Numerical Weather Prediction

The impact of observations on Mesoscale Model forecasts of 3-hourly rainfall accumulations

Forecasting Research Technical Report No. 250

Simon Anderson, Richard Graham and Michael Bader

email: nwp_publications@metoffice.com

The Impact of Observations on Mesoscale Model Forecasts of 3-hourly Rainfall Accumulations

S.R. Anderson, R.J. Graham and M.J. Bader

Contents

The Impact of Observations on Mesoscale Model Forecasts of 3-hourly Rainfall Accumulations

S.R. Anderson, R.J. Graham and M.J. Bader

1. Introduction

1.1 Background and aims of study

The aim of this work is to investigate the impact of observations on the accuracy of forecasts of precipitation accumulation from the UKMO Mesoscale NWP model. Results from such experiments provide useful evidence on which to base the design of observational networks over and around the UK, and should therefore help to continue the trend of improving our national precipitation forecasts. All observations made within the Mesoscale model domain have been considered. A full description of the 17km resolution, 31-level model (operational until June 1998, before changes in resolution and domain took place) can be found in Cullen (1993).

1.2 Overview of this report

The paper is divided into the following sections. Section 2 contains an analysis of the procedure used to select the cases, and gives useful guidance on the frequency with which we might expect observations to deliver marked forecast benefits. The case study method used to obtain the results is described in Section 3. For a greater understanding, however, the reader is strongly recommended to follow the steps in Section 4, which contains a worked example of a typical case.

A summary of all the results is given in Section 5. A discussion of the results is offered in Section 6, and conclusions and recommendations in Section 7.

2. Case selection

2.1 Technique

Cases in which initial data had a marked benefit on forecasts of widespread rainfall were identified by looking for notable improvements in forecast skill between consecutive operational forecasts valid at the same time (following Graham and Anderson, 1995). Improved skill may be attributed to the additional observations available for assimilation in the later run (provided impact from the more recent boundary conditions used in the later run can be discounted). Cases in which objective measures of forecast skill indicated a significant improvement in the T+12 forecast over the T+18 forecast, or the T+6 forecast over the T+12 were short-listed for study.

More details of the criteria used for the case selection procedure are given in Appendix 1, together with the definitions used for 'marked benefit' and 'widespread'. In all, thirteen cases were identified for detailed study, and it is the results from these which are presented in this paper.

2.2 Frequency of significant data impact on rainfall forecasts

In this sub-section, we provide an estimate of the frequency with which we can expect significant data impacts on forecasts of rainfall accumulation. This calculation (originally carried out by Graham et al., 1997, and reproduced here for completeness) is based on results of the case selection procedure used to select our first nine cases, between the dates of 8th March 1996 and 12th November 1996. (Our remaining four cases were chosen using a similar procedure, but during the autumn of 1997).

Between March and November 1996, a total of 936 operational Mesoscale model forecasts were checked for sensitivity to initial data. Of these, 512 forecasts (~55%) were for cases in which widespread rain was observed over the UK area at the verifying time. Of these 512 cases, 137 (~27%) showed positive impacts either from recent observations (i.e. observations from within the Mesoscale domain assimilated in the 6-hr period preceding the forecast) or from updated model boundary data.

Subjective criteria, described in Appendix 1, were used to select cases from the 'shortlist' of 137 as a representative sample. Nine cases were selected. It transpired that in all these cases, beneficial impact from observations was found to dominate over benefit from updated model boundary conditions. Assuming the cases are representative, we may tentatively conclude that when marked improvements occur between consecutive model forecasts, it is observational input, rather than updated model boundary information, which is usually responsible. Moreover we can estimate that such marked data impacts occur in about 25% of all forecasts of widespread rain (corresponding to about 15% of the time).

3. Method for assessing impact of observations

3.1 Experiment format

To assess the impact of the observations, a series of model re-runs was performed for each case. For this study, all the data impact experiments comprised a 6-hr period of assimilation (using the UKMO analysis correction scheme - see Lorenc et al., 1991) followed by either a 6-hr or 12-hr forecast.

For each case, two control runs were performed. The ALL OBS run, which used all the available data (i.e. as used in the operational Mesoscale runs), and the NO OBS run which entailed a dummy assimilation using no data. The NO OBS run represents the earlier, poor forecast, whilst the ALL OBS run represents the next consecutive operational model forecast, with all the benefit of the 6 hours of observations made between the two runs.

