Met Office

Meteorology Research and Development

Assessing the OPERA pilot data hub European radar composite for NWP applications

Technical Report No. 521

M. Mittermaier, B. Macpherson, M. Naylor, R. Scovell, D. Harrison and P. Earnshaw

email:nwp_publications@metoffice.gov.uk

Assessing the OPERA pilot data hub European radar composite for NWP application

M. Mittermaier, B. Macpherson, M. Naylor, R. Scovell, D. Harrison and P. Earnshaw

Document Change History

First draft, M.Mittermaier	28/08/2008
2nd draft, B.Macpherson	21/11/2008
Final, M.Mittermaier	12/12/2008

Abstract

The OPERA pilot data-hub project is a EUMETNET initiative which aims to unify and standardise radar processing across Europe to enable continental-scale radar networking. The project is producing a European composite using data from the radar networks of approximately 17 countries.

This report documents the results of data quality impact analyses conducted at the Met Office. Most of the data quality issues are due to the many differing approaches of deriving precipitation estimates from radar reflectivities, and how well the raw radar data are "cleaned".

The usefulness of the product was assessed for two applications: verification and data assimilation (DA). The verification component focused on data quality on the daily time-scale, whereas the data assimilation trials used higher temporal resolution products.

The data quality studies showed that there are strong regional signals. A new method for assessing the quality of the radar composite using model fields showed considerable success in identifying problem areas. The DA trials showed that whilst there is potential benefit of assimilating radar data over areas where we currently have no coverage, the data are not of sufficient quality to produce a consistent result.

These results are influential in shaping the user requirement for the next (operational) phase of the OPERA programme.

1 Introduction

The fundamental objective of the OPERA programme is "to provide a European platform wherein expertise on operationally-oriented weather radar issues is exchanged and holistic management procedures are optimized. With the establishment of its Data Hub, OPERA is now organized to support the application of radar data from the European Weather Radar Network. Another important objective of OPERA is to act to harmonize data and product exchange at the European level."

The concept of a continental-scale European composite has been on the table for a long time and offers many exciting possibilities. For the Met Office there were two aspects to this work: investigating data quality in terms of verification, and assessing potential impacts (in terms of forecast accuracy improvements or degradation) as a result of increased data coverage (positive) and quality (negative if poor).

The aim of this work was not to verify the model but to use the inherent continuity of the model forecast field to assess the discontinuity in the radar composite, based on the diagnostic successes reported by (Mittermaier, 2008) who used 12-month running-mean time series to identify a data-processing problem with Radarnet IV which affected some (but not all) radar data streams.

The last few years have also seen an unprecedented rise in interest and activity in the field of assimilating radar data into Numerical Weather Prediction (NWP) models. Some of this work was surveyed in chapter 10 of Meischner (2004). The COST-717 Action (Rossa, 2006) gave a detailed report on European efforts in this area and suggested follow up work. This analysis therefore picks up on one of the COST-717 recommendations.

The BUFR format OPERA data are on an azimuthally equidistant grid. For our purposes the data are transformed onto a UK 5 km National Grid (like our own radar data). The aims of the study can therefore be summarised as follows:

- For verification compare the NAE forecast accumulations with the daily OPERA accumulations to establish:
  - how significant the data frequency may be on quality;
  - how "clean" the data are, e.g. clutter, anaprop, bright band;
  - how variable the quality is across the composite;
  - compare and contrast the conventional and maximum OPERA composites;
  - how the OPERA and Nimrod fields compare over the UK.
  - how large the model biases are, compared to gauges;
- For DA assess the impact of:
  - data quality, and
  - increased coverage

Section 2 briefly describes the verification statistics used in this document. Section 3 discusses the findings relevant to verification, whereas Section 4 summarises the DA experiments and results. Concluding remarks follow in Section 5.

2 Review of categorical statistics

Categorical statistics are calculated from a contingency table (Table 1) which summarises the forecast performance in terms of the hits, false alarms, misses and correct rejections of whether a predicted event above a prescribed threshold actually occurred. These statistics form the backbone of most verification strategies for assessing the skill of NWP forecasting systems.

