

# Numerical Weather Prediction

An assessment of the Met Office's ability to forecast air traffic disruption  
due to thunderstorm activity



Forecasting Research Technical Report No. 408

Helen A. Watkin and Deborah J. Hoad

*email: [nwp\\_publications@metoffice.com](mailto:nwp_publications@metoffice.com)*

©Crown Copyright

# **An assessment of the Met Office's ability to forecast air traffic disruption due to thunderstorm activity**

Helen A. Watkin and Deborah J. Hoad

April 2003

## **Abstract**

To ensure passenger safety and comfort pilots often fly around thunderstorms, causing problems for Air Traffic Control. Advance notice of thunderstorm activity in a particular area would enable Air Traffic Management to adjust air traffic flow before the disruption occurs. In this study, Met Office forecasts of air traffic disruption due to thunderstorm activity are produced using several different forecast tools, lead times, and spatial resolutions. The forecasts are verified, against observations of air traffic disruption produced from 3D rain rate analyses, using several statistical forecast quality measures over a sample of days throughout the summer of 2002.

The forecasting scheme first finds the maximum rain rate in the vertical and uses this to replicate a pilot radar display. Using information on pilot behaviour these replicated pilot radar displays are used to assess what percentage of aircraft in an area will need to deviate from their expected flight paths. If this percentage passes a threshold value, air traffic disruption due to thunderstorm activity is forecast.

The results from the main study show that the disruption forecasts produced by extrapolating a 3D rain rate analysis outperform the forecasts produced from both Gandolf and Nimrod rain rate forecasts out to a 120-minute lead time. When compared with a linear-trended 3D extrapolation forecast and a persistence forecast in a supplementary study, the 3D extrapolation forecast is still shown to be of the overall highest quality. Forecast accuracy is shown to increase as the spatial resolution of the forecast decreases. Forecast quality is generally very good at the short lead times, but tails off with increasing lead time. A cost-benefit analysis carried out with or by NATS would help to specify a lead time up to which the forecasts would be of use to air traffic management.

The highest quality forecast was compared to Terminal Control weather disruption reports for several case studies. Despite several limitations, this analysis provided a useful first step to verifying the forecasts against actual data of air traffic disruption.

## Document review history

<b>Date</b>	<b>Version</b>	<b>Action/comments</b>	<b>Approval</b>
28/03/03	1.0	Reviewed by Clive Pierce	
15/04/03	2.0	Amended after review	Bob Lunnon

## CONTENTS

	Page
<b>Abstract.....</b>	<b>2</b>
<b>Document review history .....</b>	<b>3</b>
<b>1. Thunderstorms and Air Traffic Control: a background.....</b>	<b>5</b>
<b>2. Previous work into forecasting thunderstorms for Air Traffic Control .....</b>	<b>6</b>
<b>3. Forecasting and verification methods.....</b>	<b>7</b>
3.1 Choosing the sample of days.....	7
3.2 Forecast lead times, spatial resolutions and coverage .....	7
3.3 Stage 1: Rain rate data input.....	8
3.4 Stage 2: Obtaining maximum in-cloud rain rate .....	8
3.5 Stage 3: Replicating pilot radar displays.....	9
3.6 Stage 4: Classifying deviation pixels.....	10
3.7 Stage 5: Forecasting air traffic disruption .....	10
3.8 Stage 6: Verification .....	11
<b>4. Comparison between forecast types .....</b>	<b>12</b>
4.1 Results from the main study(using 22 sample days) .....	12
4.2 Results from the supplementary study (using 6 sample days) .....	14
<b>5. Effect of spatial resolution on forecast skill .....</b>	<b>16</b>
<b>6. Analysis of actual air traffic disruption data.....</b>	<b>18</b>
<b>7. Discussion.....</b>	<b>19</b>
7.1 Influence of forecast type on forecast accuracy.....	19
7.2 Impact of spatial resolution and lead time on forecast accuracy .....	19
7.3 General trends and limitations of the forecasts used in this study.....	21
<b>8. Summary and Conclusions .....</b>	<b>22</b>
<b>9. Recommendations.....</b>	<b>24</b>
<b>10. References .....</b>	<b>25</b>
<b>11. Acknowledgements .....</b>	<b>26</b>
<b>12. List of acronyms.....</b>	<b>27</b>
<b>Appendix A: List of sample days .....</b>	<b>28</b>
<b>Appendix B: Contingency table and formulae for verification statistics .....</b>	<b>29</b>
<b>Appendix C: Statistical measure scores for the main study .....</b>	<b>30</b>
<b>Appendix D: Statistical measure scores for the supplementary study.....</b>	<b>36</b>
<b>Appendix E: Graphs of verification scores: main study .....</b>	<b>42</b>
<b>Appendix F: Graphs of verification scores: supplementary study .....</b>	<b>45</b>
<b>Appendix G: Results from the case studies comparing 3D_extrap forecasts to Terminal Control weather disruption reports .....</b>	<b>48</b>
<b>Appendix H: Maps showing the places mentioned in the Terminal Control weather disruption reports.....</b>	<b>51</b>

# 1. Thunderstorms and Air Traffic Control: a background

National Air Traffic Services (NATS) is responsible for the safe separation of aircraft flying in UK controlled airspace, which is made up of terminal control areas that surround the major airports and airways that link these control areas. The London Area Control Centre at Swanwick, Hampshire handles most aircraft flying over England and Wales and the surrounding seas. The London Terminal Control Centre (LTCC) in West Drayton, Middlesex handles air traffic below 24,500 feet flying to or from London's airports, with their controllers handling up to 2 million movements each year (nats.co.uk). The London Terminal Control Area contains some of the busiest airspace in Europe, and is frequently running at full capacity, with aircraft directed to holding stacks during busy periods. If an outside factor causes aircraft to turn off a standard route or heading, or prevents the use of a stack, then the workload of the controllers significantly increases. It becomes more difficult to keep aircraft safely separated, and ascending and descending in an orderly stream, and air traffic can be said to have been 'disrupted'.

The outside factor causing the disruption is often the weather. The safety and comfort of the aircraft and its passengers is the responsibility of the airlines, and thus pilots tend to fly around any weather that may pose a risk to this safety. Thunderstorms fall into this category because they contain several hazards to aircraft, including severe wind shear, turbulence, icing, hail, lightning, and reduced visibility. These hazards can extend for some distance around the thunderstorm, with guides to Aviation Meteorology recommending that pilots avoid the convective activity by between 10 and 20 nm dependent on storm severity and aircraft altitude (Wilson, 1997; Thorn, 1997). In the UK thunderstorms usually occur in the summer with the area of most frequent activity concentrated in the South East (Soul et al, 2002), coinciding with the busiest time and area in terms of air traffic (nats.co.uk), which increases the impact when aircraft deviate around thunderstorms. Pilots enable a deviation either by requesting a vector from air traffic control, or in busy times by simply changing direction without informing air traffic control. The effect is an increase in workload for the controller on that sector and also on adjoining sectors which aircraft might stray into. More co-ordination is required between sectors, delays increase and maintaining separation becomes more difficult. Blair (2002) found that en-route delay due to thunderstorms contributed 5% of the overall yearly delay in the UK Terminal Control area, although thunderstorms could also have contributed to delays due to airfield weather or sector capacity. Hauf and Sasse (2002) studied the impact of thunderstorms on landing traffic at Frankfurt airport, finding that delays increased by 600% on thunderstorm days when compared to non-thunderstorm days, with thunderstorm-related delays costing c.£850,000 pa. In the US thunderstorms contribute to more than half of air traffic delays, in addition to many accidents and incidents (Souders and Showalter, 2002).

If the time and location of air traffic disruption due to thunderstorm activity could be forecast in advance then traffic flow could be reduced proactively before the disruption occurs rather than reactively as at present. For example, the control manager could decide to reduce the minimum departure interval (MDI) on a route from an airport, or to reduce the TSF in an air traffic sector. TSF is defined as “the number of aircraft entering an ATC sector per hour, which it is judged can be safely handled by the operational staff on the sector, without an excessive workload.” (Blair, 2001). By taking these steps the workload of controllers would be protected, the safety of aircraft ensured, and delays minimised.

This report attempts to assess the Met Office's ability to forecast air traffic disruption due to thunderstorm activity. Section 1 has provided a background to the effect of thunderstorms on air traffic management. Section 2 details previous studies in the UK and the US on this subject. Section 3 then explains how the forecasts of air traffic disruption due to

thunderstorms have been produced and verified. Section 4 provides a comparison between the forecast types, while section 5 details the effect of spatial resolution on forecast skill. An analysis of data of actual air traffic disruption is carried out in section 6. Section 7 then discusses important themes to come out of the results, with sections 8 and 9 providing some conclusions and recommendations resulting from the study.

## **2. Previous work into forecasting thunderstorms for Air Traffic Control**

The Met Office decided to develop and verify the accuracy of a forecasting product that would predict air traffic disruption due to thunderstorm activity. Hoad (1999) investigated an early version of the GANDOLF (Generating Nowcasts for the Deployment of Operational Land-based Flood forecasts) system, which provided short-range forecasts of thunderstorm cells up to 3 hours ahead. For several days with thunderstorm activity over southeast England, GANDOLF forecasts were verified against radar data. The results showed that the forecasts provided useful information up to about 90 minutes ahead. Watkin et al (2002) derived statistical forecast quality measures for forecasts of air traffic disruption due to thunderstorm activity using Mesoscale, Nimrod and CDP data, at a number of different spatial resolutions and lead times. Forecast accuracy declined with increasing spatial resolution and lead time, while forecast quality was shown to be greatest for disruption forecasts prepared from Nimrod data and following a method that uses both storm intensity and coverage to predict deviations.

In the US, several investigations have been carried out into aircraft encounters with thunderstorms in en-route and terminal airspace (Rhoda et al 2002; Rhoda et al 2000; Rhoda and Pawlak, 1999). These studies indicate that it is possible to use data of the intensity and coverage of thunderstorms to predict pilot behaviour in or close to terminal air space. Several thunderstorm nowcasting forecast products are available for US air traffic control. Many are based on extrapolation methods, but the newly developing products make more attempt to account for storm development or decaying and in the future may attempt to create and kill off cells. An Integrated Terminal Weather System (ITWS) has been installed at 47 busy airports and integrates sensor and model data to create tailored weather products such as the point location of storm cells and their extrapolated position out to 30 minutes (Souders and Showalter, 2002). The National Convective Weather Forecast (NCWF) converts NEXRAD weather radar data into a Vertically Integrated Liquid water content (VIL) and combines this with lightning data to produce a four-colour hazardous weather depiction and 1-hour forecast (Kulesa et al, 2002). The Collaborative Convective Forecast Product (CCFP) uses forecaster input to identify areas with at least 25% coverage of thunderstorms with tops of at least 25,000 feet up to six hours in the future. There is no attempt to depict individual thunderstorms because this is well beyond the state of the science (Hudson and Foss, 2002). The Regional Convective Weather Forecast (RCWF), demonstrated in 2002, combines radar data with convective weather prediction algorithms to forecast the probability of NWS-level 3 and greater radar reflectivity up to 2 hours ahead (Boldi et al, 2002).

Thus US research shows how aircraft will react given a set of conditions, and US forecast products predict where and when thunderstorms will occur, but as of yet the two strands of research have not been rolled together to produce a forecast of where the disruption will occur. This project will attempt to do this over UK airspace.

### 3. Forecasting and verification methods

#### 3.1 Choosing the sample of days

Lightning data were used to identify thunderstorm events across the UK during summer 2002. Data was archived from 16 days with significant thunderstorm activity, and also from 6 random days to ensure that a fair sample of data was being used. A list of the days used is available in Appendix A.

#### 3.2 Forecast lead times, spatial resolutions and coverage

On these days, forecasts of air traffic disruption due to thunderstorm activity were produced at lead times out to 360 minutes, at 15-minute intervals (up to 180 minutes when using Gandolf data). Four spatial resolutions of disruption forecast were produced: 10km, 20km, 40km and 80km. For clarification, 10km is referred to as high resolution and 80km as low resolution for the rest of this report. The area for which forecasts and analyses were produced was from 80E, 520N to 720E, -120N. However, some of the outer parts of this area have missing data and hence the area used in the verification process is smaller, covering all of England and Wales apart the very north of England.

The forecasting and verification methods used are similar to those used in Watkin et al (2002). A summary is provided in Figure 3.2.1. Details of the methods used in each stage are given below.

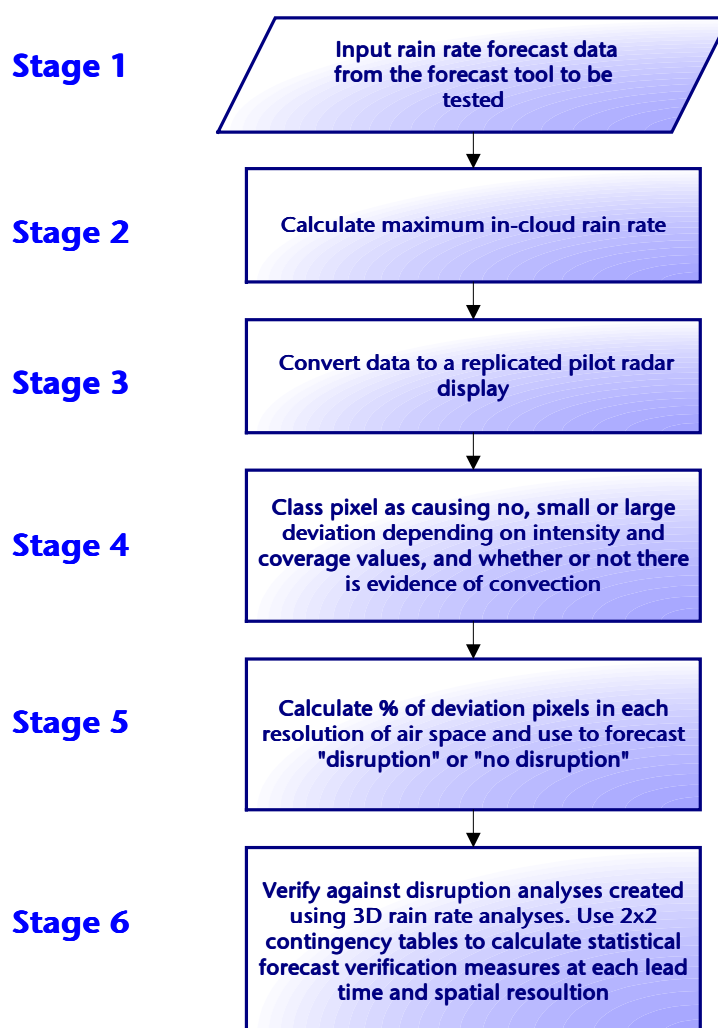


Figure 3.2.1: Summary of forecast and verification methods used in the main study

### 3.3 Stage 1: Rain rate data input

Three types of rain rate forecast data are trialled in the main study and used to produce air traffic disruption forecasts for the 22 days:

#### 1. Nimrod 5km surface rain rate data

The Met Office's Nimrod nowcasting system integrates recent observations with Mesoscale Model products to provide forecasts up to six hours ahead. Watkin et al (2002) found disruption forecasts prepared using Nimrod rain rate data to be of a higher quality than Mesoscale and CDP data, so Nimrod data is tested alongside the new forecast tools. In this report, the disruption forecasts produced using this rain rate data are referred to as *Nimrod* forecasts.

#### 2. GANDOLF 2km surface rain rate data

The Met Office's GANDOLF system combines a 2km resolution Nimrod advection forecast scheme integrated with an improved version of the Object-Oriented convective life cycle Model (OOM) developed by Hand (1996) and improved by Pierce (Pierce and Cooper, 2000). The OOM is only run when severe convection develops in the presence of strong wind shear, because the advection scheme outperformed the OOM in all but these rare occasions (Pierce and Cooper, 2000). The disruption forecasts produced using the Gandolf surface rain rate data are referred to as *Gandolf* forecasts.

#### 3. 2km extrapolation forecast of Gandolf 3D rain rate analyses

The third forecast data type is a new 3D extrapolation forecast of rain rate developed by JCHMR at Wallingford for trial in this study. Peak rain rate analyses are produced for four meteorologically significant layers (ground to Convective Cloud Base, Convective Cloud Base to 0°C, 0°C to -10°C, -10°C to Convective Cloud Top). An advection forecast is then produced for each layer using an optic flow advection scheme (Bowler and Pierce, 2002). The disruption forecasts produced using this new 3D rain rate data are referred to as *3D\_extrap* forecasts for the remainder of this report.

Two further forecast data types were trialled in a supplementary study, but due to time constraints these could only be verified for a sample of six thunderstorm days:

#### 4. 2km linear-trended extrapolation forecast of 3D rain rate analyses.

This was also developed by JCHMR at Wallingford for trial in this study. Similar to forecast type 3 it is produced from 3D analyses of rain rate, but this time produced for four CAPPI's (Constant Altitude Plan Position Indicators), rather than meteorological layers. The altitudes used were 1) ground to 1km, 2) 1-3km, 3) 3-5km, and 4) 5km and above. The forecasts were produced on each layer using an advection scheme that trends the rain rate with time. The disruption forecasts produced using this rain rate data are referred to as *3D\_trended* forecasts from this point forward.

#### 5. Persistence.

Here the 3D rain rate analyses from the previous forecast times are used to see how a persistence forecast compares to the more sophisticated forecast data types.

### 3.4 Stage 2: Obtaining maximum in-cloud rain rate

When deciding whether or not it is safe to fly through an area of convective rain, pilots are able to tilt their weather radar in order to ascertain the maximum reflectivity in the vertical. Since the rain rate forecast data in Nimrod and Gandolf are for the rate at the surface, a rain rate profile must be assumed in order to convert this into the maximum rain rate in the cloud. Watkin et al (2002) assumed a scaled rain rate profile of a mature thunderstorm when converting surface convective rain rate into the maximum in-cloud rain rate, but this overestimated the rain rate when compared to radar observations. This project trials a straight vertical profile to convert the Nimrod and the Gandolf surface rain rate data into the maximum in-cloud rain rate. This profile has been chosen after consulting studies carried out by the radar meteorology community, who have been trying to solve the opposite problem of how to obtain a surface rain rate from radar reflectivity measured above the



surface. Several authors conclude that when dealing with convective rain, the median vertical reflectivity profile is more vertically homogeneous than with stratiform rain due to the absence of the bright band (Smyth and Illingworth, 1998; Fabry, 1995; Gray et al, 2002). Gibson (2001) states that the current method of converting data uses an idealised profile taken from Kitchen et al (1994) in stratiform situations, but that "...pixels which are classified as strongly convective have no VPR correction applied other than the orographic correction." Smyth and Illingworth (1998) and Gray et al (2002) both concede that there is a large standard deviation in the vertical profile of convective rain due to the change with life stage. For example, in developing and mature thunderstorms the predominance of updraughts can result in a top-heavy reflectivity profile as shown by Hand (1996) and Zipser and Lutz (1994). However, Smyth and Illingworth state that these profiles are short lived because hail quickly descends by virtue of its large terminal velocity.