All our re-run experiments were designed to mimic as closely as possible the operational model runs. However, there was a difference in how we utilised the model boundary data. Operationally, the boundary data were updated every 3 hours, the fields being created from the UKMO Limited Area Model (operational until April 1998). In our re-runs, we used identical boundary conditions for both the NO OBS and ALL OBS runs. This meant that, provided the ALL OBS forecast still showed a marked improvement over the NO OBS forecast, we could conclude that the improvement in the forecast had come from the observations, rather than through the model boundary.

For the remaining experiments in each case study, each observation type was assimilated individually, and its impact recorded (as described in Sub-section 3.3). This method of assessing the impact of observation types separately provides useful insight into which data types play key rôles in an observation network - a requirement for network design. (The alternative method of omitting each observation type in turn from the assimilation has the disadvantage in that some degree of redundancy would normally be expected in the benefit from many data types. If experiments are run omitting one data type, then it is likely that its absence will be compensated by other remaining types).

Further runs were usually carried out to gain more insight into which combinations of observation types provided the most benefit in individual cases. More details of the method can be found in Section 4, where we describe the procedure and results from an example case study.

3.2 Observation types used

The observation types used in the UKMO Mesoscale model include radiosonde winds, temperatures and humidities, aircraft winds and temperatures, surface observations (of PMSL, winds from SYNOPS, moored buoys and ships, and temperatures and humidities from SYNOPS), satellite temperatures and cloud-track winds. All these data types were assessed separately. A map showing the locations of the SYNOP stations, moored buoys, rigs and radiosonde stations whose data were assimilated in the model is given in Appendix 2. (Note that profile data from aircraft within the domain were only available from some planes serving the major London airports).

In addition to the above, a 3-D moisture field, generated by the Moisture Observation Processing System (MOPS), is used. For a detailed description of this system, see Macpherson et al. (1996). Briefly, the purpose of MOPS is to generate a 3-D analysis of cloud fraction from satellite and radar rainfall imagery, selected SYNOP observations and a short-period forecast. The elements of the SYNOP used are cloud cover, base and type and present weather. The analysis of cloud fraction is then used to derive relative humidity profiles which are assimilated as "pseudo radiosondes", with one profile per grid square.

3.3 Forecast verification

Forecast skill was assessed by comparing the accumulation of precipitation over the final 3 hours of the forecast with an analysis of 3-hour accumulation derived from the UK weather radar network by the UKMO's very short range forecasting system, Nimrod. The Nimrod analyses include corrections for range and bright band effects, and for orographic enhancement of precipitation (Kitchen et al., 1994). Before use for forecast verification, the Nimrod analyses, which have a grid resolution of 5km, were smoothed by assigning the average value over a square array of 9x9 grid points to the central point in the array. The forecast accumulations were smoothed in a similar way using a 3x3 grid-point array. The smoothing yielded observed and forecast precipitation patterns that had similar scale representation (~50km). Verification of the model forecasts was restricted to the area covered by the UK weather radars (Fig.1).

Equitable Threat Scores (ETS) were calculated at two thresholds of rainfall accumulation - 0.5mm/3hrs (for precipitation of light intensity and greater) and 2.0mm/3hrs (for rain of moderate intensity and greater). These rates were based on the thresholds for verification applied operationally at the UK Met. Office (namely 1.0mm/6hrs and 4.0mm/6hrs). The ETS (see Schaefer, 1990, for further details) is widely respected, and is used as the main verification measure for precipitation in the UKMO's UK Index. It penalises both missed forecasts of observed rain and the forecasting of rain in the wrong place, and includes a correction for the rainfall events which may have been expected to be correctly forecast by chance. It has a theoretical range from 1.0 (highest skill) to -0.33.

The skill of each forecast is expressed as a percentage of the difference in skill between the ALL OBS and NO OBS controls. For example, for a forecast using data from MOPS,

$$\text{BENEFIT OF MOPS} = \frac{[\text{ ETS(MOPS) - ETS(NO OBS) }]}{[\text{ ETS(ALL OBS) - ETS(NO OBS) }]} \times 100\%. \qquad (1)$$

Hit rates, False Alarm rates and Root Mean Square Factor (rmsf) errors are other measures commonly applied to precipitation verification. They were also used, therefore, to assess the skill of the rainfall forecasts at the lower threshold of 0.5mm/3hrs.

