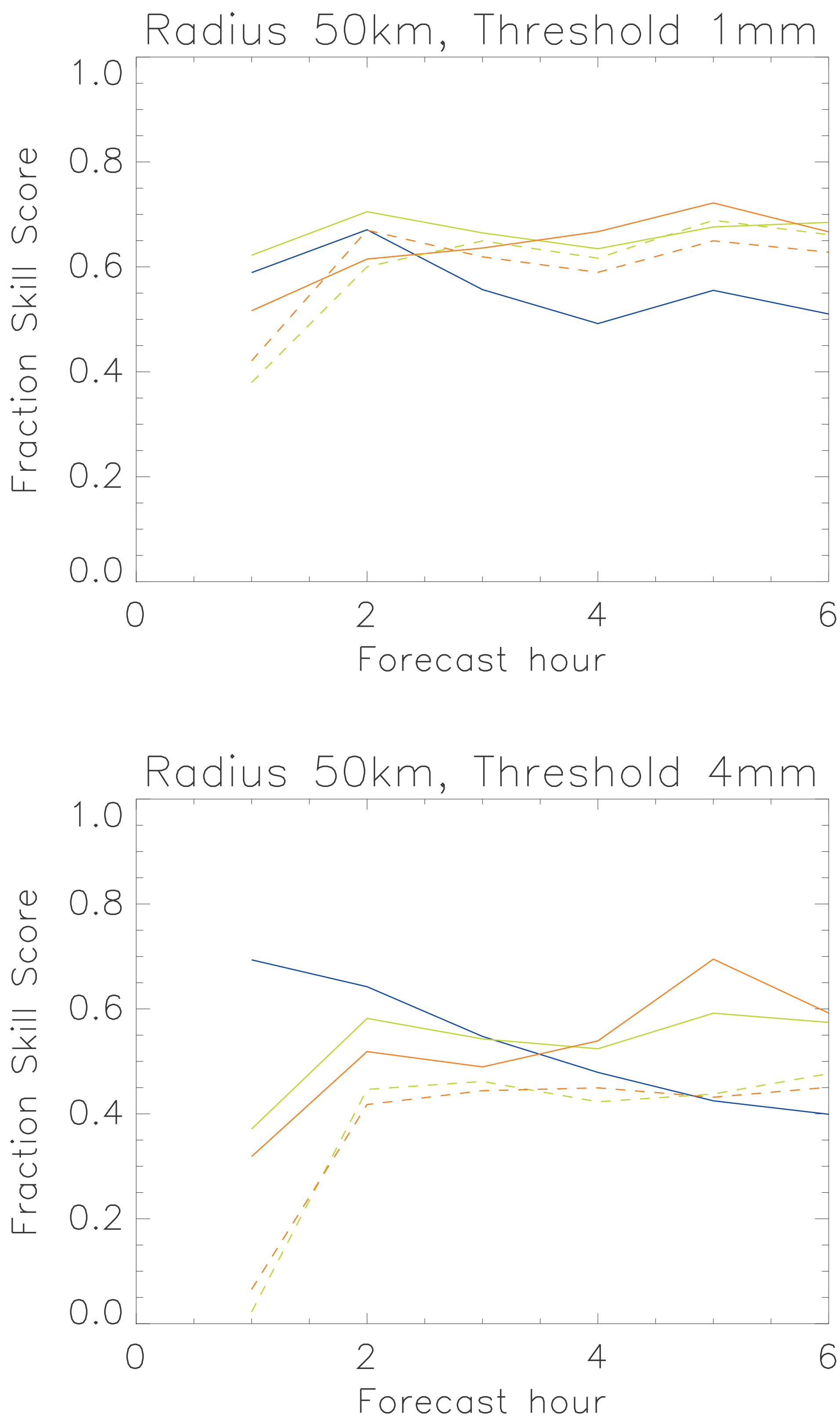
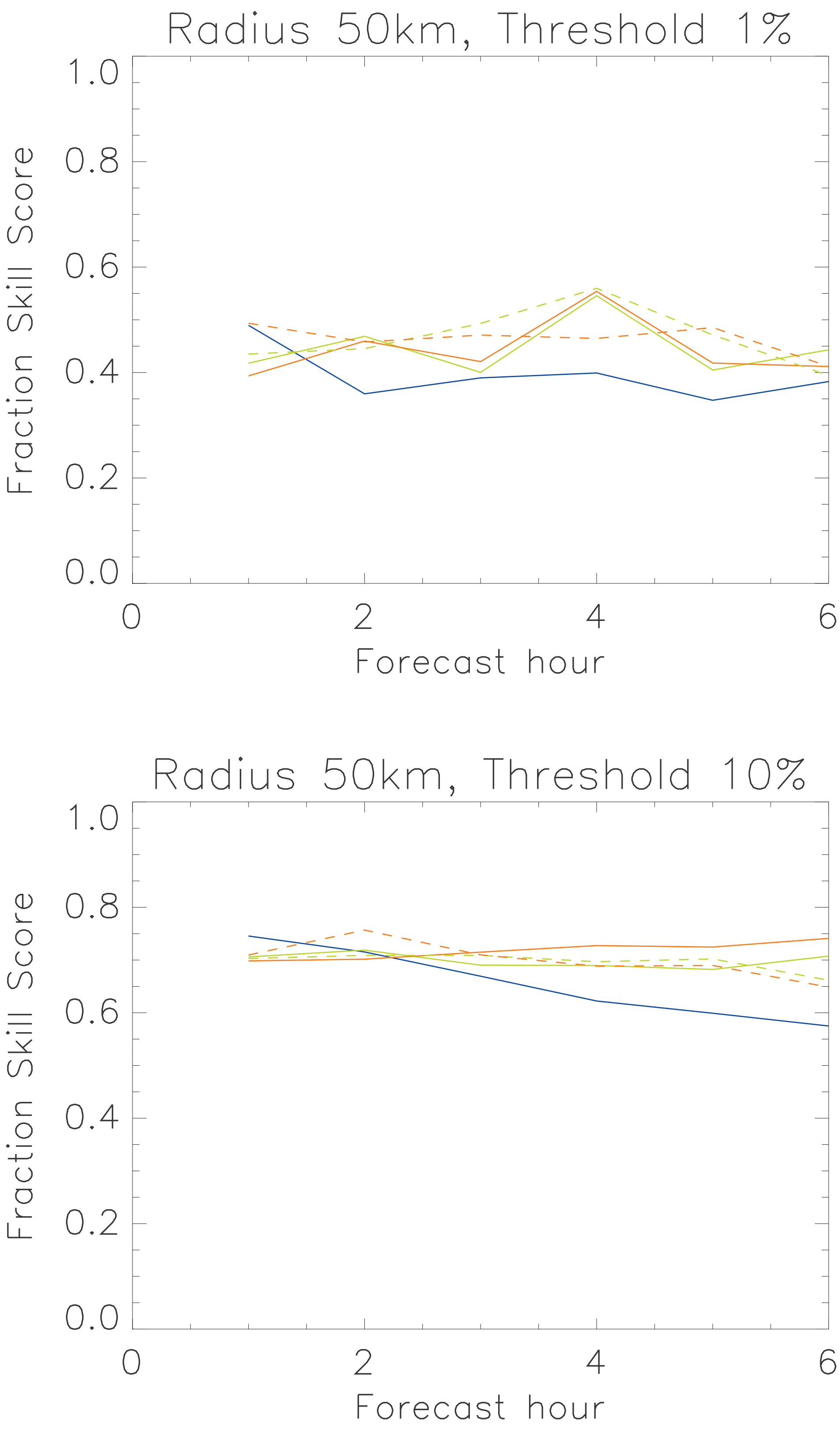


Figure 5.11 Fraction skill scores for hourly accumulations plotted against forecast hour aggregated over all runs of all cases for a 50km sampling radius. Blue line is 12km model, green is 4km and red is 1km. Solid lines are assimilation suites, broken are spin up.



Figures 5.12 Fraction skill scores for hourly accumulations relative thresholds plotted against forecast hour aggregated over all runs of all cases for a 50km sampling radius. Blue line is 12km model, green is 4km and red is 1km. Solid lines are assimilation suites, broken are spin up.