For the 3D extrapolation and 3D trended forecasts of rain rate from JCHMR the maximum is taken from the four levels and used as the maximum in-cloud rain rate. The ability to give more detailed information on the reflectivities aloft is a real advantage of these new forecast types.

### 3.5 Stage 3: Replicating pilot radar displays

Aircraft radar returns are the primary tool used by pilots to inform their decision as to whether to penetrate through or deviate around convective weather. Thus the rain rates obtained in Stage 2 are converted into a replicated pilot radar display showing black (<1mm/hr), green (1-4mm/hr), amber (4-12mm/hr) and red (>12mm/hr) pixels. Figure 3.5.1 shows an example replicated pilot display, overlaid with mid and low level air traffic sectors and the major airport locations.

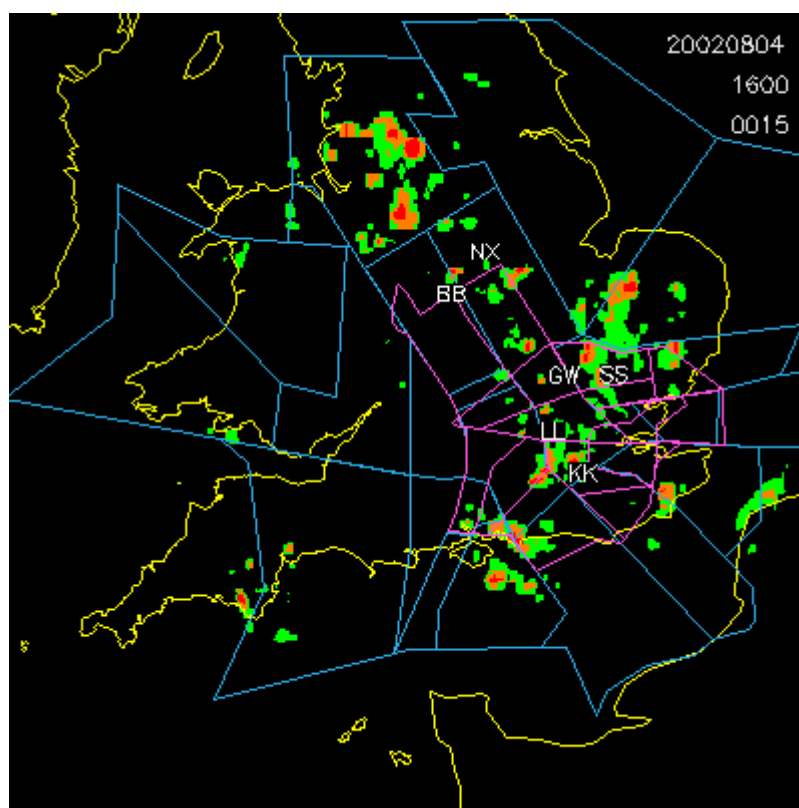


Figure 3.5.1: A replicated pilot display for 4 August 2002, using the 3D\_extrap rain rate forecast at a 15 minute lead time. The plot is overlaid with mid (blue) and low (pink) level air traffic sectors and the major airport locations: Birmingham (BB), East Midlands (NX), Luton (GW), Stanstead (SS), Heathrow (LL) and Gatwick (KK).

### 3.6 Stage 4: Classifying deviation pixels

Previous research (Rhoda and Pawlak, 2000; Watkin et al, 2002) and consultation with pilots (Rankin, 2001) has shown that both the intensity and coverage of convective weather, as shown by the radar returns, influence the type of deviation requested or carried out. Table 3.6.1 is used to calculate which pixels will cause which type of pilot action.

Replicated radar return of individual pixel	Type of convective activity in surrounding 70 by 70km area N=no deviation, S=small deviation, L=large deviation		
	Isolated (<20% non-black pixels)	Widespread and weak (>20% non-black and <20% amber-red pixels)	Widespread and strong (>20% non-zero and >20% amber-red pixels)
Black	N	N	N
Green	N	N	L
Amber-red	S	S	L

**Table 3.6.1: Using the replicated radar colour of the pixel and the amount of convective activity in the surrounding area to classify the pilot behaviour. In isolated or widespread and weak convection pilots will fly through green radar returns and make small deviations around amber or red returns, but if convection is widespread and strong they will try to avoid the whole area causing a large deviation.**

The following checks are then carried out to establish whether or not there is convection in the area:

1. Lightning check:

If the pixel is within 50km of where Nimrod lightning analysis/forecast data indicates a rate of greater than 0 strikes per minute then it is established that there is convection in the area.

2. Thunderstorm probability check:

Forecast data of the probability of there being a heavy thunderstorm is now available from Nimrod. Convective Available Potential Energy (CAPE), wind shear, precipitation rate and cloud top temperature are used as predictors in a Bayesian approach to produce an overall probability of there being a severe thunderstorm in a given 105 by 105km area (Hand, 2002). Each pixel in our forecast area is tested to see if it is covered by a >0% risk of a severe thunderstorm. If it is, it is established that there are indications of convection in the area.

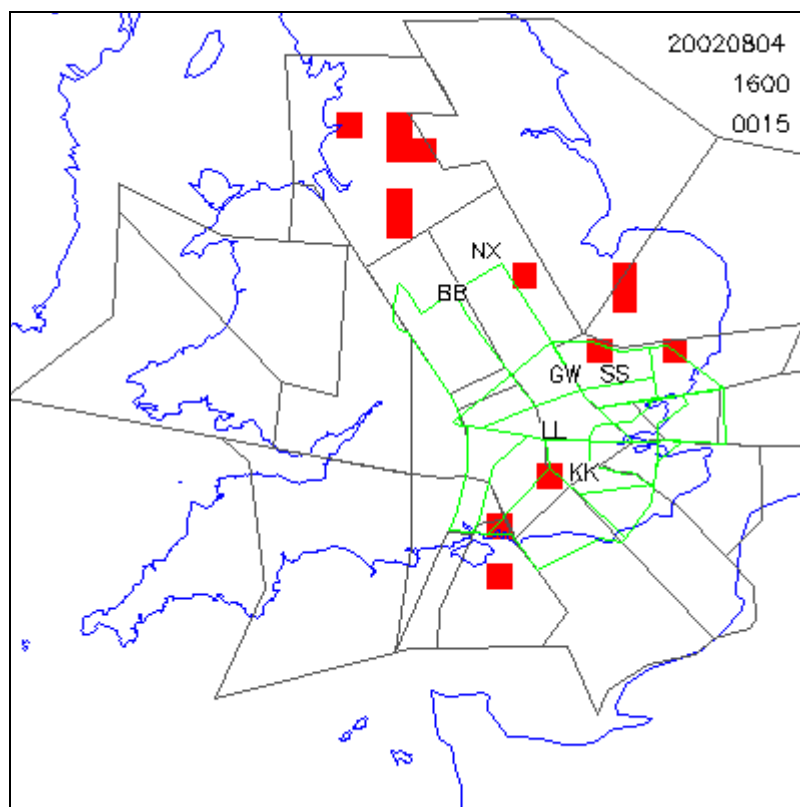
If neither condition is met then the pixel is classified as not causing a deviation, because pilots will not tend to deviate around dynamic rain.

### 3.7 Stage 5: Forecasting air traffic disruption

In the UK, significant disruption is caused to the controllers when 25-30% of aircraft are asking for a deviation, or holding in unfamiliar patterns (Layton, 2002). Disruption analyses/forecasts are produced at each of the four spatial resolutions of air space chosen to be tested, using the rules given in Table 3.7.1. It is more likely that these thresholds will be passed at the higher resolutions than at the lower resolution, since it requires less convection to satisfy these thresholds in a 10km area than in an 80km area. This is satisfactory because in a larger area the controller has more scope to re-route aircraft and limit disruption than if a small important area such as a stack is affected. Thus on some days with moderate but not severe thunderstorm activity, disruption may only be forecast at the higher resolutions. An example forecast of air traffic disruption is shown in Figure 3.7.1.

% Large deviation pixels	Operator	% Total deviation pixels	Class
$\leq 20$	AND	$\leq 30$	No disruption
$> 20$	OR	$> 30$	Disruption

**Table 3.7.1: Classifying disruption based on the percentage of pixels in a given size of air space that have been classified as causing a deviation.**



**Figure 3.7.1: A 20km resolution forecast of air traffic disruption due to thunderstorm activity for 4 August 2002, using the 3D\_extrap rain rate forecast at a 15 minute lead time. Disruption is marked in red, and is overlaid with mid (grey) and low (green) level air traffic sectors and the major airport locations: Birmingham (BB), East Midlands (NX), Luton (GW), Stanstead (SS), Heathrow (LL) and Gatwick (KK).**

### 3.8 Stage 6: Verification

The verification data used in the main study are the 3D rain rate analyses produced by Gandolf. The maximum rain rate is taken from the four analysis layers and processed in the same way as the forecast data so as to create a disruption/no disruption analysis. In the absence of reliable data of actual air traffic disruption, and given that the reflectivity aloft is the key factor causing planes to deviate from their flight path, the processed 3D rain analyses provide the best available surrogate for observed air traffic disruption.

The forecasts and analyses are compared and a 2x2 contingency table created for each forecast resolution and lead time. These tables are used to calculate the Kuipers Skill Score, Odds Ratio, Bias, Probability of Detection, Critical Success Index, and False Alarm Rate for each forecast type, lead time and spatial resolution. Table 3.8.1 provides a definition of these statistical measures while Appendix B gives an example contingency table and provides the formulas used in their calculation. The variety of statistical measures calculated is necessary because each one measures a different aspect of the forecast performance.

Measure	Definition	Numbers
Probability of detection (POD)	fraction of times disruption type was forecast when it had been observed i.e. hits / (hits + misses)	Perfect forecast =1 Completely wrong=0
Critical Success Index (CSI)	hit rate once 'disruption not forecast and not observed' values are removed	Perfect forecast =1 Completely wrong=0
False Alarm Rate (FAR)	fraction of times disruption type was forecast to occur and didn't occur	Perfect forecast =1 Completely wrong=0
Kuipers Skill Score (KSS)	a skill score calculated by comparing the accuracy of the forecasts with unbiased random forecasts	Perfect forecast =1 Random forecasts = 0 Forecast inferior to random if < 0
Log of Odds Ratio	the odds ratio is the ratio of the odds of making a hit to the odds of issuing a false alarm	When you take the log of the Odds Ratio: Completely wrong = 0 Well forecast = significantly > 0
Bias Ratio	the ratio of the number of times a disruption category was forecast to the number of times it was observed	Overforecast if >1 Underforecast if < 1

**Table 3.8.1: Definitions of the statistical forecast quality measures used in this project**

The supplementary study based on six cases is also verified against a 3D analysis of rain rate using the statistical measures described above. However, actual data of air traffic disruption due to thunderstorms, as recorded in "Terminal Control weather disruption reports" provided by Phil Layton (Manager Terminal Control Operations and Projects, LTCC), are used as an additional verification source (see section 6).

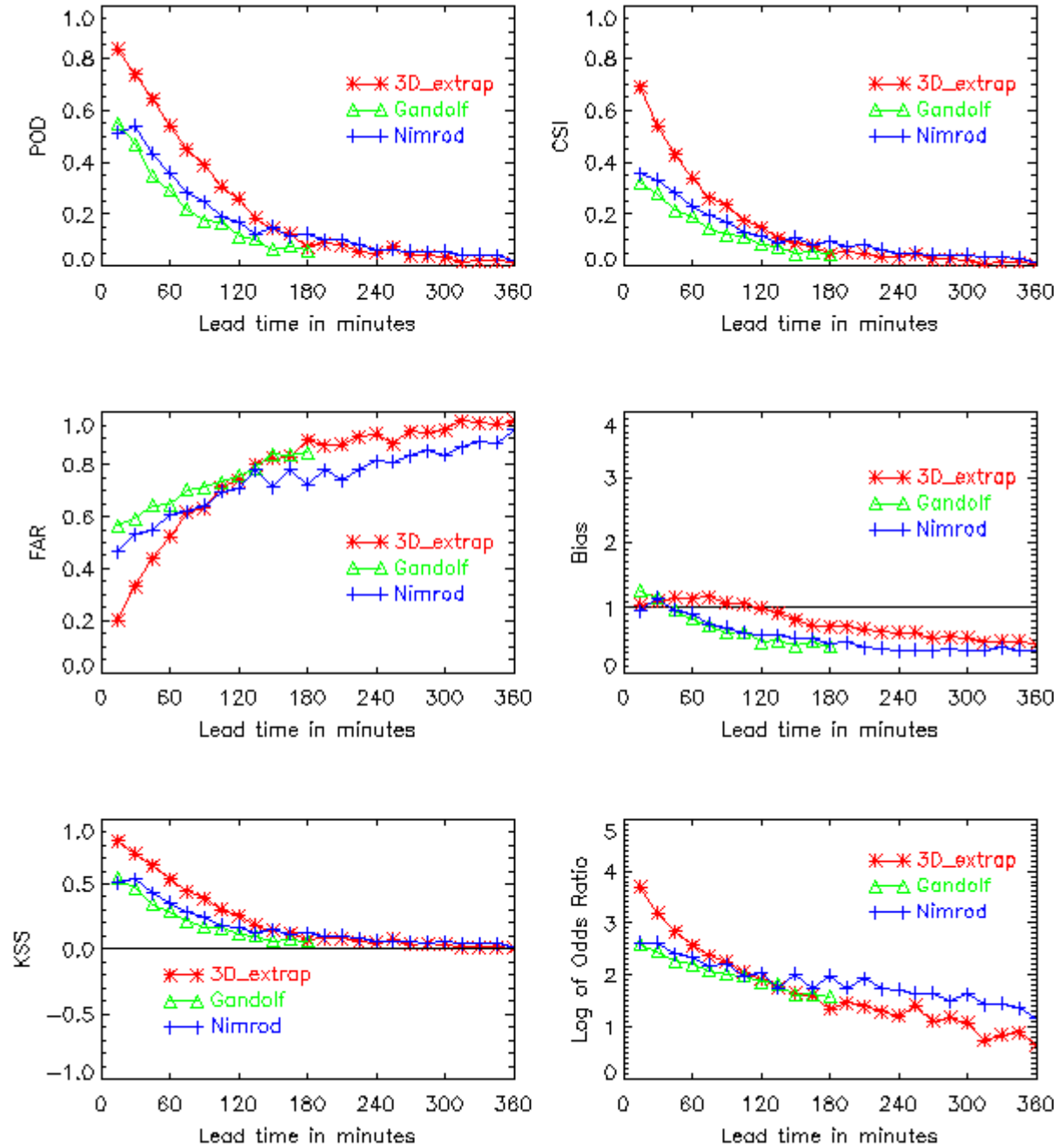
## 4. Comparison between forecast types

### 4.1 Results from the main study (using 22 sample days)

Appendices C1-C3 contain tables of the verification scores for the disruption forecasts produced from 3D rain rate extrapolation, Gandolf rain rate, and Nimrod rain rate forecast data, compared for each resolution over lead times from 15 to 360 minutes. In order to directly compare the quality of the different forecast types, Figure 4.1.1 shows the six statistical forecast verification measures for forecasts produced at the 40km resolution from each of the forecast tools.

The Kuipers Skill Score (KSS) is widely thought to provide one of the best overall estimates of the forecast skill. At the smallest lead time, 15 minutes, the 3D\_extrap forecast has a very high KSS of 0.85, indicating a near-perfect forecast. This very high score can be explained by the fact that the 3D\_extrap forecast is created from the analyses that it is then verified against, and hence performs very well at the small lead times. As the lead time increases, forecast skill decreases steadily before levelling out at around 0.1 at 180 minutes lead time, indicating that for lead times greater than three hours the forecasts show only slightly more skill than a random forecast. The Gandolf and Nimrod forecasts show similar patterns, but start from lower KSS values of 0.5-0.6 at 15 minutes lead time. This is because the rain rate forecast data in these forecast tools is not derived from the 3D rain rate analysis, but from a surface rain rate. For lead times of 150 minutes or more, the KSS of the Nimrod forecasts is equal to or better than the 3D\_extrap forecasts. This could reflect the input of model data to

the Nimrod forecasts, which begins to dominate the observation component of the system at this point.



**Figure 4.1.1: Statistical measures for the 40km resolution disruption forecasts for each of the forecast types, plotted against lead time in minutes.**

A similar pattern of results is apparent on the graph showing the Log of the Odds Ratio for each different forecast type. The 3D\_extrap forecast has a significantly higher odds ratio for the small lead times, but this declines rapidly such that lead times of over 120 minutes the forecasts produced from Gandolf and Nimrod data have a odds ratio of equal to or better than the 3D\_extrap forecast. This again reflects that the limits of the pure extrapolation method are reached around this time and the input of model data is desirable.

All three forecast types have a Bias of greater than 1 at the very small lead times, but this does not rise above 1.3 for any of the forecasts, indicating that they all overforecast the disruption, but at an acceptable level. The Gandolf and Nimrod forecasts drop below 1 to underforecast the disruption by a 45-minute lead time and decline steeply to reach a Bias of 0.5 by a 120-minute lead time. The 3D\_extrap forecast retains a Bias of close to the optimal 1 until a 120-minute lead time, where it also begins to underforecast the disruption.