The hit rate represents the percentage of points where rainfall accumulations above the threshold were forecast correctly. Hence, high scores are preferred. The false alarm rate represents the percentage of points where rainfall accumulations were forecast but did not occur. Low scores are desirable.

The root mean square factor error can be thought of as an average ratio, in a root-mean-square sense, of the forecast to observed precipitation. Low scores are once again preferred. It has the advantage of concentrating on the rain areas rather than being dominated by dry events, as, for every grid-point where the score is calculated, at least one of the forecast value or observation has to exceed the threshold value. (Note that the threshold value for the rmsf error is used as a lower limit in order to avoid division by zero). The main disadvantage, however, is that it condenses all the information about the forecast behaviour at different rainfall intensities into one score. Hence it is considered inferior to the ETS, which may be calculated at any number of rainfall thresholds.

4. Worked example from one case study

In this section we work through a typical case (Case 13: verifying at 18UTC, 6th October 1997) to demonstrate the procedure for calculating the impact of the observation types and to give the reader a better understanding of the magnitudes of the skill scores that contribute towards the results from all thirteen cases. The method of objective verification used to obtain the results is presented, and we also give a subjective evaluation of the observational impact on forecast skill for both low and high levels of rainfall accumulation (i.e. for thresholds of 0.5mm/3hrs and 2.0mm/3hrs).

Synoptic situation and precipitation distribution

The surface analysis for 18UTC on 6th October 1997 is shown in Fig. 2a. A low pressure system was located to the north west of the British Isles, with a cold front extending from Eastern Scotland to S.W. England. There was also a thundery trough ahead of the front. The significant weather chart for 18UTC (Fig. 2b) shows widespread precipitation associated with the front, in a band from N.E. England, through North Wales towards Cornwall. Precipitation of moderate intensity and thunderstorms associated with the trough were reported over East Anglia.

Fig. 3a shows the contoured precipitation accumulation field for the 3 hours 15-18UTC, and Fig. 3b shows the grid-point location of accumulations greater than 2.0mm, as derived from the Nimrod accumulation analysis. The main areas of accumulations in excess of this greater threshold are over Cumbria, Wales and Cornwall, and also over Eastern England in a region around The Wash. The Nimrod analysis is clearly consistent with the significant weather reports (Fig. 2b).

Control runs

The corresponding 6-hr forecasts from the NO OBS and ALL OBS control runs are shown in Figs. 4a&b and 5a&b respectively. In Figs. 5a&b, black shading shows the locations for which an accumulation exceeding 2.0mm/3hrs was correctly forecast (a Hit), whilst grey shading shows the locations for which such an accumulation was forecast, but did not occur (a False Alarm).

Both NO OBS and ALL OBS forecasts produce a spurious rainfall area over Southern England and across the English Channel. In fact, the NO OBS forecast in Fig. 4a shows heavy rainfall in this area, with accumulations peaking at 26mm. This run has also produced spurious rain over Southern Scotland. The ALL OBS run (Fig. 5b), however, has a better forecast of the rain along the front and over Eastern England, and has correctly eliminated the rainfall over Southern Scotland. Comparison with the analysis (Fig. 3b) shows clearly that the ALL OBS run gives useful guidance for the heavier rainfall whilst the NO OBS run is very misleading.

The statistics for the two control runs for this case are shown in Table 1. The superiority of the ALL OBS run is reflected in a higher hit rate at the 0.5mm/3hrs threshold (59.9% compared to 53.6% for the NO OBS run), a lower false alarm rate and rmsf error, and higher ETS scores at both thresholds (0.184 compared to only 0.054, for example, at the 2.0mm/3hrs threshold). The greater skill of the ALL OBS run shows that observational data introduced in the 6-hr assimilation cycle prior to the ALL OBS forecast delivered a marked benefit.

	Threshold	**E.T.S.**	**Hit Rate**	**False Alarm rate**	**RMS Factor error**
NO OBS	0.5:	0.096	53.6	43.4	5.02
	2.0:	0.054			
ALL OBS	0.5:	0.235	59.9	29.6	3.89
	2.0:	0.184			

Table 1 - Scores from different verification measures for precipitation forecasts at T+12 for Case 13 for the control forecasts.