Table 1: *Contingency table showing the meaning of individual entries.*

	Obs = Yes	Obs = No	Marginal Sum
Fcst = Yes	a = Hits	b = False alarms	$(a + b)$
Fcst = No	c = Misses	d = Correct rejections	$(c + d)$
Marginal Sum	$(a + c)/n$ = Base Rate	$(b + d)$	n

One of the most important descriptive measures is the frequency bias which is the ratio of the observed and forecast events. For a perfect forecast system the frequency bias is equal to 1.

$$FB = \frac{a+b}{a+c} \quad (1)$$

A multitude of categorical skill score formulations are listed in the literature (see e.g. Mason (2003) in Jolliffe and Stephenson (2003)). Despite being sensitive to the base rate the ETS remains one of the most popular scores (Eq. 2).

$$ETS = \frac{a - a_r}{a + b + c - a_r} \; where \; a_r = \frac{(a+b)(a+c)}{n} \quad (2)$$

3 Data quality assessment

A NimrodToFDB task completes the data conversion of the OPERA BUFR files to pp format. As required by NimrodToFDB, a counts file which says how many 15 minute accumulations are available every 6 hours for each grid square is created. This is used in the initial "quality control" step. We impose a 90% data availability threshold.

3.1 Data frequency

Across the UK we now have 5 minute data update times which means that 12 time slices make up every hour. For the OPERA composite a reduced time frequency of 4 slices (every 15 minutes) is used. Figure 1 shows the impact of this reduction.

It is important to highlight how much data are excluded from the UK Nimrod product and the detail this adds to the radar mask. The other aspect to notice is that the OPERA range is about twice that used for Nimrod with a marked range effect towards the edges. The difference field shows that the OPERA accumulations over the W Scotland are much lower (by more than 20 mm). This can be ascribed to the data frequency, i.e. the upscaling of the rainfall rates (5 minute or 15 minute) to produce hourly accumulations. Overall Fig. 2(a) shows that over the last years the NAE forecast bias is greater than 1 (i.e. over-forecasting), but comparable to all the other models in the European