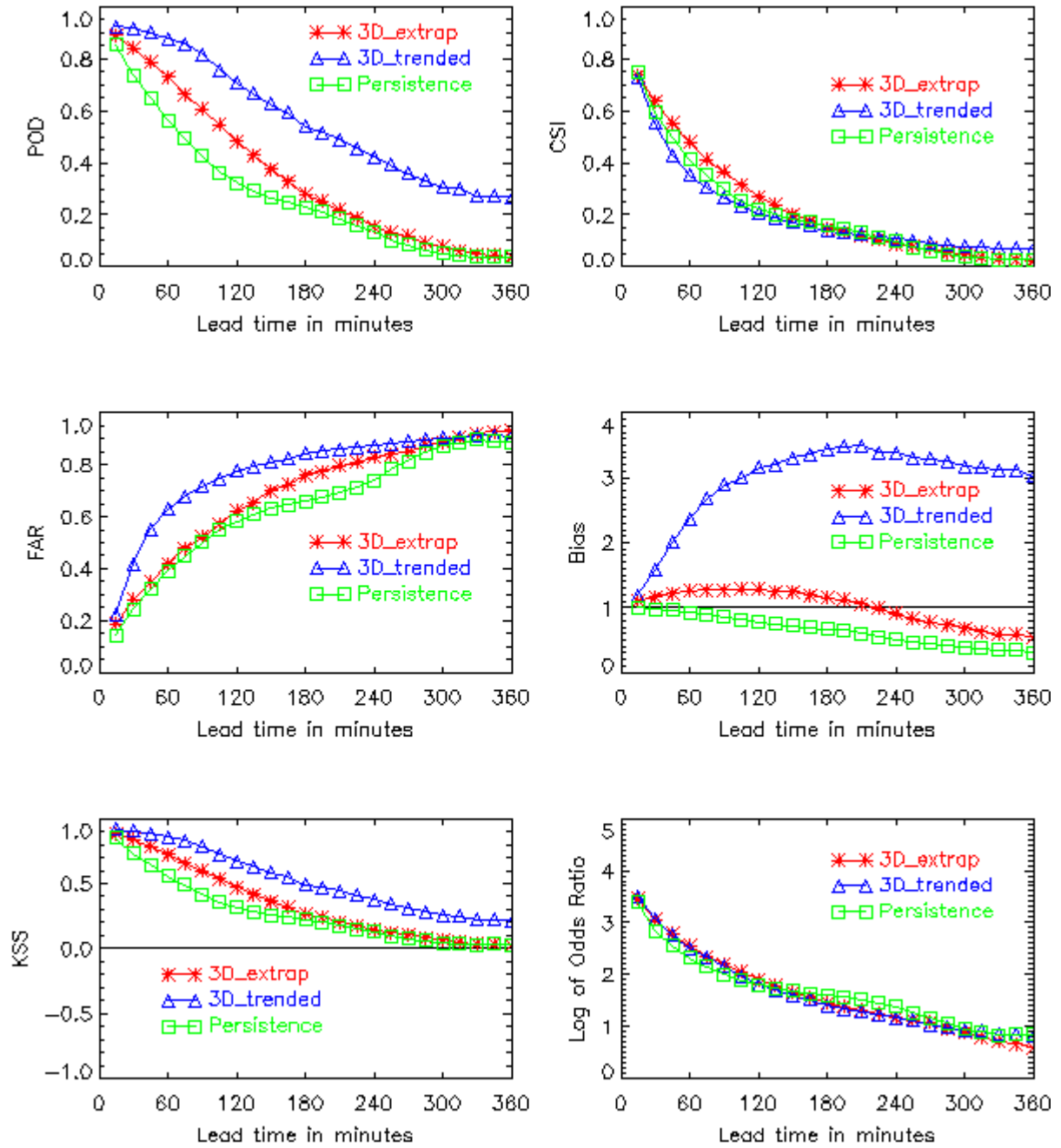
The POD and CSI show similar curve shapes to the KSS, with the 3D\_extrap forecast again starting from a higher level than the forecasts produced from Nimrod and Gandolf data. Again the curves show a steadily declining accuracy from 15 minutes to 180 minutes lead time, before levelling out, indicating that the limitations of the forecast have been reached and are no longer providing useful information. Forecasts produced from Gandolf and Nimrod data have a higher FAR up to 90 minutes lead time, but for longer lead times the Nimrod forecasts have a significantly lower FAR, again indicating the positive influence of the model input at these lead times.

The Gandolf results are largely similar to those shown by Hoad (1999), but show a slightly lower Bias and false alarm rate, likely to be due to the fewer number of times that the Object Orientated method is run in the current as compared to the earlier Gandolf version. The Nimrod KSS results are slightly better than those shown by Watkin et al (2002), possibly representing an improvement in the way that the Nimrod data is used in the forecasting scheme.

Given the assumption that the 3D analyses are the best available surrogate for observed air traffic disruption, the 3D\_extrap forecasts appear to be of significantly higher quality than the disruption forecasts prepared from Gandolf and Nimrod data for lead times of up to 120 minutes. This indicates that the forecasts produced using a surface rain rate and a straight vertical profile do not forecast the reflectivity aloft as well as the forecast produced from 3D rain rate data. The surface rain rate forecasts will not show the higher reflectivity aloft that is present during developing and mature thunderstorms. The initial identification of disruption will also be delayed for surface rain rates because the critical thresholds on the replicated pilot radar displays will be passed later at the surface than in the developing core of convection. The 3D rain rate forecast data is able to give a more accurate estimate of what the pilots will see in their aircraft radar when making the decision of whether or not to deviate around convective weather. Hence the 3D\_extrap forecast outperforms the Gandolf and Nimrod forecasts in terms of its ability to forecast air traffic disruption due to thunderstorm activity.

## **4.2 Results from the supplementary study (using 6 sample days)**

In the supplementary study, the 3D\_extrap disruption forecast is compared with two new types of disruption forecast, prepared from a linear-trended 3D rain rate extrapolation forecast (3D\_trended) and from a persistence forecast of 3D rain rate (Persistence) over six sample days. The full results are tabled in Appendices D1-D3. Figure 4.2.1 shows the six statistical forecast verification measures for forecasts produced on a 40km resolution from each of the forecast tools, so as to directly compare the forecast quality.



**Figure 4.2.1: Statistical measures for the 40km resolution disruption forecasts for each of the forecast types in the supplementary study, plotted against lead time in minutes.**

All three of the forecast types are derived in some way from the 3D rain rate analyses and thus at the shortest lead time, 15 minutes, they perform very similarly and very well in all of the verification statistics. As the lead time increases however, marked differences between the three appear. In terms of the KSS, the 3D\_trended rain rate forecast seems to be the best performer, followed by the 3D\_extrap forecast and then the Persistence forecast. The same order is present in the POD scores, where even at a 360-minute lead time the 3D\_trended forecast shows a POD of 0.25, when the other forecasts are very close to the worst possible score of zero. If these two measures are taken in isolation then the linear-trended 3D forecast appears to be the highest quality forecast. However, some important problems of this linear-trended approach are highlighted by the other statistical measures. The FAR graph shows that at most lead times the linear-trended forecast has the highest false alarm rate. The

graph of the Bias results also shows the tendency of the 3D\_trended forecast to overforecast the air traffic disruption due to thunderstorm activity very quickly as the lead time increases, reaching a Bias of 3 by 90 minutes. In contrast the non-trended 3D\_extrap forecast stays within  $\pm 0.5$  of the perfect value of 1 at all lead times, being slightly above 1 at lead times up to 210 minutes and below 1 thereafter. The Persistence forecast starts at 1 but steadily declines such that it is underforecasting at all lead times. Thus in terms of the Bias the 3D\_extrap forecast appears to be the highest quality forecast. This is also true for the Critical Success Index, which represents the hit rate once 'disruption not forecast and not observed' values are removed, and shows the 3D\_extrap forecast to be the most accurate forecast type at the crucial lead times of 0 to 180 minutes. These findings mirror the conclusions of previous work comparing steady-state extrapolation forecasts with intensity trended extrapolation forecasts (Wilson et al, 1998).

When weighing up which is the best forecast tool to be supplied to air traffic control the operational considerations of the controllers are just as important as the verification statistics. It may be that failing to receive a forecast of air traffic disruption due to thunderstorm activity and thus not applying restrictions to the air space, has more severe consequences than applying restrictions which turn out not to be necessary. Hence, the forecast that has the highest POD would be most useful, even if it had a high FAR. In this case it could be argued that the linear-trended forecast provides the greatest 'quality' to air traffic management. However, it could also be argued a large number of false alarms would lead the controllers to lose confidence in the forecast such that a correct forecast may not be acted upon. Too many false alarms could also prove very costly in terms of unnecessary delays. What is ideally needed is for air traffic management to carry out a cost-benefit analysis using the statistics presented in this report in order to establish the best type, lead time and spatial resolution of forecast to receive. However, in the absence of this information and after weighing up all of the pros and cons of the different forecast tools, the 3D\_extrap forecasts seem to be of the highest quality. Thus it is the 3D\_extrap forecasts that will be taken forward into the next two sections to examine the effect of spatial resolution on forecast skill and to see how well the disruption noted by the controllers at LTCC on these days was forecast.

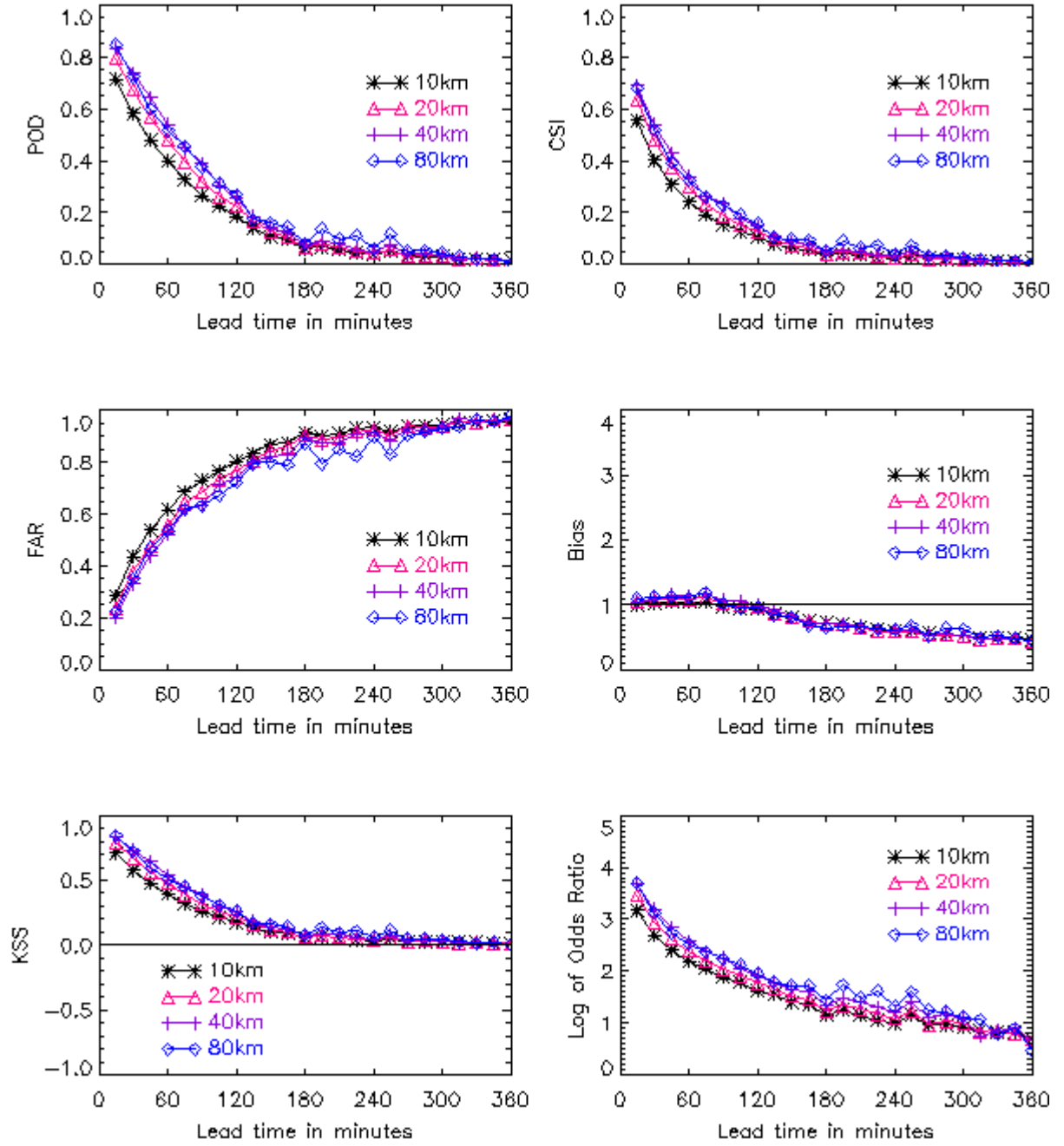
## 5. Effect of spatial resolution on forecast skill

Figure 5.1 shows the six statistical forecast verification measures for the 3D\_extrap forecasts, compared for each spatial resolution over lead times from 15 to 360 minutes. The 3D\_extrap forecasts are chosen for the analysis as they were shown in the previous section to exhibit the best overall verification scores. Similar graphs of all of the forecast types evaluated in the main and supplementary studies can be found in Appendices E and F respectively.

The Kuipers Skill Score and the Odds Ratio are greater for the lower spatial resolutions (40km and 80km), than the higher resolutions (20 km and 10km). This is to be expected, as it is easier to forecast whether or not there will be disruption within a larger area, than to specify exactly which 10km area will have that disruption. The curves for the 40km and 80km resolutions are quite close together, indicating that an optimal forecast area size may be reached around 40km.

The graphs of the Probability of Detection and Critical Success Index also indicate that forecast accuracy increases as the spatial resolution decreases up to 40km. The FAR is higher for the higher resolutions, which again indicates that the lower resolutions (40km and 80km) are of a better quality.





**Figure 5.1: Statistical measures of the 3D\_extrap disruption forecasts plotted for the four spatial resolutions against lead time in minutes.**

It is difficult to draw conclusions on the best forecast resolution because this depends on the operational needs of the air traffic controller as well as the forecast accuracy. It is likely that the controller would want the highest quality forecast possible, at the highest resolution possible, leaving a delicate trade-off between the two. The 40km forecast appears to be better quality than the 10 and the 20km resolution, and not significantly worse than the 80km resolution, which is why the 40km forecasts were used for the forecast type comparison in section 4.

## 6. Analysis of actual air traffic disruption data

For each of the six events analysed in the supplementary study, data of the actual air traffic disruption were available through the Terminal Control weather disruption reports. The entries in this log describe any problems caused by weather, including occurrences of weather avoidance by pilots, any Minimum Departure Interval (MDI) restrictions placed upon routes from the major airports, and any sector or landing regulations. Due to the very varied nature of the entries it was not possible to automatically classify how well the forecast had performed. Thus a subjective analysis of how well the information contained in the log has been picked up by the 3D extrapolation forecast was carried out, for each of six sample thunderstorm days. The results of one day, 09/08/2002, are analysed in full below, since they illustrate many of the general findings. The results of the other five sample days are contained in tabular form in Appendix G. Maps showing the locations of the places mentioned in the Terminal Control reports are contained in Appendix H.

### 09/08/2002

*There is a swirling mass of convection throughout much of the West of England from lunchtime to around 1900, causing widespread disruption for air traffic management. In West Berkshire the highest 30 minute fall of rain for 15 years was experienced, with the highest one-minute intensity being 144mm/hr at 1750 (Brugge, 2002).*

The first note in the Terminal Control Weather Disruption report is at 1206, when it states that there is weather avoidance present. It doesn't give a location for this weather disruption and hence it is difficult to analyse if the forecasts have accurately picked up the disruption. We are not forecasting any disruption over the London TMA (Terminal Manoeuvring Area), but some disruption due to thunderstorm activity is forecast in the Bristol and Lakes sectors up to about 2 hours ahead.

Then around 1325 disruption is felt in the immediate approach and departure routes from LL (London Heathrow). This is picked up in the 10km forecast at 1315 up to a 60-minute lead time, but not in the lower resolutions, because much of the surrounding area is clear and thus there is insufficient coverage of the larger areas to warrant a forecast of disruption. This shows that even if convective activity is very localised, it can still cause disruption if it is in an important location for air traffic (e.g. covering a hold, or a key approach or departure route), emphasising the need for higher resolution data in these key areas.

Traffic is stopped on the CPT (Compton, West Berkshire) route from 1330 onwards, with disruption in the CPT area being successfully picked up by the 10km, 20km and 40km forecasts up to 90 minutes ahead. At the same time LL BUZ (Buzad) and GW (London Luton) OLNEY routes are subject to release due to the weather. Disruption is forecast in these areas up to around 90 minutes ahead but only in the 10km resolution indicating that there was not sufficient convection to warrant a disruption forecast in the other resolutions. It should be remembered that the 10km resolution forecasts have been shown by section 5 to be of a lower quality than the lower resolutions. It may appear from this case study that the higher resolution forecasts are more successful at picking up the air traffic disruption as recorded in the disruption reports. This is misleading as it is largely due to the fact that it is more common for their required thresholds to be passed; a trait that will also lead to more false alarms. If required, the thresholds for the lower resolution forecasts could be reduced in the LTMA area so that the disruption noted here would also be forecast in these resolutions.

At 1410 a Gatwick arrival regulation is applied. No disruption is forecast at any lead time or resolution in the Gatwick area (even in the analysis), but this regulation could be caused by the severe weather problems being experienced out to the West in the Hurn, Berry Head and Bristol sectors. This is an example of the problems of trying to verify the forecasts on this

logbook information. It is sometimes difficult to tell from the entry exactly where the disruption is that has caused the regulation. The area that is experiencing the disruption might not be where the thunderstorms themselves are but may be feeling the knock on effects of convection elsewhere.

Also at 1410 departures at BB (Birmingham airport) and NX (East Midlands airport) are stopped due to the weather. We do forecast some disruption out to the West of BB which could have necessitated this move, but in the immediate area of these two airports no disruption was forecast at any lead time or spatial resolution because there were not the required indications of convection in the area.

At 1435 all sectors in the Terminal Control (TC) area have weather rates applied. Disruption is correctly forecast in the TC area in all spatial resolutions up to around 90 minutes ahead. After this we continue to forecast disruption at the 10 and 20km resolutions out to 180 minutes, and at lead times greater than this disruption is still forecast, but only out to the West of the TC area. By 1535 the sectors which had been subject to a regulation have this lifted by 10%. Our forecasts do not seem to indicate that there is any less disruption around than an hour earlier. This could be because the 3D\_extrap forecast is an extrapolation forecast and hence does not attempt to kill off cells.

At 1605 there is a general entry in the log stating that weather is circling around the Midlands, the west of the TMA and the Southwest. Our forecasts show this clearly up to 120 minutes lead time, but after this they overforecast the disruption. At 1630 the regulations are taken off the NE departures and the DAGGA sector. There is no disruption at any lead time or spatial resolution in these areas, indicating that the forecasts could have been used to plan ahead that it would be safe to take the regulations off.

By 1945 a log entry remarks that the weather is moving south leaving clearer weather for the aircraft. This movement to the South leaving clearer weather behind is picked up well by the forecasts, which only forecast disruption in the Berry Head sector and not any further north (up to a 150-minute lead time).

## **7. Discussion**

### **7.1 Influence of forecast type on forecast accuracy**

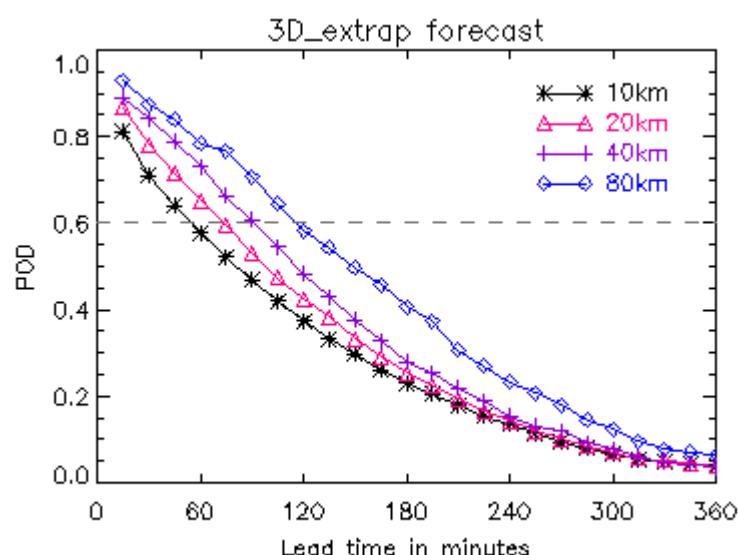
The disruption forecast produced using a 3D extrapolation forecast of rain rate has been shown to provide the highest overall quality of all the forecast types trialled in the main and supplementary studies, when verified against the processed 3D analyses. However, to produce the optimum forecast of air traffic disruption due to thunderstorm activity a combination of all of the forecast tools compared in this study is needed. It has been proven that it is essential to be able to predict the reflectivity aloft in addition to the surface rain rate. However, the value of model input in addition to a pure extrapolation-based forecast has also been shown. Thus an operationally-produced 3D rain rate from a system such as Gandolf would be ideal. Thunderstorm forecasts for air traffic management will also benefit from further research into the early detection of thunderstorms, as is currently being carried out in the joint GANDOLF-RDT project carried out by the Met Office and Meteo-France. Further research of how storm development or decaying can be built into the forecast tools would also be of benefit.