This table shows the results of the two control runs from our example case.
Column 2 indicates which rainfall threshold the scores refer to - either the lower threshold of 0.5mm/3hrs, or the higher one of 2.0mm/3hrs.
For an explanation of the verification measures in columns 3-6, see Sub-section 3.3.

Impact of individual observation types

We now investigate the rôle that the individual observation types played in the impact. The forecasts were re-run, and each time a different observation type was assimilated, and its effect on the calculated forecast skill scores noted. Table 2 shows the results of these runs for Case 13.

The ETS results show that the most beneficial individual observation type for this case was the moisture field created by MOPS. The score of 0.241 at the lower threshold represents a 104% benefit compared to the ALL OBS benefit, when calculated using Equation (1). The other significant observation types at the lower threshold were Satellite temperatures (SATEMS, 33%) and Surface observations (25%). At the higher threshold, the four most beneficial observation

	Threshold	E.T.S.	Hit Rate	False Alarm rate	RMS Factor error
MOPS	0.5:	0.241	69.7	33.3	3.57
	2.0:	0.144			
RsH	0.5:	0.076	37.1	41.6	5.92
	2.0:	-0.020			
RsW	0.5:	0.130	54.0	39.5	4.72
	2.0:	0.096			
RsT	0.5:	0.107	53.4	42.0	4.75
	2.0:	0.123			
SURFACE	0.5:	0.131	54.0	39.4	4.79
	2.0:	0.099			
AcW	0.5:	0.110	56.2	42.4	4.91
	2.0:	0.076			
AcT	0.5:	0.082	52.7	44.8	5.17
	2.0:	0.040			
SATEMS	0.5:	0.142	55.6	38.6	5.35
	2.0:	0.064			
SATOBS	0.5:	0.106	55.2	42.6	5.00
	2.0:	0.057			

Table 2 - Scores from different verification measures for precipitation forecasts at T+12 for Case 13 for the individual observation type forecasts.

This table shows the results from the individual observation type runs from our example case. The thresholds and verification measures are the same as in Table 1.
The observation types tested were MOPS (see Sub-section 3.2 for details), Radiosonde humidities (RsH), Radiosonde winds (RsW), Radiosonde temperatures (RsT), Surface observations, Aircraft winds (AcW), Aircraft temperatures (AcT), Satellite temperatures (SATEMS) and Cloud-track winds (SATOBS).
All scores should be compared to the control run scores shown in Table 1.

types were MOPS (69%), Radiosonde temperatures (53%), Surface observations (35%) and Radiosonde winds (32%).

MOPS was the most influential observation type for all verification measures in this case.

Summarizing, the results for the low rainfall threshold show that :-
For the ***ETS***, the three most beneficial observation types were MOPS, SATEMS and surface observations.
For the ***hit rate***, MOPS, aircraft winds and SATEMS gave the highest individual skills.
For the ***false alarm rate***, MOPS, SATEMS, surface observations and radiosonde winds had the most impact.
For the ***root mean square factor error***, the best individual scores were achieved with MOPS, radiosonde winds, radiosonde temperatures and surface observations.

The results for the higher rainfall threshold show that :-
For the ***ETS***, MOPS, radiosonde temperatures and surface observations and radiosonde winds gave the four highest individual skills.

For Case 13, three or more of the verification measures show the importance of MOPS, surface observations and/or SATEMS.

Fig. 5c shows the forecast using the data from MOPS only. A fair attempt has been made at forecasting the heavy frontal rain, and much of the spurious rainfall over Southern England seen in the NO OBS run is missing. However, more false alarms have been forecast over Scotland. The forecast using only Surface observations is shown in Fig. 5d. The frontal rainband positioning is very similar to the ALL OBS run, and ahead of the front, less rain has been forecast, leading to fewer false alarms but also fewer hits. Looking at Figs. 4c&d, we see that although the Surface observations help to produce a better forecast of the frontal rain, it is only the MOPS information which eliminates the spurious totals over the English Channel.

Impact of combined observation types

Further runs were carried out combining observation types in order to find a small subset of types that would produce a forecast closely resembling the ALL OBS control run. For Case 13, two further experiments resulted in a run being performed using MOPS, radiosonde temperatures and surface observations. The results from this run are given in Table 3, and show how only three observation types can almost match (or even improve on) the effect of all the observations used in the ALL OBS forecast.