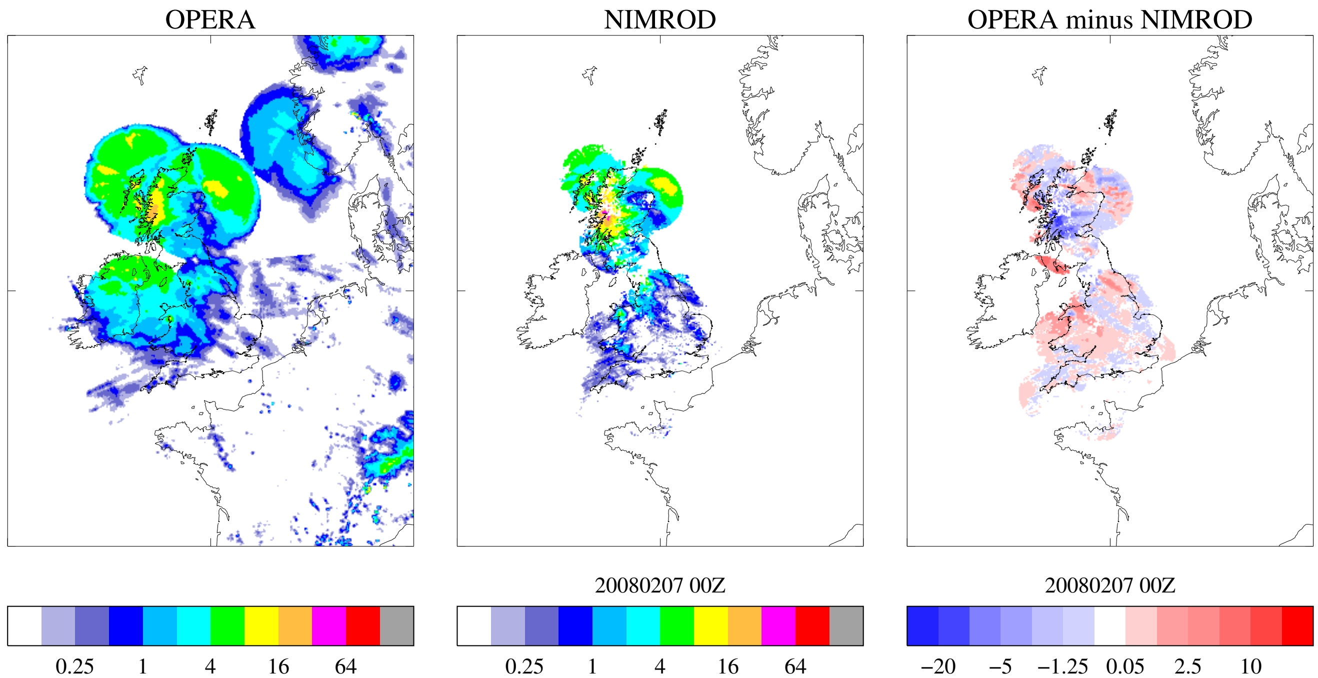

Figure 1: *Accumulation differences over the UK.*

intercomparison project. The ETS in Fig. 2 shows that we outperform the other models. This is for a 4 mm.d^{-1} threshold.

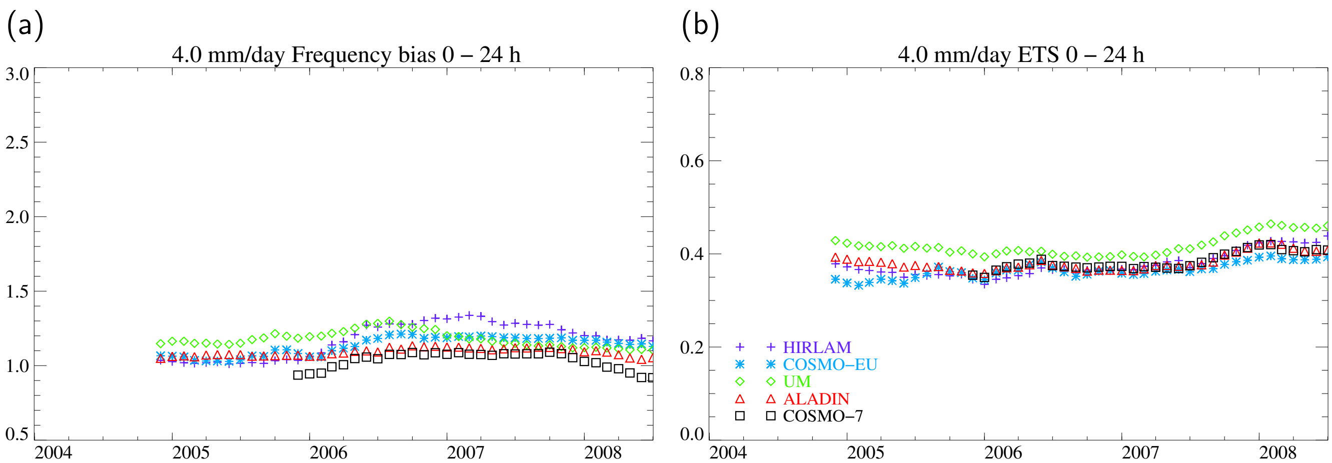

Figure 2: *NAE 12-month running means of (a) the frequency bias and (b) the ETS, against Nimrod radar-only analysis.*

The differences over the OPERA domain between model and radar composite are shown in Figure 3. The model field continuity is immediately apparent, i.e. how it fills the radar coverage mask to the edges. By contrast the actual individual radar coverages sometimes end a long way from the "mask edge" (e.g. Norway). This is largely due to the nature of the precipitation (snow) with sub-zero surface temperatures (i.e. radar sampling only in the ice) which means that the signal is weaker and the effective range of the radar is reduced.