### **7.2 Impact of spatial resolution and lead time on forecast accuracy**

The results indicate that the quality of the disruption forecasts produced is higher for lower resolution forecast areas than for higher resolutions. This is to be expected as it is easier to forecast that a larger area will be affected by thunderstorms somewhere within it, than to

say where exactly within these larger areas the disruption will be located. Forecast accuracy also declines with increasing lead-time for all of the forecast types. The speed of this drop off varies between forecast types, with persistence forecasts dropping off most rapidly.

One interesting line of research in thunderstorm forecasting for air traffic control is the complimentary relationship between forecast lead time, the nature of the decisions made at that time, and forecast accuracy, as described by Boldi et al (2002) and Vigeant-Langlois and Jansman Jr. (2002). Higher resolution forecasts are required for reactive and tactical uses at short lead times when these can be produced with greater accuracy. The box radius can be increased for greater lead times so as to maintain accuracy, and still be of use for strategic planning. Comparing the accuracy of Met Office disruption forecasts over several lead times and spatial resolutions allows an evaluation of the practicality of such an approach. Suppose, for example, air traffic controllers found that they could make efficient use of a disruption forecast if it had a POD of above 0.6. Figure 7.2.1 shows that using the 3D\_extrap forecast, they could receive a high resolution 10km forecast up to a 45 minute lead time, a 20km forecast at 60 and 75 minutes, 40km at 90 minutes and an 80km forecast from 105 to 120 minutes. The graph here shows the results of the supplementary study where the differences between the four resolutions are more pronounced than in the main study. This is because the main study included random days where there was little or no disruption, when all four forecasts faired equally well.



**Figure 7.2.1: Probability of detection of the 3D\_extrap forecast from the supplementary study. The dashed line shows that a POD of 0.6 or higher can be maintained for higher lead times by increasing the box size of the forecast area.**

At present, the different spatial resolutions all use the same threshold of pixels causing a deviation (see table 3.2) to decide whether or not to classify the air space as "disrupted". This means that it is more common for disruption to be forecast in the higher resolution forecasts, since there has to be a lot of convection around to cause the threshold to be passed in airspaces of 80km. This is acceptable however, because if the area under consideration is the size of an air space sector, the controller has more scope to re-route aircraft than if a small area such as a stack is affected. An additional way of utilising this fact would be to provide variable sized disruption forecasts for different areas. For example, in the areas immediately surrounding the London airports, or the holding stacks, the forecast resolution would be higher, whereas in areas where air traffic is less busy a lower resolution forecast could be provided.

### 7.3 General trends and limitations of the forecasts used in this study

If the forecasts of air traffic disruption due to thunderstorm activity are to be used by air traffic control it is essential that they are aware of the limitations involved in the forecasts:

#### *Effect of type and stage of thunderstorm event on forecast accuracy*

What is hidden by the statistics is the relationship between the type, and stage, of the thunderstorm event and the Met Office's ability to forecast it. For example it is easier to predict organised thunderstorm activity, for example embedded convection within a front, rather than isolated and unorganised convection. If an area experiences convective activity for, say, a six-hour period, then the forecasts will be of a higher quality for forecast times in the middle of the event than for those times nearer the start or the end of the event. This is because all of the rain rate forecast tools used in this report (including Gandolf and Nimrod) rely heavily on the extrapolation of observations at small lead times. They will not initiate the convection, and thus will tend to under-predict the disruption at the start of an event. Once an event is underway the forecasts are better able to forecast the movement of the storms and thus predict which areas will experience disruption when. At the end of an event the forecasts will tend to over-predict the disruption. The case study analysis contained in section 6 and Appendix G illustrates these general patterns.

#### *Fluctuation of controller workload and effect on air traffic disruption*

At present, none of the disruption forecasts take account of the fluctuation of workload for the controllers, with the key thresholds in the forecasting scheme remaining constant and independent of the time of day, day of week, or area covered by the forecast. From analysing the disruption reports received from LTCC it has become apparent that in fact whether or not "disruption" is reported to have occurred depends on several factors in addition to the weather. For example, the same size of thunderstorm event over the London Terminal Manoeuvring Area will cause more disruption than if it were over an airway which was not working at capacity. Similarly, weekdays, due to the higher number of flights, are more prone to disruption than weekends, as are certain key times of the day. This fluctuating controller workload could of course be incorporated into the forecasting scheme if required.

#### *Limitations of the forecast data used in this study*

It is difficult to compare the forecast skill of the Nimrod and the Gandolf forecasts due to the different resolutions of the original forecast data (5km for Nimrod and 2km for Gandolf), and the fact that the Gandolf forecasts only ran up to 180 minutes during the summer of 2002. The disruption forecasts produced from Gandolf data used the surface rain rate produced by the Gandolf system that was operational over the summer of 2002. At the end of September a new version of Gandolf was implemented to extend the maximum range of its precipitation nowcasts out to 6 hours, and incorporate an optic-flow based extrapolation scheme. This scheme was shown to be superior to the existing advection algorithms implemented within Gandolf during cases of embedded convection (Bowler and Pierce, 2002). Thus the quality of disruption forecasts produced from the new Gandolf system is likely to be higher than the statistics presented in this report. The 3D extrapolation forecast and linear-trended extrapolation forecasts were both produced using the optic flow advection scheme and thus benefited from the enhanced accuracy.

A further problem is that on the 7<sup>th</sup> August a new version of the Met Office Unified Model (UM5) became operational. Changes were made to several areas of the model, which may have affected the quality of forecasts produced from each method before and after this date, since all the methods use model data at some point. Unfortunately the implementation date fell so that approximately half of our sample days fell before and half after. Thus it was decided that the problem of the discontinuity of data was unavoidable.

### *Limitations of the verification data used in this study*

Two different types of verification data have been used in this study. The first, disruption analyses produced from 3D rain rate analyses, are a useful data source for testing the overall ability of the various different forecasts to forecast the location and extent of thunderstorm activity. However, the data are not fully independent, as some of the forecast types use the 3D analyses as the base from which the forecasts are produced. Furthermore, the 3D analyses are converted into disruption analyses using the same method that is used to produce the forecasts, and thus the verification process assumes that we have a reliable and accurate method for transferring a rain rate forecast to a forecast of air traffic disruption. In reality, there is currently insufficient data available to assess how well our replicated pilot radar displays match the actual displays, and how well our forecasting scheme has characterised the relationship between the in-flight radar displays and the air traffic disruption. Thus more reliable data from in-flight pilot displays or of actual air traffic disruption is required in order to make definite assertions regarding the relative quality of the forecast types and forecast resolutions.

The second type of verification data used are the Terminal Control Weather Disruption Reports. The case study analysis carried out in section 6 and Appendix G has proved a useful first attempt at checking how often the forecasts have correctly forecast the actual disruption experienced by controllers by LTCC, and the need for the various regulations that were put into force. There were, however, a number of problems with using this form of analysis data, which restricted the conclusions that could be drawn from the results:

- Due to the complex and varied nature of the entries, the data could only be analysed subjectively.
- The entries into the log were made by the duty control manager, which varies between shift. Thus the entries themselves are subjective, and will vary in their content from controller to controller.
- The report contains usually refers to weather, and not specifically to thunderstorms and thus in some instances the required action may be due to another type of weather.
- A regulation may be required on a route or a sector either because of a thunderstorm in that area, or due to the knock on affects of convection elsewhere. The location of the thunderstorms that have caused the disruption is often not stated in the report, making it difficult to assess how well the forecasts have done.

A possible compromise between the two sets of analyses would be to examine flight track data to see where aircraft have deviated from their planned route in order to avoid thunderstorms.

## **8. Summary and Conclusions**

This paper has investigated the quality of forecasts of air traffic disruption due to thunderstorm activity produced from several different types of rain rate forecast data. The forecasts are verified against disruption analyses prepared from 3D rain rate analyses, which are acting as a surrogate for actual air traffic disruption. The results from the main study show that the disruption forecasts produced by extrapolating a 3D rain rate analysis outperform the forecasts produced from both Gandolf and Nimrod rain rate forecasts for lead times out to 2 hours. This proves the importance of forecasting the reflectivity aloft in addition to the surface rain rate. The value of model input, in addition to a pure extrapolation-based forecast is also apparent, as the Gandolf and Nimrod forecast perform better than the 3D extrapolation at lead times of greater than two hours.

When compared with a linear-trended 3D extrapolation forecast and a persistence forecast in a supplementary study, the 3D extrapolation forecast is still shown to be of the overall highest quality. Although the linear-trended 3D forecast performs better in some of the

statistical measures, its high Bias and false alarm rate would restrict its operational use, emphasising the importance of considering a range of verification statistics when deciding on the optimum forecast tool.

The highest quality forecast was compared to Terminal Control weather disruption reports for several case studies. Despite several limitations, this analysis provided a useful first step to verifying the forecasts against actual data of air traffic disruption.

Forecast accuracy has been shown to increase as the spatial resolution of the forecast decreases, with the differences between the spatial resolution most apparent in the 3D forecasts and in the supplementary study. There is some evidence in the main study that an optimum size of forecast area may be reached at 40km. Forecast quality is generally very good at the short lead times, especially for the 3D extrapolation forecasts, but tails off with increasing lead time. Without information on the relative costs or a hit and a false alarm from NATS it is difficult to give a lead time up to which the forecasts would be of use to air traffic management. However, it is likely that the forecasts would be of limited use up to lead times of around 120 minutes, after which the limits of the forecasts appear to have been reached. It is of course true that thunderstorms are by their very nature, very difficult to forecast and thus forecast quality drops off rapidly in the first hour of forecasting. It may never be possible to provide good quality forecasts of air traffic disruption due to thunderstorm activity produced from rain rate data out to more than two hours. At lead times greater than this a probabilistic forecast such as that described by Hand (2002) may be of more value.

The possibility of incorporating the fluctuating controller workload over time and area into the forecasting scheme should also be considered, in order to reflect more fully the outcome for air traffic management of a particular storm. Both the thresholds used and the box size of forecast received could vary with time and area depending on the controller workload at each time and location. For example, in an area that is important to the operations of air traffic control such as a stack or the immediate departure and arrival routes, even localised thunderstorms can cause a problem and thus a high resolution forecast may be required, even if this raises the false alarm rate. The box size of the forecast could also vary with lead time to reflect the complimentary relationship between forecast lead time, the nature of the decisions made at that time, and forecast accuracy.

This study took the decision to produce a forecast of where thunderstorms were of sufficient coverage and intensity to cause pilot deviations and hence air traffic disruption. In doing so we have tried to add value to a standard weather product by incorporating the outcomes of the weather for Air Traffic Management into the forecasting and verification scheme. By converting the forecast data into the four disruption forecast resolutions tested here, errors in the intensity of each pixel are smoothed and thus a higher quality forecast product is available to longer lead times. A different option would be to supply air traffic control with a forecast in the form of a replicated radar display at the maximum resolution possible (2km). The statistics produced in this report would still be of use as a general guide to how well the convective activity within a particular air space size are forecast.

An important conclusion from this study is that in order to forecast thunderstorms for air traffic control, the ideal rain rate input would be an operationally-produced 3D rain rate forecast from a system such as Gandolf. In the limited number of cases where life-cycle modelling has been shown to be of benefit, the 3D rain rate could be produced taking into account the cell characteristics, and in all other cases a 3D extrapolation forecast of rain rate would be used. Model input could also be incorporated. This forecast should be able to take advantage of any improvements in life-cycle modelling which may result from the joint Gandolf-RDT work to be carried out in 2003.

## 9. Recommendations

1. In light of this report, air traffic management (ATM) need to consider their user requirements for thunderstorm forecasts. If the preferred option is a forecast of air traffic disruption due to thunderstorm activity then it is important that a cost-benefit analysis be carried out using the verification statistics presented here. This will allow the air traffic controllers to choose the combination of forecast type, spatial resolution and lead time which offers them the maximum benefit. ATM also need to decide if they would like the fluctuating controller workload to be incorporated into the forecasting scheme. If the relative cost of a hit, miss and false alarm can be evaluated, this information could also be used to tune the model used to generate the disruption forecasts.
2. A 3D rain rate forecast should be developed and operationally produced from GANDOLF. This would combine inputs from a 3D extrapolation method, life cycle modelling using cell characteristics, and NWP data, and is likely to prove the best forecast tool for providing any type of thunderstorm forecast to air traffic control. This forecast would then benefit from the ongoing improvements to the GANDOLF system.
3. Any future work would benefit enormously from the routine collection of data of actual air traffic disruption due to thunderstorm activity by LTCC and LACC. This data would ideally be produced in a standardised and objective format, and would help to clarify the relationship between the replicated pilot displays and actual air traffic disruption.
4. Flight track data is suggested as a possible alternative source of verification data in future studies. It could be used to assess where aircraft have deviated around thunderstorms and caused disruption, and also provide an interesting line of research into pilot behaviour around thunderstorms.



## 10. References

- Blair, K., 2001: Personal communication
- Blair, K., 2002: Developing a Business Case for Weather Radar. *NATS General Research Programme Bright Idea 93*, NATS.
- Brugge, R., 2002: <http://www.met.rdg.ac.uk/~brugge/diary2002.html#0208>
- Boldi, R.A., Wolfson, M.M., Dupree, W.J., Johnson, R.J., Theriault, K.E., Forman, B.E. and Wilson, C.A., 2002: An automated, operational two hour convective weather forecast for the corridor integrated weather. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 116-119
- Bowler, N. and Pierce, C.E., 2002, Development of a short-range QPF algorithm based on optic flow techniques. *Forecasting Research Technical Report No. 390*, Met Office.
- Fabry, F., 1995: Vertical profiles of reflectivity and precipitation intensity. *III International Symposium on hydrological applications of weather radars*, 344-350
- Gibson, M., 2001: Report on current vertical profile of reflectivity correction methods. *Forecasting Research Technical Report No. 346*, Met Office.
- Gray, W.R., Uddstrom, M.J. and Larsen, H.R., 2002: Radar surface rainfall estimates using a typical shape function approach to correct for the variations in the vertical profile of reflectivity. *International Journal of Remote Sensing*, **23**(12), 2489-2504
- Hand, W.H., 1996: An object-orientated technique for nowcasting heavy showers and thunderstorms. *Meteorological Applications*, **3**, 31-41
- Hand, W.H., 2002: Operational documentation of method to provide probabilities of severe convective weather. *Internal Met Office website*. Accessed September 2002.
- Hauf, T. and Sasse, M., 2002: The impact of thunderstorms on landing traffic at Frankfurt airport (Germany) - a case study. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 160-161
- Hoad, D.J., 1999: Verification of short range thunderstorm forecasts using radar data to assess their benefit to the aviation community. *Forecasting Research Technical Report No. 284*, Met Office.
- Hudson, H.R. and Foss, F.P., 2002: The Collaborative Convective Forecast Product from the Aviation Weather Center's perspective. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 73-76
- Kitchen, M., Brown, R. and Davies, R.G., 1994: Real-time correlations of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation. *Quarterly Journal of the RMS*, **120**
- Kulesa, G.J., Kirchoffer, P.J., Pace, D.J., Fellner, W.L., Sheets, J.E., Travers, V.S., 2002: New weather products developed by the federal aviation administration's aviation weather research program. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 18-19
- Layton, P., 2002: Personal communication
- <http://www.nats.co.uk/services/latcc.html>. Accessed July 2002.
- Pierce, C.E. and Cooper, A.M., 2000: Comparison of the performance of 2 km resolution Object-Orientated Model and Nimrod advection precipitation nowcast schemes. *Forecasting Research Technical Report No. 350*, Met Office.
- Rankin, J., 2001: Personal communication
- Rhoda, D.A. and Pawlak, M.L., 1999: The thunderstorm penetration/deviation decision in the terminal area. *8<sup>th</sup> Conference on Aviation, Range, and Aerospace Meteorology*, 308-312
- Rhoda, D.A., Boorman, B.G., Bouchard, E.A., Isaminger, M.A., and Pawlak, M.L., 2000: Commercial aircraft encounters with thunderstorms in the Memphis terminal airspace. *9<sup>th</sup> Conference on Aviation, Range, and Aerospace Meteorology*, 37-42
- Rhoda, D.A., Kocab, E.A. and Pawlak, M.L., 2002: Aircraft encounters with thunderstorms in enroute vs. terminal airspace above Memphis, Tennessee. *10<sup>th</sup> Conference on Aviation, Range, and Aerospace Meteorology*, 162-165

- Smyth, T.J. and Illingworth, A.J., 1998: Radar estimates of rainfall rates at the ground in bright band and non-bright band events. . *Quarterly Journal of the RMS*, **124**, 2417-2434
- Souders, C.G. and Showalter, R.C., 2002: An update to the FAA's National Airspace System weather architecture. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 1-4
- Souders, C.G. and Showalter, R.C., 2002: The role of ITWS in the National Airspace System modernization: an update. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 5-8
- Soul, K.M., Archibald, E.J., Hardaker, P.J. and Hounsell, A., 2002: Using the GANDOLF system as a tool to aid the forecasting of lightning strikes. *Meteorological Applications*, **9**(2), 229-238
- Thorn, T., 1997: The air pilot's manual 2: Aviation law and meteorology. Airline Publishing, Shrewsbury, England
- Vigeant-Langlois, L. and Jansman, Jr., R.J., 2002: Trajectory-based performance assessment of aviation weather information. *10<sup>th</sup> Conference on Aviation, Range and Aerospace Meteorology*, 136-139
- Watkin, H.A., Scott, T.R. and Hoad, D.J., 2002: Spatial and lead-time accuracies of thunderstorm forecasts for air traffic control. *Forecasting Research Technical Report No. 379, Met Office*.
- Wilson, J.W., Cook, N.A., Mueller, C.K., Sun, J. and Dixon, M.: Nowcasting thunderstorms: a status report. *Bulletin of the American Meteorological Society*, **10**, 2078-2099
- Wilson M., (1997) Meteorology for pilots. Airline Publishing, Shrewsbury, England
- Zipser, E.J. and Lutz, K.R.: 1994: The vertical profile of radar reflectivity of convective cells: a strong indicator of storm intensity and lightning probability? *Monthly Weather Review*, **122**(8), 1751-1759

## 11. Acknowledgements

The authors of this report would like to thank the following:

- Neill Bowler of the Joint Centre for Hydro-Meteorological Research (JCHMR) for providing the data for the 3D extrapolation and 3D trended rain rate forecasts
- Clive Pierce, also of JCHMR, for his help and guidance throughout the study
- Phil Layton, Manager of Terminal Control Operations and Projects at LTCC, for providing the very useful Terminal Control weather disruption reports
- Katherine Blair of NATS for her advice on gaining data of air traffic disruption;
- Tania Scott and Bob Lunnon for their help and guidance during the course of the study

## 12. List of acronyms

Acronym	Explanation
ATC	Air Traffic Control
BB	Birmingham airport
BIG	Refers to the hold at Biggin Hill to the SE of London
BNN	Bovingdon hold to the NNW of LL
BPK	Northeast route to CLN (Clacton sector)
BUZ	BUZAD - West of Stanstead
CDP	Convective Diagnosis Project
CLN	Clacton sector
COW	COWLY sector
CPT	Compton - West Berkshire
CSI	Critical Success Index
DAGGA	DAGGA sector
DTY	Daventry sector
DVR	Dover sector
ESSEX	Essex sector
FAR	False Alarm Rate
GANDOLF	Generating Nowcasts for the Deployment of Operational Land-based Flood forecasts
GW	London Luton airport
JCHMR	Joint Centre for Hydro-Meteorological Research
KK	London Gatwick airport
KSS	Kuipers Skill Score
LAM	Lambourne sector to the NE of LHR (also a hold)
LL	London Heathrow airport
LACC	London Area Control Centre
LTCC	London Terminal Control Centre
LTMA	London Terminal Manoeuvring Area
MDI	Minimum Departure Intervals on a route (eg. 1/4 is 1 per 4 mins)
NATS	National Air Traffic Services
NWS	US National Weather Service
NX	East Midlands airport
OLNEY	Outbound route from Luton
POD	Probability of detection
RDT	Rapid Developing Thunderstorm product
SAM	Southampton
SS	London Stanstead airport
TC	Terminal Control
TMA	Terminal Manoeuvring Area
TSF	Target sector flow
UM5	Met Office Unified Model version 5
WEL	TC WELIN sector - Northamptonshire
WOB	Northerly departure from Heathrow

## Appendix A: List of sample days

A list of the 22 days analysed in this report, consisting of 16 thunderstorm days, picked out using sferics data, and six random days. The six thunderstorm events used in the additional study are in bold text.