	Threshold	**E.T.S.**	**Hit Rate**	**False Alarm rate**	**RMS Factor error**
MOPS + RsT	0.5:	0.183	61.8	36.2	3.46
	2.0:	0.152			
MOPS + RsT	0.5:	0.211	61.6	33.2	3.53
+ SURFACE	2.0:	0.188			

Table 3 - Scores from different verification measures for precipitation forecasts at T+12 for Case 13 for the combined observation type forecasts.

This table shows the results from the combined observation type runs from our example case. The thresholds and verification measures are the same as in Table 1.
Note how all the skill scores show good results for the combination run `MOPS+ RsT+SURFACE', when compared to the ALL OBS scores in Table 1.

Figs. 5e&f show the forecasts from the runs using MOPS and radiosonde temperatures, and then using MOPS, radiosonde temperatures and surface observations respectively. With only the two observation types (Fig. 5e), there is still a spurious area of rainfall forecast over the Scottish borders, but when surface observations are included (Fig. 5f), these false alarms decrease and the pre-frontal rain is also better represented. The ETS has risen to 0.188, slightly higher than the ALL OBS value. The contoured accumulation chart for this run (Fig. 4f) compares very well with the ALL OBS forecast (Fig. 4b) over virtually the whole of the Nimrod domain.

5. Results from all the cases

Results like those obtained in Case 13 were computed for each case, and combined to produce the information in Figure 6. The histogram in Fig. 6a displays the results for ETS at the lower threshold, whilst Fig. 6b represents the higher threshold rainfall results. The vertical scale represents the number of times that each individual observation type has had a noticeable benefit on the model forecast. In more detail, each bar shows the number of times that the benefit exceeded 25% of the ALL OBS benefit, and also when the benefit exceeded 50%. Results for the 6-hr and 12-hr forecasts have been combined.

The results of the thirteen case studies, using ETS as the measure of skill, are summarised as follows.

- At the lower threshold of 0.5mm/3hrs, the most beneficial data impacts on rainfall accumulation forecasts were obtained with **MOPS**, **Radiosonde humidities** and **Surface observations**. Other profile information in the form of Radiosonde winds and temperatures plays a supporting rôle.

- At the higher threshold of 2.0mm/3hrs, the most beneficial impacts were obtained with **Surface observations** and **Radiosonde temperatures and winds**. **MOPS** data also play a strong rôle.

- Few examples of marked impact of satellite data were found, either from Satellite temperatures or Cloud-track winds. Both of these observation types only played a supporting rôle at the lower rainfall threshold, and had no impact on the forecasts of heavier rain.

The equitable threat score is considered to be the best statistic for assessing forecasts of rainfall accumulations. As mentioned in Section 3, however, we did calculate other objective skill scores (all at the lower rainfall threshold of 0.5mm/3hrs). A brief summary of these results is as follows.

- The most beneficial observation types on scores of hit rate, over all cases, were **MOPS**, **Surface observations** and **Radiosonde humidities**.

- When considering the effect on scores of false alarm rate, the most beneficial observation types were **MOPS**, **Radiosonde humidities** and **Radiosonde winds**.

- For the Root Mean Square Factor error, scores were improved the most by **MOPS**, **Radiosonde winds** and **Surface observations**.

- Few examples of marked benefits on any of these three skill scores were found from either Satellite temperatures or Cloud-track winds.

These results re-enforce those found using the ETS skill score.

6. Discussion of results

a. ***Consistency of results from the cases***

It is encouraging that even with just 13 cases, a consistent pattern emerges in that certain observation types provide benefit to the forecasts more frequently than others. Another encouraging feature is the consistency of the results using the different verification measures.

b. ***Dependence on precipitation 'type'***

A qualitative analysis of the cases shows that there is no apparent preference for certain observation types over the Mesoscale model domain to influence forecasts of a particular 'type' of rainfall over a particular area. The restriction to cases with widespread precipitation may filter out many purely convective cases. However, Case 13 shows the benefit of such observations in forecasting an area of convective rainfall ahead of a cold front, and also of precipitation over hilly regions.