Date	Thunderstorm or random day
11/07/2002	Thunderstorm
20/07/2002	Thunderstorm
22/07/2002	Random
28/07/2002	Thunderstorm
<b>29/07/2002</b>	<b>Thunderstorm</b>
<b>30/07/2002</b>	<b>Thunderstorm</b>
31/07/2002	Thunderstorm
03/08/2002	Thunderstorm
<b>04/08/2002</b>	<b>Thunderstorm</b>
05/08/2002	Thunderstorm
<b>07/08/2002</b>	<b>Thunderstorm</b>
08/08/2002	Thunderstorm
<b>09/08/2002</b>	<b>Thunderstorm</b>
10/08/2002	Thunderstorm
12/08/2002	Random
<b>15/08/2002</b>	<b>Thunderstorm</b>
19/08/2002	Thunderstorm
23/08/2002	Thunderstorm
26/08/2002	Random
02/09/2002	Random
09/09/2002	Random
16/09/2002	Random

## Appendix B: Contingency table and formulae for verification statistics

Below is an example of a 2 x 2 contingency table. Each box gives the number of times a particular combination of observed and forecast disruption level occurs. For example 'a' represents the number of times that air traffic disruption due to thunderstorm activity was forecast and observed.

		Observed		Total
		Air traffic disruption due to thunderstorms	No disruption	
Forecast	Air traffic disruption due to thunderstorms	a	b	a+b
	No disruption	c	d	c+d
Total		a+c	b+d	n=a+b+c+d

a = number of hits (correctly forecast disruption events)

b = number of false alarms (incorrectly forecast no disruption events)

c = number of misses (incorrectly forecast disruption events)

d = number of correct rejections (correctly forecast no disruption events)

Forecast quality measures can be calculated from the above table using the following equations:

$$KSS = \frac{ad - bc}{(a + c)(b + d)}$$

$$Odds \ Ratio = LOG\left(\frac{ad}{bc}\right)$$

$$Bias = \frac{a + b}{a + c}$$

$$POD = \frac{a}{a + c}$$

$$CSI = \frac{a}{a + b + c}$$

$$FAR = \frac{b}{a + b}$$

## Appendix C: Statistical measure scores for the main study

### Appendix C1: 3D\_extrap forecast

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 80 km	15	0.848	0.681	0.224	1.092	0.847	3.687
	30	0.724	0.519	0.352	1.118	0.722	3.124
	45	0.6	0.393	0.467	1.125	0.597	2.741
	60	0.517	0.325	0.534	1.107	0.513	2.522
	75	0.455	0.262	0.617	1.188	0.451	2.377
	90	0.376	0.229	0.631	1.02	0.373	2.242
	105	0.313	0.191	0.672	0.954	0.309	2.116
	120	0.268	0.158	0.723	0.967	0.264	1.952
	135	0.175	0.105	0.794	0.847	0.171	1.781
	150	0.163	0.099	0.8	0.813	0.159	1.714
	165	0.143	0.093	0.788	0.675	0.14	1.719
	180	0.08	0.051	0.874	0.636	0.077	1.429
	195	0.139	0.091	0.792	0.668	0.136	1.736
	210	0.098	0.063	0.851	0.656	0.094	1.48
	225	0.112	0.074	0.822	0.631	0.109	1.623
	240	0.062	0.04	0.898	0.609	0.059	1.308
	255	0.116	0.075	0.828	0.677	0.113	1.603
	270	0.052	0.035	0.902	0.526	0.049	1.246
	285	0.054	0.034	0.916	0.643	0.051	1.219
	300	0.045	0.028	0.929	0.628	0.041	1.118
	315	0.032	0.022	0.936	0.495	0.029	1.071
	330	0.019	0.013	0.963	0.519	0.016	0.773
	345	0.021	0.015	0.956	0.479	0.019	0.899
	360	0.008	0.006	0.981	0.451	0.006	0.464
Spatial resolution: 40 km	15	0.834	0.69	0.201	1.044	0.833	3.696
	30	0.736	0.539	0.332	1.101	0.734	3.182
	45	0.644	0.429	0.438	1.144	0.641	2.839
	60	0.54	0.34	0.522	1.13	0.537	2.574
	75	0.449	0.261	0.616	1.168	0.445	2.364
	90	0.391	0.233	0.633	1.064	0.387	2.252
	105	0.305	0.174	0.712	1.057	0.301	2.044
	120	0.259	0.149	0.739	0.992	0.255	1.918
	135	0.185	0.107	0.798	0.916	0.181	1.772
	150	0.146	0.087	0.822	0.817	0.142	1.649
	165	0.125	0.078	0.829	0.727	0.121	1.617
	180	0.075	0.046	0.894	0.706	0.071	1.351
	195	0.088	0.055	0.875	0.703	0.085	1.463
	210	0.082	0.052	0.876	0.662	0.079	1.405
	225	0.059	0.038	0.905	0.622	0.056	1.29
	240	0.051	0.033	0.917	0.616	0.048	1.208
	255	0.075	0.049	0.879	0.619	0.072	1.407
	270	0.039	0.026	0.927	0.539	0.036	1.118
	285	0.042	0.028	0.924	0.551	0.039	1.172
	300	0.035	0.023	0.935	0.533	0.032	1.074
	315	0.015	0.01	0.968	0.468	0.012	0.752
	330	0.019	0.013	0.959	0.475	0.017	0.845
	345	0.02	0.014	0.958	0.482	0.018	0.897
	360	0.012	0.009	0.972	0.436	0.01	0.66

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 20 km	15	0.795	0.637	0.238	1.043	0.794	3.451
	30	0.674	0.479	0.377	1.082	0.672	2.938
	45	0.567	0.375	0.474	1.078	0.564	2.622
	60	0.48	0.299	0.557	1.083	0.476	2.387
	75	0.397	0.232	0.641	1.105	0.393	2.209
	90	0.319	0.188	0.687	1.02	0.315	2.033
	105	0.263	0.152	0.734	0.988	0.259	1.906
	120	0.223	0.129	0.767	0.959	0.218	1.766
	135	0.167	0.098	0.807	0.864	0.162	1.669
	150	0.126	0.075	0.843	0.802	0.122	1.514
	165	0.105	0.064	0.86	0.749	0.101	1.443
	180	0.068	0.042	0.903	0.704	0.064	1.233
	195	0.078	0.048	0.888	0.699	0.074	1.334
	210	0.068	0.043	0.895	0.649	0.064	1.258
	225	0.05	0.032	0.916	0.592	0.046	1.158
	240	0.043	0.028	0.926	0.583	0.04	1.084
	255	0.057	0.037	0.902	0.585	0.054	1.237
	270	0.031	0.021	0.941	0.53	0.028	0.963
	285	0.035	0.023	0.934	0.535	0.032	1.031
	300	0.033	0.022	0.937	0.521	0.03	0.996
	315	0.02	0.014	0.958	0.473	0.017	0.825
	330	0.023	0.016	0.952	0.476	0.019	0.86
	345	0.018	0.012	0.963	0.48	0.015	0.778
	360	0.015	0.01	0.967	0.443	0.012	0.685
Spatial resolution: 10 km	15	0.712	0.555	0.285	0.996	0.711	3.164
	30	0.582	0.402	0.435	1.03	0.58	2.691
	45	0.481	0.308	0.539	1.043	0.477	2.409
	60	0.398	0.243	0.616	1.038	0.394	2.185
	75	0.33	0.192	0.685	1.047	0.326	2.046
	90	0.266	0.155	0.728	0.977	0.261	1.883
	105	0.223	0.129	0.765	0.948	0.218	1.783
	120	0.184	0.105	0.803	0.931	0.178	1.623
	135	0.141	0.082	0.835	0.851	0.136	1.547
	150	0.107	0.064	0.865	0.796	0.103	1.397
	165	0.097	0.058	0.872	0.753	0.092	1.372
	180	0.062	0.037	0.914	0.72	0.057	1.144
	195	0.069	0.042	0.902	0.701	0.065	1.242
	210	0.058	0.037	0.91	0.651	0.054	1.151
	225	0.043	0.028	0.929	0.613	0.039	1.054
	240	0.04	0.026	0.934	0.601	0.036	1.001
	255	0.051	0.033	0.915	0.598	0.047	1.142
	270	0.034	0.022	0.939	0.555	0.03	0.958
	285	0.032	0.021	0.942	0.549	0.028	0.956
	300	0.03	0.02	0.943	0.535	0.026	0.919
	315	0.022	0.015	0.954	0.488	0.019	0.84
	330	0.022	0.015	0.955	0.484	0.018	0.815
	345	0.021	0.014	0.958	0.495	0.018	0.817
	360	0.017	0.012	0.964	0.467	0.013	0.695

## Appendix C2: Gandolf forecast

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 80 km	15	0.557	0.305	0.598	1.385	0.554	2.539
	30	0.462	0.259	0.63	1.248	0.458	2.37
	45	0.376	0.23	0.627	1.01	0.373	2.306
	60	0.303	0.203	0.621	0.799	0.301	2.236
	75	0.245	0.164	0.669	0.74	0.243	2.171
	90	0.179	0.124	0.711	0.619	0.177	2.018
	105	0.171	0.117	0.733	0.643	0.169	1.971
	120	0.124	0.092	0.737	0.472	0.122	1.902
	135	0.108	0.077	0.784	0.5	0.106	1.83
	150	0.074	0.054	0.837	0.456	0.073	1.622
	165	0.083	0.06	0.824	0.468	0.081	1.652
	180	0.086	0.064	0.8	0.429	0.084	1.73
	195	-	-	-	-	-	-
	210	-	-	-	-	-	-
	225	-	-	-	-	-	-
	240	-	-	-	-	-	-
	255	-	-	-	-	-	-
	270	-	-	-	-	-	-
	285	-	-	-	-	-	-
	300	-	-	-	-	-	-
	315	-	-	-	-	-	-
	330	-	-	-	-	-	-
	345	-	-	-	-	-	-
	360	-	-	-	-	-	-
Spatial resolution: 40 km	15	0.548	0.321	0.564	1.257	0.545	2.591
	30	0.471	0.28	0.592	1.155	0.468	2.452
	45	0.346	0.214	0.642	0.964	0.343	2.262
	60	0.292	0.19	0.647	0.825	0.289	2.194
	75	0.217	0.144	0.701	0.727	0.215	2.086
	90	0.174	0.121	0.714	0.608	0.172	2.011
	105	0.167	0.114	0.733	0.625	0.165	1.981
	120	0.115	0.085	0.756	0.47	0.113	1.863
	135	0.105	0.076	0.786	0.489	0.103	1.828
	150	0.065	0.049	0.84	0.409	0.064	1.613
	165	0.077	0.055	0.837	0.476	0.076	1.627
	180	0.062	0.047	0.845	0.403	0.061	1.588
	195	-	-	-	-	-	-
	210	-	-	-	-	-	-
	225	-	-	-	-	-	-
	240	-	-	-	-	-	-
	255	-	-	-	-	-	-
	270	-	-	-	-	-	-
	285	-	-	-	-	-	-
	300	-	-	-	-	-	-
	315	-	-	-	-	-	-
	330	-	-	-	-	-	-
	345	-	-	-	-	-	-
	360	-	-	-	-	-	-



	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 20 km	15	0.558	0.323	0.565	1.283	0.554	2.554
	30	0.458	0.268	0.607	1.163	0.454	2.364
	45	0.356	0.214	0.65	1.018	0.353	2.207
	60	0.285	0.179	0.676	0.88	0.282	2.081
	75	0.218	0.138	0.726	0.794	0.215	1.99
	90	0.17	0.112	0.752	0.685	0.167	1.88
	105	0.153	0.1	0.777	0.685	0.15	1.822
	120	0.11	0.077	0.797	0.54	0.107	1.711
	135	0.098	0.067	0.822	0.549	0.096	1.677
	150	0.064	0.045	0.869	0.485	0.061	1.467
	165	0.074	0.051	0.862	0.535	0.071	1.499
	180	0.06	0.043	0.87	0.46	0.057	1.445
	195	-	-	-	-	-	-
	210	-	-	-	-	-	-
	225	-	-	-	-	-	-
	240	-	-	-	-	-	-
	255	-	-	-	-	-	-
	270	-	-	-	-	-	-
	285	-	-	-	-	-	-
	300	-	-	-	-	-	-
	315	-	-	-	-	-	-
	330	-	-	-	-	-	-
	345	-	-	-	-	-	-
	360	-	-	-	-	-	-
Spatial resolution: 10 km	15	0.533	0.309	0.576	1.258	0.529	2.49
	30	0.419	0.242	0.636	1.15	0.415	2.256
	45	0.321	0.188	0.688	1.028	0.317	2.096
	60	0.252	0.154	0.717	0.891	0.248	1.956
	75	0.201	0.124	0.755	0.821	0.198	1.896
	90	0.149	0.096	0.789	0.707	0.146	1.755
	105	0.136	0.086	0.812	0.724	0.133	1.703
	120	0.1	0.067	0.832	0.595	0.097	1.582
	135	0.086	0.057	0.855	0.591	0.083	1.546
	150	0.063	0.043	0.882	0.535	0.06	1.393
	165	0.066	0.044	0.882	0.562	0.063	1.397
	180	0.051	0.035	0.9	0.507	0.048	1.294
	195	-	-	-	-	-	-
	210	-	-	-	-	-	-
	225	-	-	-	-	-	-
	240	-	-	-	-	-	-
	255	-	-	-	-	-	-
	270	-	-	-	-	-	-
	285	-	-	-	-	-	-
	300	-	-	-	-	-	-
	315	-	-	-	-	-	-
	330	-	-	-	-	-	-
	345	-	-	-	-	-	-
	360	-	-	-	-	-	-

## Appendix C3: Nimrod forecast

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 80 km	15	0.541	0.345	0.513	1.112	0.538	2.569
	30	0.575	0.345	0.536	1.239	0.572	2.643
	45	0.424	0.266	0.584	1.018	0.42	2.346
	60	0.396	0.25	0.595	0.978	0.393	2.384
	75	0.271	0.177	0.662	0.8	0.268	2.098
	90	0.269	0.176	0.664	0.799	0.266	2.173
	105	0.2	0.131	0.726	0.729	0.197	1.925
	120	0.164	0.111	0.747	0.649	0.162	1.939
	135	0.112	0.078	0.796	0.547	0.109	1.712
	150	0.142	0.101	0.743	0.552	0.14	1.936
	165	0.106	0.074	0.8	0.529	0.104	1.697
	180	0.134	0.105	0.673	0.41	0.133	2.08
	195	0.094	0.07	0.787	0.441	0.092	1.726
	210	0.104	0.083	0.714	0.366	0.103	1.979
	225	0.082	0.061	0.806	0.424	0.08	1.668
	240	0.06	0.047	0.826	0.343	0.058	1.676
	255	0.053	0.041	0.852	0.359	0.051	1.506
	270	0.037	0.029	0.889	0.336	0.036	1.436
	285	0.041	0.031	0.892	0.382	0.039	1.343
	300	0.06	0.047	0.814	0.321	0.058	1.703
	315	0.053	0.041	0.839	0.329	0.051	1.542
	330	0.052	0.038	0.873	0.41	0.05	1.498
	345	0.041	0.032	0.879	0.341	0.039	1.388
	360	0.015	0.011	0.958	0.358	0.013	0.953
Spatial resolution: 40 km	15	0.511	0.354	0.464	0.954	0.509	2.623
	30	0.539	0.333	0.533	1.154	0.536	2.615
	45	0.432	0.283	0.55	0.96	0.429	2.408
	60	0.357	0.23	0.607	0.909	0.355	2.339
	75	0.283	0.194	0.62	0.744	0.28	2.18
	90	0.247	0.171	0.641	0.689	0.245	2.206
	105	0.187	0.131	0.694	0.613	0.185	1.981
	120	0.168	0.119	0.709	0.577	0.166	2.028
	135	0.126	0.087	0.781	0.574	0.123	1.754
	150	0.153	0.111	0.712	0.529	0.151	2.013
	165	0.114	0.081	0.783	0.525	0.112	1.743
	180	0.124	0.094	0.722	0.444	0.122	1.977
	195	0.102	0.075	0.781	0.465	0.1	1.742
	210	0.103	0.079	0.739	0.393	0.101	1.928
	225	0.083	0.064	0.778	0.372	0.081	1.737
	240	0.062	0.049	0.812	0.329	0.061	1.722
	255	0.065	0.051	0.81	0.339	0.063	1.641
	270	0.056	0.044	0.832	0.335	0.055	1.656
	285	0.053	0.04	0.853	0.357	0.051	1.5
	300	0.056	0.043	0.838	0.346	0.055	1.632
	315	0.045	0.035	0.868	0.341	0.043	1.435
	330	0.044	0.033	0.887	0.393	0.043	1.44
	345	0.041	0.031	0.884	0.35	0.039	1.365
	360	0.023	0.018	0.933	0.346	0.022	1.178

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 20 km	15	0.557	0.365	0.485	1.082	0.553	2.569
	30	0.525	0.317	0.554	1.177	0.521	2.509
	45	0.421	0.265	0.584	1.012	0.417	2.28
	60	0.338	0.215	0.629	0.91	0.335	2.23
	75	0.258	0.167	0.677	0.798	0.254	1.997
	90	0.218	0.141	0.714	0.762	0.215	1.988
	105	0.178	0.117	0.747	0.704	0.175	1.802
	120	0.156	0.103	0.767	0.672	0.153	1.833
	135	0.122	0.081	0.807	0.631	0.119	1.623
	150	0.128	0.087	0.789	0.609	0.126	1.764
	165	0.098	0.066	0.832	0.585	0.095	1.539
	180	0.1	0.071	0.805	0.512	0.098	1.708
	195	0.079	0.055	0.845	0.513	0.077	1.486
	210	0.086	0.062	0.818	0.472	0.084	1.663
	225	0.067	0.048	0.853	0.456	0.065	1.454
	240	0.058	0.042	0.868	0.437	0.056	1.482
	255	0.062	0.046	0.851	0.412	0.059	1.455
	270	0.048	0.035	0.889	0.433	0.046	1.391
	285	0.053	0.038	0.877	0.427	0.05	1.352
	300	0.051	0.037	0.882	0.433	0.049	1.417
	315	0.041	0.031	0.895	0.392	0.039	1.266
	330	0.045	0.031	0.907	0.48	0.042	1.291
	345	0.038	0.028	0.907	0.406	0.035	1.202
	360	0.029	0.02	0.938	0.459	0.026	1.095

Spatial resolution: 10 km	15	0.479	0.334	0.475	0.913	0.476	2.484
	30	0.44	0.284	0.557	0.992	0.437	2.41
	45	0.34	0.226	0.598	0.844	0.336	2.166
	60	0.271	0.179	0.653	0.781	0.268	2.118
	75	0.204	0.139	0.694	0.666	0.2	1.9
	90	0.177	0.121	0.723	0.638	0.174	1.923
	105	0.139	0.094	0.772	0.61	0.136	1.69
	120	0.122	0.083	0.792	0.583	0.119	1.73
	135	0.095	0.066	0.825	0.544	0.092	1.527
	150	0.097	0.068	0.818	0.535	0.095	1.645
	165	0.072	0.051	0.856	0.502	0.069	1.416
	180	0.076	0.055	0.835	0.463	0.074	1.583
	195	0.061	0.044	0.866	0.455	0.058	1.374
	210	0.067	0.049	0.845	0.429	0.065	1.548
	225	0.051	0.038	0.874	0.406	0.049	1.336
	240	0.044	0.033	0.887	0.387	0.042	1.378
	255	0.044	0.033	0.878	0.363	0.042	1.313
	270	0.039	0.029	0.898	0.381	0.037	1.324
	285	0.038	0.029	0.896	0.368	0.036	1.231
	300	0.036	0.027	0.904	0.379	0.035	1.289
	315	0.032	0.025	0.907	0.348	0.03	1.171
	330	0.031	0.023	0.925	0.419	0.029	1.163
	345	0.03	0.022	0.918	0.363	0.027	1.107
	360	0.021	0.015	0.948	0.41	0.019	0.983

## Appendix D: Statistical measure scores for the supplementary study

### Appendix D1: 3D\_extrap forecast

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 80 km	15	0.929	0.813	0.133	1.071	0.926	3.736
	30	0.873	0.694	0.228	1.13	0.869	3.196
	45	0.838	0.608	0.312	1.218	0.832	2.903
	60	0.783	0.53	0.379	1.26	0.775	2.639
	75	0.768	0.499	0.413	1.309	0.759	2.551
	90	0.707	0.445	0.454	1.296	0.697	2.377
	105	0.644	0.402	0.484	1.248	0.634	2.227
	120	0.582	0.353	0.528	1.231	0.57	2.082
	135	0.542	0.329	0.545	1.192	0.53	2.002
	150	0.497	0.292	0.586	1.199	0.484	1.884
	165	0.456	0.267	0.608	1.162	0.443	1.803
	180	0.406	0.232	0.65	1.16	0.392	1.686
	195	0.373	0.216	0.66	1.098	0.359	1.633
	210	0.306	0.181	0.694	1	0.293	1.515
	225	0.27	0.163	0.708	0.926	0.258	1.458
	240	0.232	0.141	0.735	0.875	0.22	1.372
	255	0.207	0.129	0.744	0.809	0.195	1.329
	270	0.179	0.113	0.764	0.76	0.167	1.262
	285	0.146	0.093	0.794	0.708	0.134	1.164
	300	0.124	0.081	0.808	0.645	0.113	1.103
	315	0.097	0.063	0.848	0.637	0.085	0.958
	330	0.077	0.051	0.869	0.587	0.066	0.872
	345	0.07	0.046	0.879	0.578	0.058	0.812
	360	0.064	0.043	0.886	0.562	0.053	0.78
Spatial resolution: 40 km	15	0.89	0.739	0.187	1.094	0.887	3.449
	30	0.842	0.635	0.279	1.169	0.838	3.06
	45	0.789	0.554	0.349	1.212	0.783	2.788
	60	0.73	0.48	0.416	1.25	0.722	2.546
	75	0.664	0.412	0.48	1.277	0.655	2.343
	90	0.607	0.364	0.523	1.271	0.597	2.196
	105	0.546	0.315	0.573	1.279	0.535	2.042
	120	0.483	0.269	0.623	1.279	0.471	1.894
	135	0.43	0.238	0.652	1.238	0.418	1.789
	150	0.377	0.202	0.698	1.248	0.364	1.654
	165	0.327	0.176	0.724	1.186	0.314	1.559
	180	0.279	0.148	0.759	1.156	0.265	1.443
	195	0.253	0.136	0.774	1.118	0.239	1.389
	210	0.218	0.118	0.795	1.062	0.205	1.313
	225	0.189	0.105	0.808	0.986	0.176	1.25
	240	0.153	0.087	0.831	0.906	0.141	1.161
	255	0.132	0.078	0.841	0.825	0.12	1.113
	270	0.118	0.071	0.848	0.78	0.107	1.078
	285	0.091	0.055	0.876	0.733	0.08	0.958
	300	0.077	0.048	0.886	0.681	0.067	0.901
	315	0.06	0.038	0.904	0.629	0.05	0.8
	330	0.047	0.031	0.919	0.587	0.038	0.704
	345	0.043	0.028	0.927	0.582	0.033	0.654
	360	0.036	0.024	0.934	0.542	0.026	0.589

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 20 km	15	0.867	0.715	0.197	1.08	0.863	3.307
	30	0.78	0.574	0.315	1.14	0.775	2.813
	45	0.715	0.484	0.4	1.192	0.708	2.536
	60	0.652	0.42	0.458	1.202	0.643	2.337
	75	0.596	0.366	0.513	1.224	0.586	2.177
	90	0.532	0.315	0.565	1.222	0.521	2.018
	105	0.474	0.273	0.608	1.21	0.463	1.887
	120	0.426	0.24	0.644	1.197	0.414	1.779
	135	0.384	0.215	0.672	1.171	0.372	1.689
	150	0.333	0.182	0.713	1.158	0.319	1.565
	165	0.292	0.159	0.742	1.13	0.278	1.47
	180	0.254	0.136	0.772	1.112	0.239	1.373
	195	0.225	0.122	0.79	1.07	0.211	1.304
	210	0.193	0.106	0.809	1.013	0.179	1.228
	225	0.166	0.093	0.824	0.943	0.152	1.166
	240	0.143	0.082	0.839	0.886	0.13	1.104
	255	0.119	0.07	0.855	0.82	0.106	1.031
	270	0.101	0.06	0.87	0.778	0.089	0.966
	285	0.085	0.051	0.884	0.734	0.073	0.895
	300	0.07	0.044	0.895	0.671	0.059	0.832
	315	0.057	0.036	0.911	0.638	0.046	0.738
	330	0.051	0.033	0.915	0.599	0.041	0.71
	345	0.045	0.029	0.923	0.581	0.035	0.656
	360	0.039	0.026	0.929	0.55	0.029	0.602
Spatial resolution: 10 km	15	0.812	0.658	0.224	1.047	0.809	3.083
	30	0.709	0.512	0.352	1.095	0.703	2.612
	45	0.639	0.426	0.439	1.14	0.631	2.357
	60	0.577	0.364	0.503	1.163	0.568	2.169
	75	0.522	0.316	0.556	1.175	0.512	2.023
	90	0.469	0.275	0.601	1.175	0.457	1.892
	105	0.421	0.241	0.639	1.165	0.409	1.783
	120	0.374	0.211	0.675	1.153	0.362	1.676
	135	0.332	0.185	0.705	1.125	0.319	1.582
	150	0.297	0.163	0.734	1.116	0.284	1.493
	165	0.261	0.143	0.759	1.083	0.248	1.409
	180	0.231	0.126	0.784	1.067	0.217	1.327
	195	0.204	0.112	0.802	1.034	0.19	1.257
	210	0.18	0.1	0.818	0.987	0.166	1.196
	225	0.153	0.086	0.834	0.923	0.14	1.126
	240	0.135	0.078	0.845	0.872	0.122	1.077
	255	0.114	0.067	0.861	0.82	0.102	1.008
	270	0.095	0.056	0.877	0.773	0.083	0.931
	285	0.08	0.048	0.89	0.722	0.068	0.866
	300	0.066	0.041	0.902	0.672	0.055	0.796
	315	0.054	0.034	0.916	0.643	0.043	0.708
	330	0.047	0.03	0.921	0.603	0.037	0.67
	345	0.043	0.028	0.926	0.583	0.033	0.632
	360	0.04	0.026	0.928	0.557	0.03	0.61

## Appendix D2: 3D\_trended forecast

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 80 km	15	0.971	0.817	0.162	1.159	0.968	4.023
	30	0.951	0.635	0.344	1.45	0.943	3.359
	45	0.934	0.488	0.495	1.848	0.918	2.954
	60	0.914	0.4	0.584	2.201	0.892	2.674
	75	0.899	0.345	0.641	2.507	0.872	2.5
	90	0.867	0.297	0.689	2.786	0.835	2.285
	105	0.822	0.272	0.711	2.849	0.786	2.097
	120	0.786	0.24	0.743	3.061	0.746	1.946
	135	0.744	0.224	0.758	3.071	0.702	1.825
	150	0.713	0.209	0.771	3.118	0.669	1.736
	165	0.676	0.195	0.784	3.135	0.63	1.642
	180	0.648	0.179	0.801	3.266	0.6	1.561
	195	0.593	0.162	0.818	3.258	0.543	1.442
	210	0.579	0.161	0.818	3.182	0.529	1.418
	225	0.554	0.154	0.824	3.145	0.504	1.369
	240	0.525	0.147	0.831	3.108	0.474	1.314
	255	0.508	0.144	0.833	3.05	0.457	1.283
	270	0.439	0.123	0.854	3.003	0.387	1.155
	285	0.397	0.112	0.865	2.932	0.344	1.077
	300	0.368	0.106	0.87	2.836	0.316	1.024
	315	0.36	0.103	0.873	2.843	0.306	0.997
	330	0.342	0.098	0.88	2.852	0.288	0.954
	345	0.339	0.098	0.879	2.807	0.283	0.941
	360	0.34	0.098	0.878	2.798	0.285	0.945
Spatial resolution: 40 km	15	0.924	0.729	0.224	1.19	0.92	3.511
	30	0.918	0.553	0.418	1.577	0.908	3.072
	45	0.901	0.428	0.551	2.007	0.885	2.753
	60	0.878	0.353	0.629	2.366	0.856	2.509
	75	0.856	0.304	0.68	2.672	0.83	2.341
	90	0.817	0.266	0.718	2.892	0.786	2.153
	105	0.757	0.234	0.747	2.996	0.724	1.961
	120	0.707	0.205	0.776	3.157	0.671	1.808
	135	0.666	0.189	0.792	3.197	0.628	1.702
	150	0.625	0.17	0.811	3.308	0.584	1.593
	165	0.594	0.158	0.823	3.355	0.551	1.516
	180	0.538	0.138	0.843	3.437	0.493	1.39
	195	0.516	0.13	0.852	3.484	0.469	1.336
	210	0.491	0.123	0.859	3.488	0.444	1.287
	225	0.456	0.116	0.865	3.39	0.408	1.224
	240	0.423	0.107	0.875	3.383	0.375	1.159
	255	0.394	0.101	0.881	3.303	0.345	1.106
	270	0.36	0.092	0.891	3.296	0.311	1.037
	285	0.332	0.085	0.898	3.251	0.283	0.981
	300	0.306	0.079	0.904	3.178	0.257	0.928
	315	0.301	0.078	0.905	3.165	0.25	0.907
	330	0.273	0.071	0.912	3.113	0.222	0.844
	345	0.273	0.071	0.912	3.116	0.221	0.84
	360	0.27	0.072	0.911	3.03	0.218	0.833

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 20 km	15	0.904	0.702	0.242	1.192	0.899	3.336
	30	0.868	0.526	0.429	1.52	0.858	2.82
	45	0.842	0.41	0.556	1.897	0.826	2.516
	60	0.806	0.337	0.633	2.195	0.785	2.276
	75	0.778	0.292	0.681	2.44	0.753	2.124
	90	0.738	0.257	0.717	2.612	0.709	1.97
	105	0.697	0.229	0.745	2.738	0.665	1.841
	120	0.643	0.2	0.775	2.855	0.608	1.697
	135	0.608	0.184	0.792	2.919	0.571	1.606
	150	0.565	0.165	0.811	2.994	0.525	1.5
	165	0.529	0.151	0.826	3.042	0.488	1.417
	180	0.491	0.135	0.842	3.115	0.448	1.328
	195	0.47	0.128	0.85	3.136	0.425	1.275
	210	0.446	0.121	0.857	3.131	0.4	1.226
	225	0.419	0.114	0.865	3.094	0.373	1.174
	240	0.388	0.105	0.874	3.07	0.341	1.112
	255	0.362	0.099	0.88	3.02	0.315	1.061
	270	0.34	0.092	0.888	3.031	0.292	1.011
	285	0.303	0.083	0.898	2.968	0.255	0.936
	300	0.279	0.077	0.904	2.908	0.231	0.882
	315	0.264	0.073	0.909	2.896	0.215	0.841
	330	0.257	0.071	0.911	2.875	0.207	0.82
	345	0.255	0.071	0.91	2.847	0.205	0.816
	360	0.255	0.072	0.909	2.815	0.205	0.812

Spatial resolution: 10 km	15	0.854	0.645	0.275	1.178	0.849	3.072
	30	0.803	0.477	0.46	1.487	0.793	2.585
	45	0.767	0.375	0.577	1.812	0.751	2.304
	60	0.734	0.312	0.648	2.088	0.713	2.105
	75	0.703	0.269	0.696	2.315	0.677	1.96
	90	0.657	0.234	0.734	2.472	0.628	1.812
	105	0.619	0.209	0.76	2.583	0.587	1.701
	120	0.577	0.185	0.786	2.699	0.543	1.588
	135	0.543	0.17	0.802	2.747	0.508	1.507
	150	0.507	0.153	0.821	2.828	0.469	1.416
	165	0.477	0.141	0.834	2.87	0.437	1.344
	180	0.446	0.128	0.848	2.941	0.404	1.268
	195	0.423	0.119	0.858	2.976	0.379	1.21
	210	0.4	0.112	0.866	2.978	0.356	1.16
	225	0.374	0.105	0.873	2.938	0.329	1.109
	240	0.354	0.099	0.879	2.933	0.309	1.064
	255	0.328	0.092	0.886	2.882	0.282	1.012
	270	0.304	0.085	0.894	2.876	0.258	0.956
	285	0.276	0.078	0.902	2.829	0.23	0.896
	300	0.251	0.071	0.909	2.769	0.204	0.837
	315	0.238	0.067	0.914	2.773	0.19	0.797
	330	0.228	0.065	0.917	2.756	0.18	0.77
	345	0.23	0.066	0.915	2.72	0.182	0.775
	360	0.227	0.065	0.916	2.694	0.178	0.762

## Appendix D3: Persistence forecast

	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 80 km	15	0.89	0.81	0.101	0.99	0.888	3.596
	30	0.8	0.682	0.178	0.973	0.796	3.049
	45	0.718	0.583	0.244	0.95	0.713	2.729
	60	0.645	0.513	0.284	0.9	0.639	2.534
	75	0.582	0.455	0.324	0.862	0.577	2.387
	90	0.52	0.399	0.368	0.823	0.514	2.247
	105	0.461	0.351	0.404	0.774	0.455	2.125
	120	0.421	0.317	0.438	0.75	0.415	2.044
	135	0.381	0.284	0.472	0.721	0.374	1.954
	150	0.365	0.272	0.483	0.706	0.358	1.924
	165	0.352	0.263	0.49	0.689	0.345	1.899
	180	0.316	0.234	0.526	0.667	0.309	1.817
	195	0.298	0.219	0.547	0.658	0.291	1.768
	210	0.252	0.187	0.577	0.595	0.245	1.685
	225	0.208	0.155	0.621	0.549	0.201	1.58
	240	0.177	0.133	0.649	0.503	0.17	1.51
	255	0.141	0.108	0.689	0.455	0.135	1.407
	270	0.116	0.088	0.732	0.432	0.109	1.304
	285	0.092	0.07	0.773	0.406	0.086	1.195
	300	0.071	0.055	0.809	0.373	0.065	1.084
	315	0.051	0.039	0.853	0.346	0.045	0.93
	330	0.045	0.035	0.863	0.325	0.039	0.888
	345	0.037	0.029	0.879	0.308	0.031	0.809
	360	0.034	0.027	0.882	0.292	0.029	0.793
Spatial resolution: 40 km	15	0.856	0.749	0.143	0.998	0.853	3.388
	30	0.735	0.594	0.244	0.973	0.731	2.836
	45	0.651	0.499	0.32	0.957	0.646	2.552
	60	0.562	0.413	0.392	0.924	0.556	2.311
	75	0.494	0.353	0.447	0.893	0.487	2.154
	90	0.425	0.298	0.502	0.853	0.418	2.004
	105	0.363	0.251	0.551	0.808	0.356	1.877
	120	0.322	0.222	0.583	0.772	0.315	1.799
	135	0.291	0.201	0.608	0.743	0.284	1.731
	150	0.266	0.183	0.631	0.721	0.258	1.675
	165	0.248	0.171	0.644	0.697	0.24	1.637
	180	0.227	0.158	0.659	0.667	0.22	1.596
	195	0.211	0.147	0.674	0.648	0.204	1.56
	210	0.185	0.131	0.693	0.603	0.178	1.511
	225	0.157	0.113	0.713	0.546	0.15	1.449
	240	0.134	0.097	0.739	0.513	0.127	1.381
	255	0.102	0.074	0.782	0.466	0.095	1.26
	270	0.083	0.061	0.814	0.448	0.077	1.167
	285	0.065	0.048	0.846	0.419	0.059	1.056
	300	0.05	0.037	0.873	0.391	0.044	0.945
	315	0.043	0.032	0.884	0.371	0.037	0.89
	330	0.037	0.028	0.895	0.35	0.031	0.827
	345	0.037	0.029	0.891	0.34	0.032	0.845
	360	0.036	0.028	0.885	0.312	0.031	0.858