c. ***Dependence on synoptic situation***

The impact of observations over the Mesoscale model domain on forecasts over the UK should depend on the forecast range and the mobility of the synoptic situation. The shorter the forecast range and the less mobile the situation, the more influence these observations should have. Conversely, the longer the forecast range and the more mobile the situation, the greater chance that observations from outside the domain will influence the forecast, through the boundary conditions. In this study, cases were chosen between the months of March and November, so the combined results from the cases reflect a slight bias towards less mobile synoptic patterns.

d. ***Importance of profile data and consequences for network design***

A feature of the results with both rainfall thresholds is the importance of humidity profile data from sondes and MOPS. This finding is perhaps not surprising with rainfall as the impact variable. Although the radiosonde data give direct measurements of humidity, the upper-air network (with a spacing of typically 300km - see map in Appendix 2) is often not dense enough for the mesoscale variability to be adequately resolved. An advantage of MOPS is that it receives more varied observational input, leading to a better humidity analysis. The results indicate that a network of humidity profile measurements, denser than that of the current radiosondes, would provide significant benefit to short-period forecasts of precipitation.

The experiments described here did not distinguish between the benefit from observations at different heights - a task worth doing in any future studies. For example, Nielsen et al. (1995) point out that a detailed representation of the forcing from the lower boundary is required to give a high quality forecast of the finer scale structure in the precipitation pattern.

e. *Importance of surface data*

The results of these experiments show that there is a key rôle for surface data, including cloud information (through MOPS). This evidence, together with the findings of Graham and Anderson (1997), shows that the impact on NWP forecasts must be considered in any plans to modify the UK surface network. (The map in Appendix 2 shows the locations of the surface stations).

f. *Comparison of results between the two rainfall thresholds*

Results with both thresholds of rainfall accumulation show the importance of surface and profile information, but radiosonde temperatures and winds become important with the high thresholds. This result is consistent with the fact that such radiosonde measurements are important for modelling the 'dynamical' ascent of air in cases of widespread moderate or heavy rain (e.g. frontal rain with embedded convection). Furthermore, for significant enhancement of rainfall over hills (e.g. as a result of the seeder-feeder process), the wind direction and speed in the lower troposphere (plus of course the humidity) are particularly important.

g. *Comparison with previous work*

It is notable in this study that aircraft winds did not play such a noticeable rôle as in Graham and Anderson (1995), in which the impact was assessed using the Global model. One reason was that, in the previous study, aircraft data (especially at cruise level over the oceans) filled important gaps in the observational network, whereas in the current study, profile data from sondes, and to a lesser extent aircraft, were distributed within the Mesoscale model domain.

The benefit of surface observations has been demonstrated before, both in the Global impact studies by Graham and Anderson (1995) and more recently in the impact studies during the FASTEX field experiment (see Anderson, 1998). Conclusions about the utility of profile information from sondes have also been made before (see e.g. Pailleux, 1998), and the results of these experiments confirm the effectiveness of such information within the current observational network.

7. Conclusions and recommendations

7.1 Conclusions

The following conclusions may be drawn from the results of our thirteen case studies looking at the impact of observations over the UKMO Mesoscale model domain on forecasts of widespread precipitation accumulation:

- The 3-D moisture field produced by the UKMO Moisture Observation Processing System (MOPS - see Sub-section 3.2 for details) has been shown to be a highly beneficial observation type (showing significant improvements in seven out of thirteen cases) when assimilated into the Mesoscale model.

- Both surface observations and profile data supplied by radiosondes have shown marked benefits to the forecasts of rainfall accumulations (in eight and seven cases respectively), and their impact should be taken into account when redesigning observing networks.

- Neither aircraft nor satellite data showed significant benefits in many of the cases.

- It is estimated tentatively that observations made within the Mesoscale model domain deliver benefits in about 25% of Mesoscale model forecasts of widespread rainfall up to T+12hrs.

- When marked improvements occur between consecutive model forecasts, it is observational input, rather than updated model boundary information, which is usually responsible.

7.2 Recommendations for further work

Recommendations for further data impact studies are summarised below.

- It is recognised that the results from just thirteen case studies may not fully represent the true significance that each observation type makes. More data impact studies are needed before firm conclusions can be made. Such studies are planned, using the new UKMO Mesoscale model which has an extended domain and improved resolution.