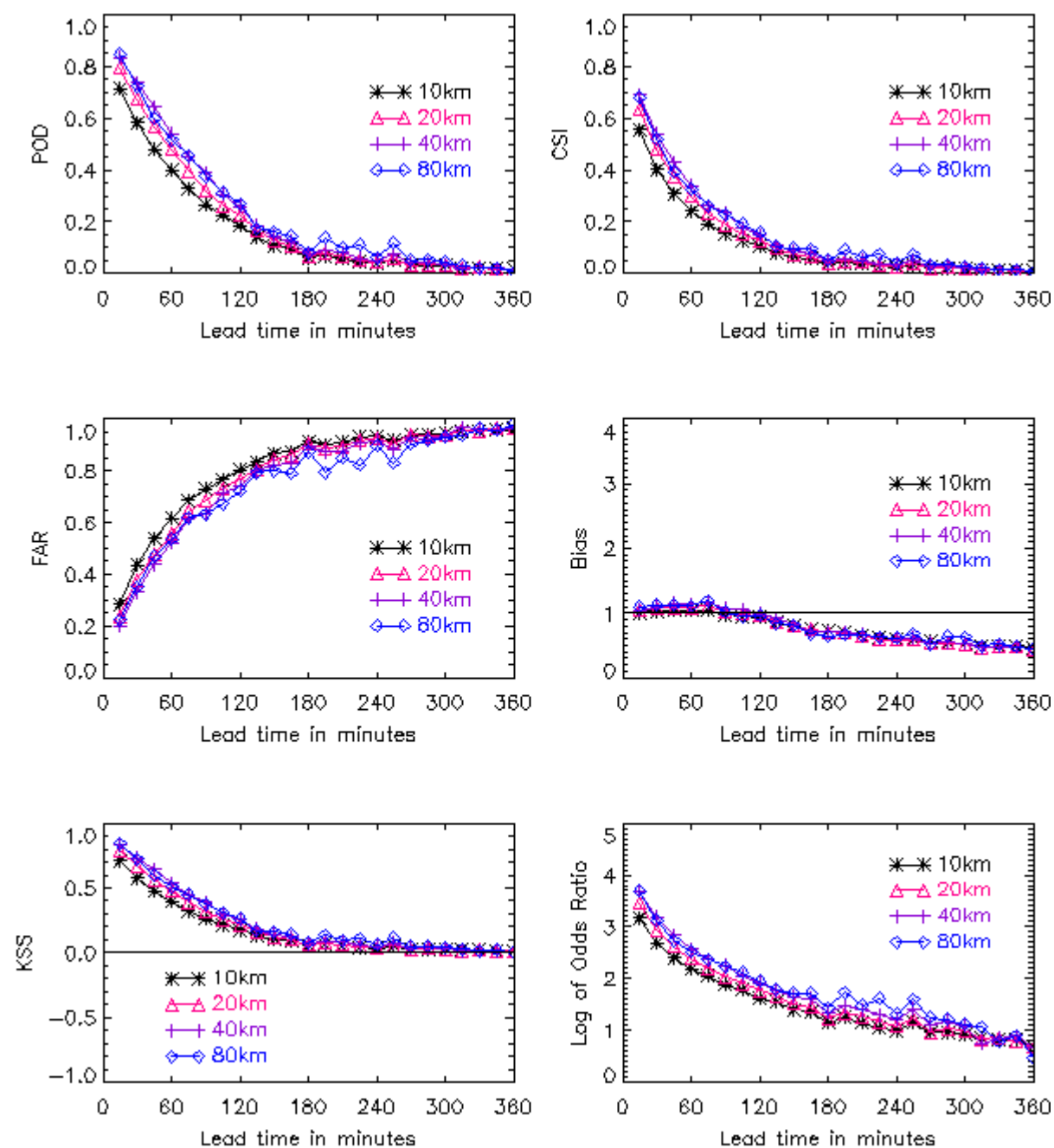


	Lead time (minutes)	POD	CSI	FAR	Bias	KSS	Odds
Spatial resolution: 20 km	15	0.812	0.69	0.179	0.989	0.809	3.132
	30	0.682	0.53	0.296	0.968	0.677	2.617
	45	0.588	0.432	0.381	0.949	0.582	2.34
	60	0.509	0.362	0.445	0.917	0.502	2.143
	75	0.441	0.304	0.505	0.892	0.433	1.987
	90	0.379	0.256	0.559	0.858	0.37	1.848
	105	0.325	0.217	0.605	0.822	0.316	1.732
	120	0.28	0.186	0.645	0.79	0.272	1.634
	135	0.253	0.167	0.669	0.765	0.244	1.571
	150	0.227	0.15	0.693	0.739	0.218	1.509
	165	0.208	0.138	0.709	0.714	0.199	1.464
	180	0.19	0.126	0.726	0.694	0.181	1.419
	195	0.178	0.12	0.733	0.667	0.169	1.395
	210	0.159	0.109	0.744	0.622	0.151	1.362
	225	0.138	0.096	0.763	0.585	0.131	1.307
	240	0.119	0.083	0.783	0.548	0.111	1.248
	255	0.103	0.074	0.796	0.506	0.096	1.205
	270	0.092	0.066	0.808	0.48	0.085	1.167
	285	0.075	0.055	0.832	0.45	0.069	1.087
	300	0.063	0.046	0.85	0.42	0.056	1.016
	315	0.055	0.041	0.862	0.4	0.049	0.962
	330	0.049	0.037	0.871	0.381	0.043	0.922
	345	0.048	0.036	0.87	0.368	0.042	0.919
	360	0.044	0.034	0.874	0.348	0.038	0.894

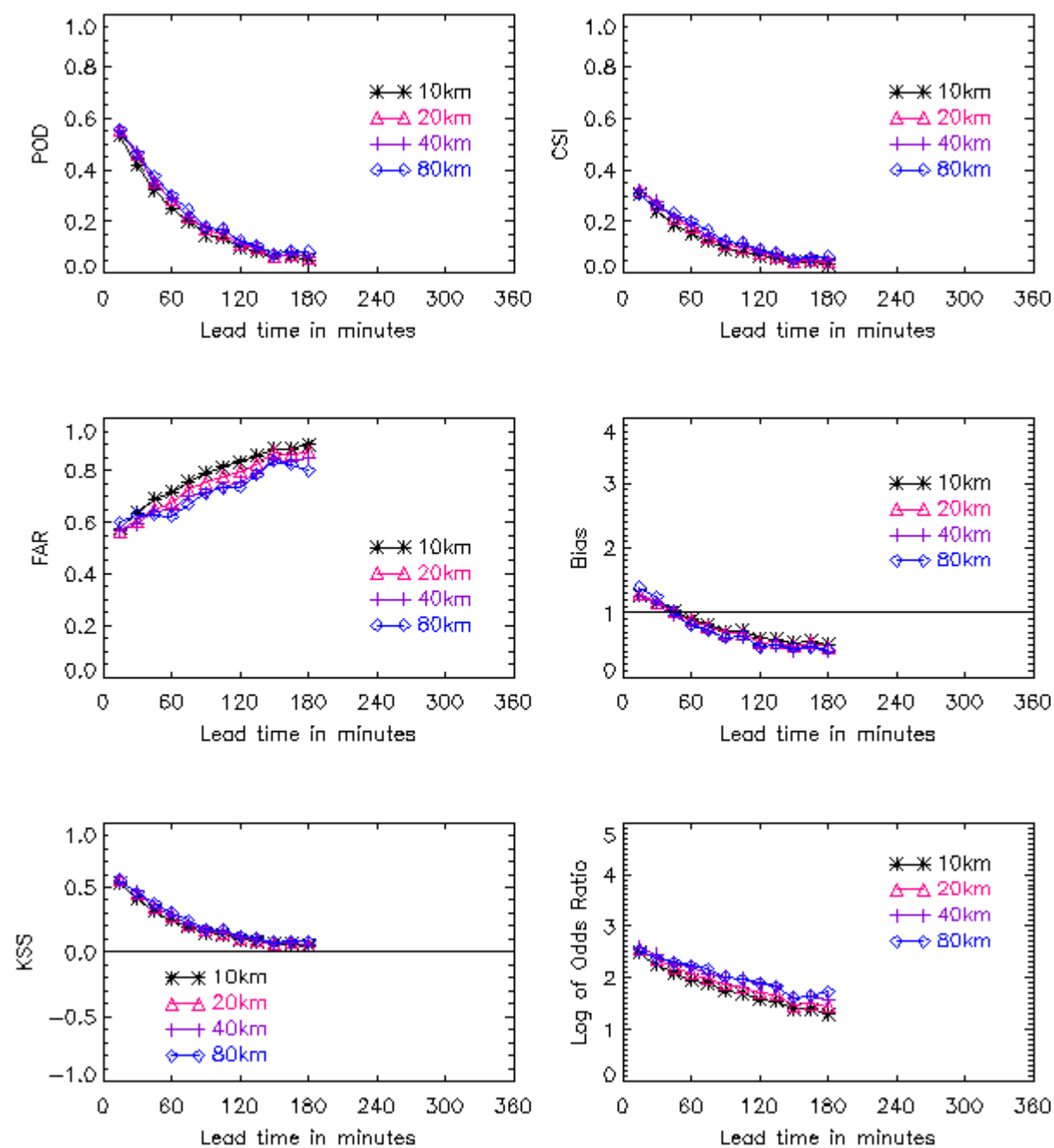
Spatial resolution: 10 km	15	0.76	0.617	0.233	0.99	0.756	2.876
	30	0.618	0.456	0.364	0.972	0.611	2.397
	45	0.528	0.37	0.447	0.955	0.52	2.155
	60	0.456	0.309	0.51	0.929	0.447	1.981
	75	0.398	0.264	0.561	0.907	0.389	1.853
	90	0.343	0.224	0.608	0.875	0.333	1.73
	105	0.296	0.192	0.648	0.841	0.286	1.629
	120	0.26	0.168	0.678	0.809	0.251	1.552
	135	0.236	0.152	0.699	0.783	0.226	1.497
	150	0.213	0.138	0.72	0.761	0.203	1.441
	165	0.195	0.127	0.734	0.733	0.185	1.397
	180	0.182	0.119	0.745	0.714	0.172	1.366
	195	0.17	0.112	0.754	0.693	0.161	1.339
	210	0.152	0.101	0.768	0.654	0.143	1.298
	225	0.136	0.092	0.779	0.616	0.127	1.262
	240	0.122	0.084	0.79	0.58	0.114	1.228
	255	0.107	0.075	0.802	0.542	0.1	1.186
	270	0.096	0.067	0.815	0.518	0.088	1.145
	285	0.081	0.058	0.832	0.486	0.074	1.084
	300	0.069	0.05	0.85	0.459	0.062	1.017
	315	0.062	0.045	0.86	0.44	0.055	0.973
	330	0.055	0.04	0.87	0.418	0.048	0.927
	345	0.051	0.037	0.874	0.403	0.044	0.899
	360	0.047	0.035	0.879	0.386	0.04	0.875

## Appendix E: Graphs of verification scores: main study

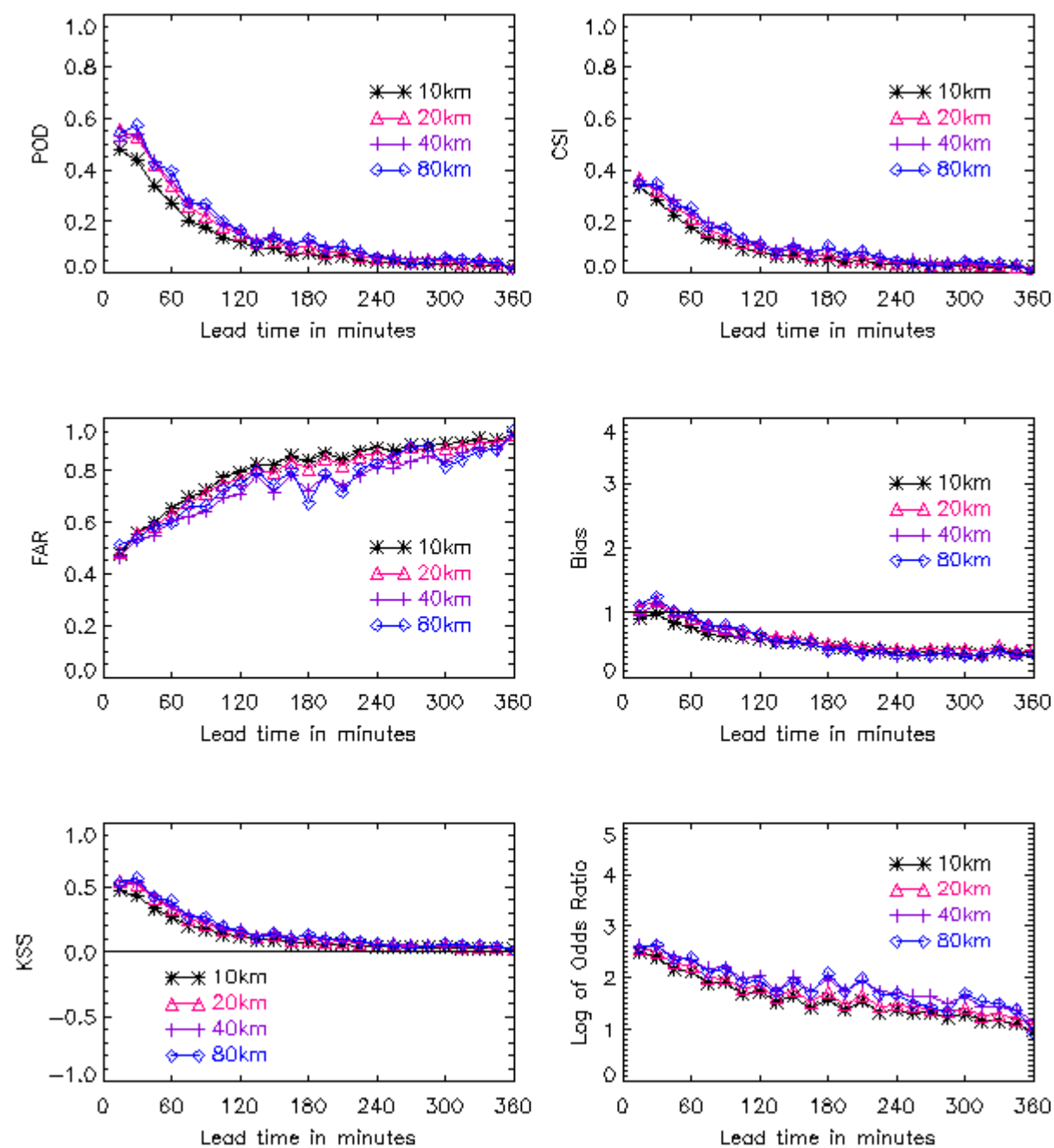
### Appendix E1: Verification statistics for the 3D\_extrap forecasts plotted for the four spatial resolutions against lead time in minutes



## Appendix E2: Verification statistics for the Gandolf forecasts plotted for the four spatial resolutions against lead time in minutes

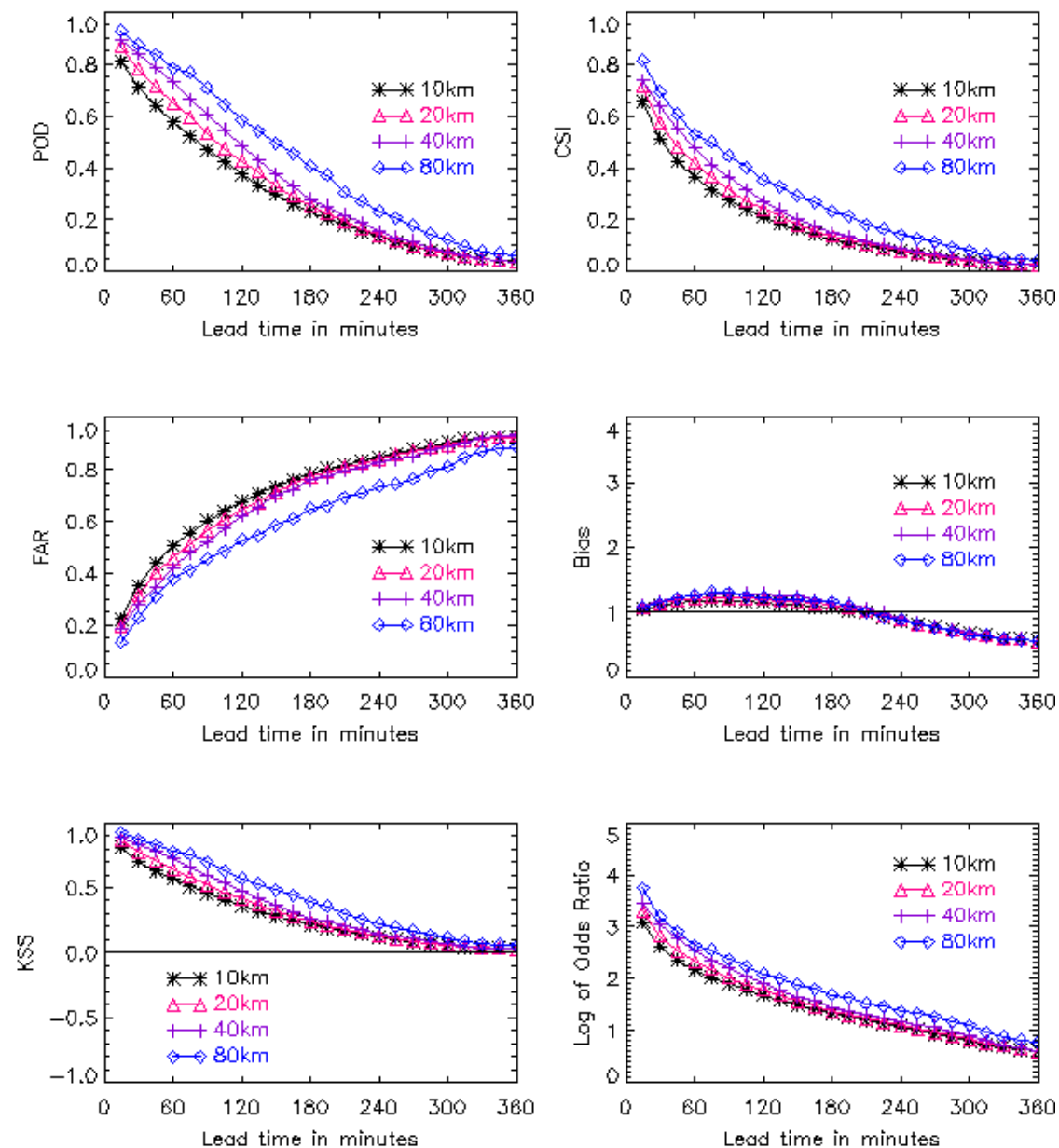


### Appendix E3: Verification statistics for the Nimrod forecasts plotted for the four spatial resolutions against lead time in minutes

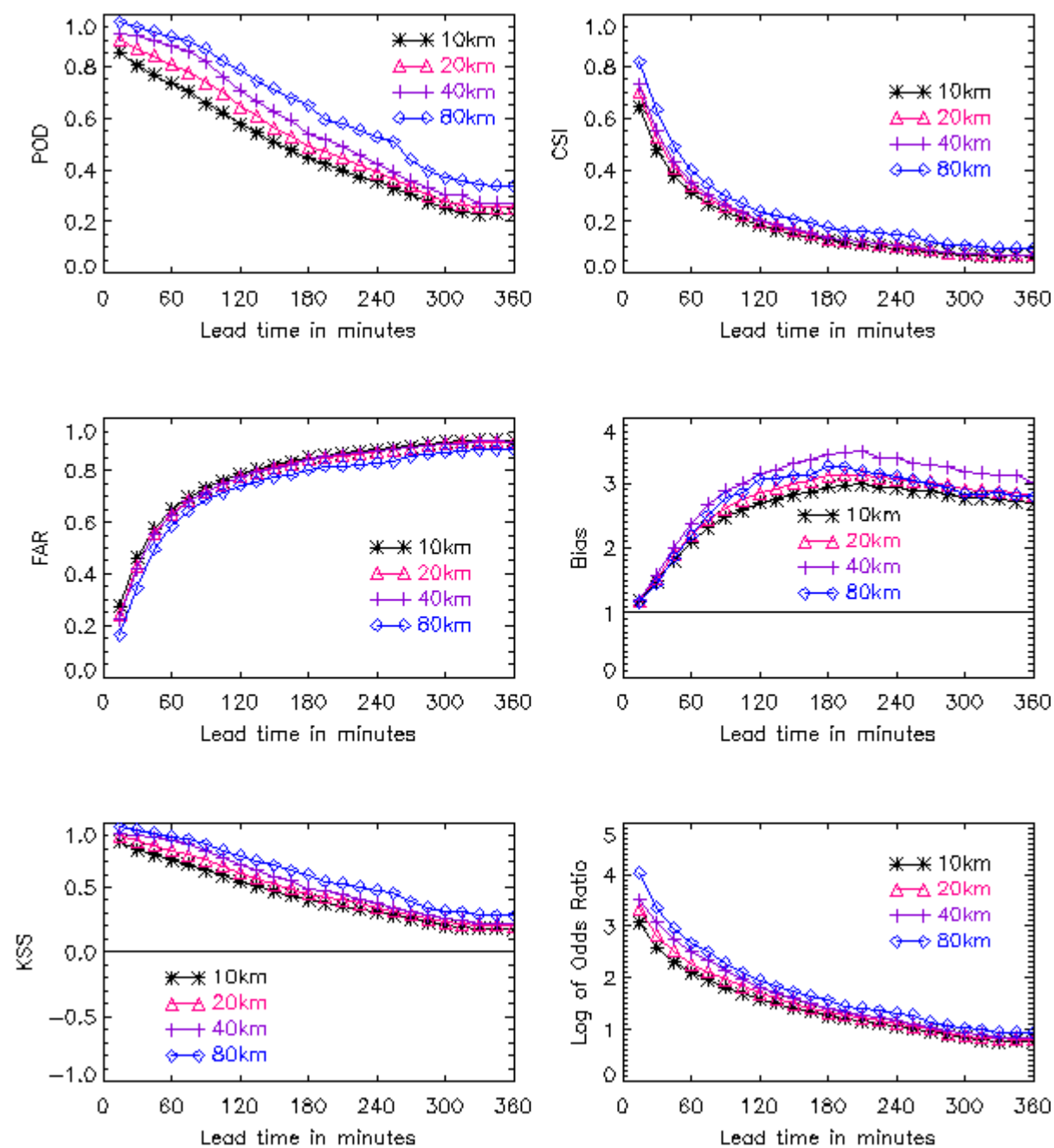


## Appendix F: Graphs of verification scores: supplementary study

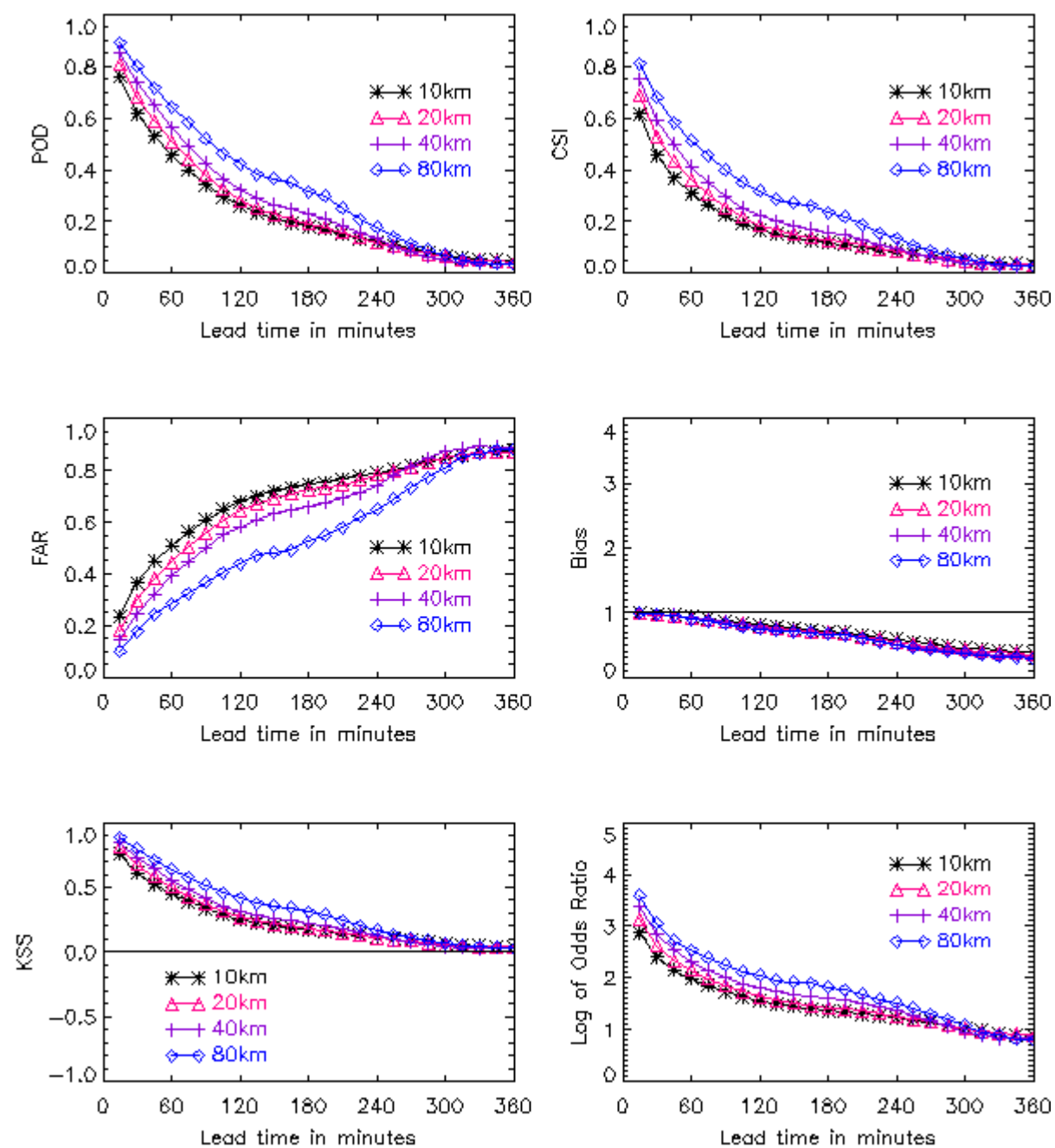
### Appendix F1: Verification statistics for the 3D\_extrap forecasts plotted for the four spatial resolutions against lead time in minutes



## Appendix F2: Verification statistics for the 3D\_trended forecasts plotted for the four spatial resolutions against lead time in minutes



### Appendix F3: Verification statistics for the Persistence forecasts plotted for the four spatial resolutions against lead time in minutes



## Appendix G: Results from the case studies comparing 3D\_extrap forecasts to Terminal Control weather disruption reports

**29/07/2002:**

A large area of convective activity appeared mid-afternoon in the northwest of the Terminal Manoeuvring Area (TMA) and moved gradually in a north-easterly direction throughout the afternoon and evening.

Time Z	Terminal Control weather disruption report	How well does the forecast pick up the disruption?
1415	Weather starting to affect small areas of the Terminal Manoeuvring Area North (TMAN) and the Bovingdon (BNN) hold, north-north-west of LL (London Heathrow)	The very small amount of convection indicated on the replicated radar display is not sufficient to be classified as causing disruption either in the analyses or the forecasts
1530	Weather starts to affect Compton (CPT) departures off LL and SS (London Stanstead)	Only small areas of green pixels are shown on the replicated radar displays and hence no disruption is forecast
1610	Weather is starting to cause problems throughout including the intermediate approach into LL	Disruption is forecast in the area up to 45 minutes ahead in the 10km resolution
1620	A Minimum Departure Interval (MDI) of 1 per 6 minutes (1/6) is placed on the Buzad (BUZ) route from LL	Although disruption is only forecast in the BUZ area up to 20 minutes ahead, lots of disruption is forecast just to the north, and may be the cause of the problem
1640	A stop is required on traffic into the Midlands sectors	This is forecast up to 75 minutes ahead, in most spatial resolutions
1704	A regulation of 35 is placed on both COWLY and WELIN sectors	These are forecast at the 10, 20 and 40km resolutions up to 45 minutes ahead and up to 90 minutes ahead for a 10km resolution
1730	The BUZ route from the London Terminal Manoeuvring Area (LTMA) is stopped	Disruption in this area is forecast up to a 40km resolution up to 75 minutes ahead
1924	Welin (WEL) sector is still affected by the weather	This is picked up in all resolutions up to 195 minutes ahead
2000	Weather is moving east and some aircraft are able to fly in the WEL airspace	This is clear in the analysis and the 15 minute forecast, but at greater lead times disruption is still forecast in the area

**30/07/2002**

This was one of the largest thunderstorm events of the summer. From late morning convection started building up in a NW-SE line over England. The strongest activity was in the North, the East Midlands and East Anglia, and persisted well into the evening.

Time Z	Terminal Control weather disruption report	How well does the forecast pick up the disruption?
0830	Swanwick requests a 1/6 MDI on the CPT route from LL due to weather on the Bristol sector	This is not picked up in the analysis or in any of the forecasts. There are no amber or red pixels shown in the area on the replicated radar display. It may be that it was a different type of weather and not thunderstorms that caused this particular disruption.
1030	Weather is starting to build on the ESSEX and DAGGA sectors	Disruption due to thunderstorm activity is forecast up to an hour ahead for the 10km resolution forecast. At this time there is not enough coverage to warrant a forecast of disruption at the lower resolutions.
1135	Stops put on all DVR (Dover) departures due to massive weather problems in DVR sector	This is forecast at 10, 20 and 40km resolutions up to 30 minutes ahead and at 10 and 20km resolutions up to a 90 minutes lead time.
1425	Flow regulation of 35 is placed on the WELIN sector	Disruption in this area is forecast at all resolutions up to an hour ahead.
1501	Weather in Daventry N is causing problems	This is forecast at all resolutions up to 120 minutes ahead.
1730	Regulations are placed on both the COWLY and WELIN sectors and weather on the	The location and severity of the disruption due to the thunderstorms is well forecast. All resolutions show



	midlands sectors is severely disrupting operations	disruption in COWLY up to a 105-minute lead time, and in WELIN up to 150 minutes ahead.
2100	Regulations taken off the COWLY and WELIN sectors	Thus no disruption should be forecast there at this time. This is true up to an hour ahead, but at longer lead times than this disruption is still forecast in the WELIN sector and in north Daventry.

### **04/08/2002**

From midday onwards there is a lot of convective activity in the LTMA, in the South East at first, then spreading, before affecting the Midlands sectors later on in the afternoon.

<b>Time Z</b>	<b>Terminal Control weather disruption report</b>	<b>How well does the forecast pick up the disruption?</b>
1245	Weather starting to make itself known but the numbers (deviating) are manageable until later. (It doesn't say if this is because the traffic flow is low or because the weather isn't severe enough yet)	We forecast some disruption in the W and SW of the TMA at the 10km and 20km resolutions up to a lead time of 180 minutes.
1400	Large amount of weather avoidance at LL and around the TMA	We forecast disruption, mostly to the W of LL, up to 180 minutes lead time, but with increasing lead time the amount that is forecast is reduced.
1410	Between 1410 and 1425 departures are stopped from LL, SS (Stanstead) and GW (Luton) to the NE, SS and LL to the NW and KK (Gatwick) on the SAM (Southampton) route	We forecast the disruption on the NE departures very well, at the 10, 20 and 40km resolutions up to 150 minutes ahead. However, we only forecast the problems on the SAM route up to 75 minutes ahead, and only in the 10km resolution. The problem on the NW departures is forecast up to an hour ahead in the 10 and 20km resolutions.
1425	Impose an MDI of 1/4 on the BUZ route from SS and the BPK (NE route to Clacton sector) route from LL	The disruption on these two routes is forecast at most lead times up to 180 minutes.
1530	Thunderstorms have spread to cover a wider area and weather in DTY (Daventry sector) is causing a problem for LL WOB (Northerly) departures	As the cells within the DTY sector are still quite small, disruption in this area is only forecast in the 10km resolution, and only up to 60 minutes ahead. It should be noted that there is a large area of disruption being forecast at all resolutions just to the north of the Daventry section. It is not possible to tell whether the disruption recorded in the log in the DTY sector is due to the small cells within it or the large area of convection to the north.
1650	Still problems in the DTY sector	Forecast in the 10km and 20km resolutions up to 120 minutes ahead.
1700	Less weather around SS so the restriction here is lifted	No disruption is forecast at near SS up to 60 minutes lead time but at longer lead times than this disruption is forecast around SS.

### **07/08/2002**

There is localised convective activity in the London TC area throughout the afternoon, spreading to the Midlands sectors as the afternoon wears on.

<b>Time Z</b>	<b>Terminal Control weather disruption report</b>	<b>How well does the forecast pick up the disruption?</b>
1327	Weather avoidance at LAM (Lambourne hold) and BNN (Bovingdon hold)	This is not well forecast. Some cells are visible on the replicated radar display but these are not of sufficient size to cause a forecast of disruption. For lead times of 30-60 mins disruption is forecast at the LAM area in the 10km resolution.
1520	Big problems on the DTY sector with weather avoidance causing Swanwick to impose MDIs.	Disruption is forecast for this area at the 10km resolution out to a lead time of 165 minutes. However, even in the analysis no disruption is forecast at the wider spatial resolutions of 20, 40 and 80 km. Thus the small cells that are present seem to be causing more disruption than would usually be

		expected.
1520	From 1520 to 1605 MDIs are placed on the LAM route from KK, WOB route from LL, BUZ route from SS, and OLNEY route from GW.	This disruption is forecast in the 10 km resolution at most lead times up to 90 minutes but not in the other resolutions because there is not sufficient coverage of the cells in the wider resolutions to indicate disruption. This is an example of how even very localised convective activity can still cause disruption if it occurs in a critical area within the LTMA.
1620	Severe disruption to the LL approach sequencing	This is forecast in the 10km and 20km resolutions, but only up to 45 minutes ahead.
1732	LAM/CLN route from KK has an MDI placed upon it	This disruption is forecast in the 10km and 20km resolutions up to 120 minutes ahead
1829	SAM and CPT routes from LL are restricted to an MDI of 1/5	This is picked up in the disruption forecasts at the 20km resolution up to 45 minutes lead time and in the 10km resolution up to 135 minutes lead time.
1835	Problem with LL inbounds from the BIG stack	Disruption is forecast for these areas up to about a 150-minute lead time in the 20km and 10km resolutions.
1900	Problem on the LL SAM departures	Disruption is forecast on this route at the 20km resolution up to 60 minutes lead time and in the 10km resolution up to 135 minutes lead time. However, it should be noted that at these lead times a lot of disruption is forecast at the 10km resolution in areas that did not experience disruption (false alarms).

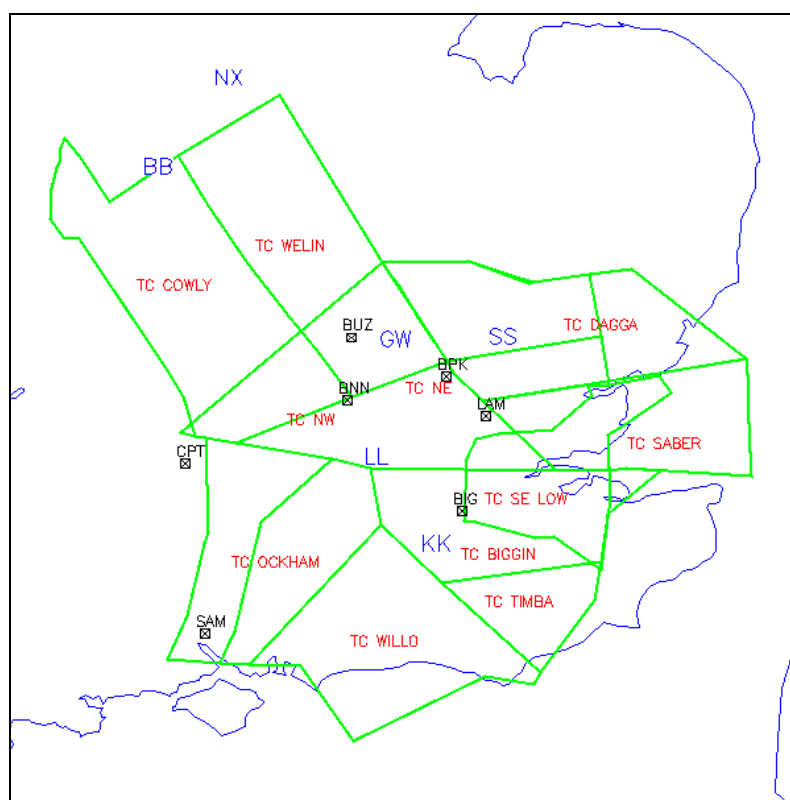
### 15/08/2002

This was a relatively small thunderstorm event affecting East Anglia only. The thunderstorms are localised and thus their coverage is not enough to cause any disruption to be forecast in the 40km and 80km resolutions.

<b>Time Z</b>	<b>Terminal Control weather disruption report</b>	<b>How well does the forecast pick up the disruption?</b>
1700	Weather avoidance in the NE quadrant. An MDI is applied off Stanstead from 1719-30.	The thunderstorms are localised and thus their coverage is not enough to cause any disruption to be forecast in the 40km and 80km resolutions. Disruption is forecast around Stanstead and in the Beeno Sector up to 90 minutes ahead in the 20km resolution forecast and up to 120 minutes ahead in the 10km forecast.

## Appendix H: Maps showing the places mentioned in the Terminal Control weather disruption reports

### London Terminal Control area



### Medium level sectors