- The cases reported here correspond to widespread rainfall events. The impact on forecasts of convective rainfall, in which the detail is often particularly difficult to predict, needs to be studied.

- It is hoped to study any impact from model boundary data identified in future cases, alongside the observational impacts.

Acknowledgements
The authors would like to thank Dawn Harrison and Brian Golding for their help and advice in the use of Nimrod rainfall analyses.

References

Anderson, S.R., 1998: The impact of observations on Limited Area Model forecasts during the FASTEX campaign. UK Met. Office FR Division Tech. Report No. 251.

Cullen, M.J.P., 1993: The Unified Forecast/Climate Model. Meteorol. Mag., 122, No. 1449, 81-94.

Graham, R.J. and S.R. Anderson, 1995: The relative utility of current observation systems to global-scale NWP forecasts; preliminary results. UK Met. Office FR Division Tech. Report No. 173.

Graham, R.J., S.R. Anderson and M.J. Bader, 1996: The utility of observations for Mesoscale model forecasts of precipitation accumulation. Seventh Conf. on Mesoscale Processes, Univ. of Reading, 9-13 Sept. 1996, 418-420.

Graham, R.J., S.R. Anderson and M.J. Bader, 1997: The impact of automating and thinning the 'visual' part of UK SYNOP reports on Mesoscale model forecasts. UK Met. Office FR Division Tech. Report No. 214.

Kitchen, M., R. Brown and A.G. Davies, 1994: Real-time correction of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation. Quart. J. Roy. Meteor. Soc., 120, 1231-1254.

Lorenc, A.C., R.S. Bell, B. Macpherson, 1991: The Meteorological Office analysis correction scheme. Quart. J. Roy. Meteor. Soc., 177, 59-89.

Macpherson, B., B.J. Wright, W.H. Hand and A.J. Maycock, 1996: The Impact of MOPS Moisture Data in the UK Meteorological Office Mesoscale Data Assimilation Scheme. Mon. Wea. Rev., 124, 1746-1766.

Nielsen, N.W., B.H. Saas and J. Jorgensen, 1995: Mesoscale forecasts with an atmospheric Limited Area Model. Met. Apps., 2, 351-361.

Pailleux, J., 1998: Impact of various observing systems on numerical weather prediction. WMO World Weather Watch Tech. Report No. 18, WMO/TD No. 868, 1-8.

Schaefer, J.T., 1990: The critical success index as an indicator of warning skill. Weather and Forecasting, 5, 570-575.

Appendix 1: Further details of case selection technique

Cases for detailed study were selected using the following sequence of criteria:

(i) For this study, only cases in which the observed rainfall was widespread were considered for selection. The criterion used for 'widespread rain' was that the analysed 3-hr rainfall accumulations must exceed 0.5mm at more than 200 Mesoscale model grid squares (corresponding to about 7% of the area covered by the radar network).

(ii) Secondly, objective methods were used to prepare a 'shortlist' of possible cases. We looked for a minimum acceptable improvement between consecutive forecasts, as measured by differences in the Equitable Threat Score (ETS - see Section 3) for rainfall accumulation. Experience showed that significant local improvements in the rainfall forecast were usually associated with improvements in the ETS of a factor of 1.2 or more (measured over the entire radar area for a threshold of 0.5mm/3hrs). This factor was therefore set as a criterion for defining a 'marked improvement'. Two further criteria were also defined; a) the ETS of the later forecast must be greater than 0.1 for the 0.5mm/3hrs threshold, and b) the hit rate of the later forecast must be greater than 50%.

(iii) Finally, subjective criteria based on examination of the rainfall distribution over the UK were applied. For example, only cases in which there were noticeable improvements in the positioning or accumulations associated with a well-formed rainband were considered. Many cases were discarded because the rainfall patterns, although 'widespread' by our definition in (i), were not organised into any recognisable structures.

Note that even at this stage, the case selection procedure has not sifted out cases where the forecast improvements may have been caused by updated boundary conditions. It was not possible to prove for sure that observational data, from within the Mesoscale model domain, provided the benefit in each case, until the two control runs (see Section 3) were performed. In all the cases studied, however, the observational data from within the Mesoscale model domain led to the improvements seen in the later forecasts.

Appendix 2: Map showing locations of surface and radiosonde observations