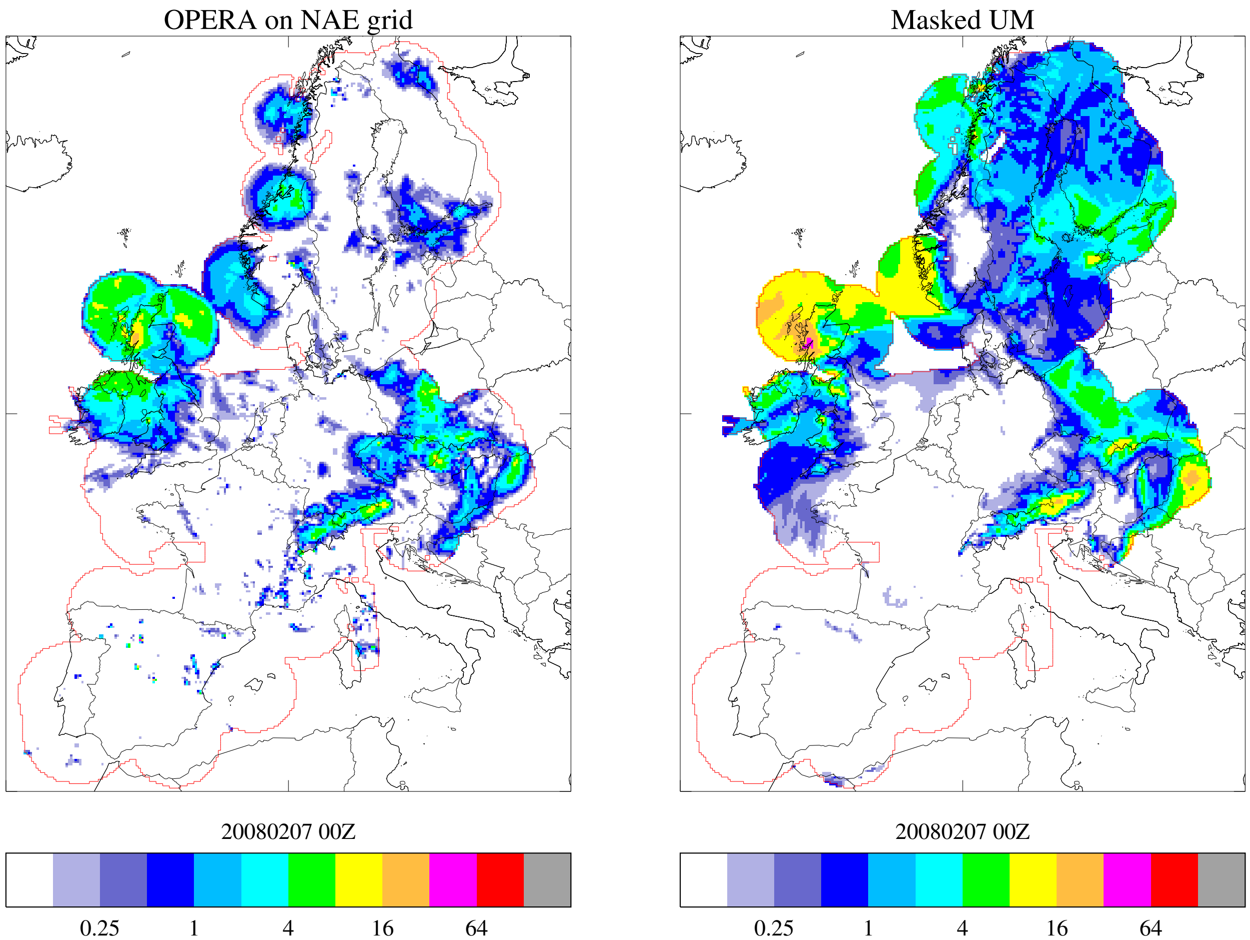

Figure 3: *OPERA and masked NAE accumulations.*

3.2 Basic data quality

This being a pilot data hub the level of quality control varies considerably across the whole composite area, which has serious consequences for downstream quantitative users. Anaprop in particular is a persistent problem, mostly over the North Sea (from the Dutch radars) but also over the Spanish interior. An example is shown in Fig. 4. Bech and Haase (2008) showed at the recent 2008 European radar conference in Helsinki just how difficult the anaprop problems are over Spain.

It has become quite apparent that many countries are not sending their "best" radar data to the pilot data hub. Some send corrected data but reduced temporal resolution, others send partially corrected data, others uncorrected data.

3.3 Data variability

Given this variability in the individual data contributions, a "summary overview" of identifying problem regions was devised, and is described by Mittermaier (2008). This is done by using the continuity in the model forecast fields at the daily time scale. To remove the impact of the bias, the radar and model accumulations are scaled by their respective maxima. This removes the bias and unit dependence, and allows evaluation in a relative sense. Figure 5 shows an example of the scaled model and corresponding radar values.

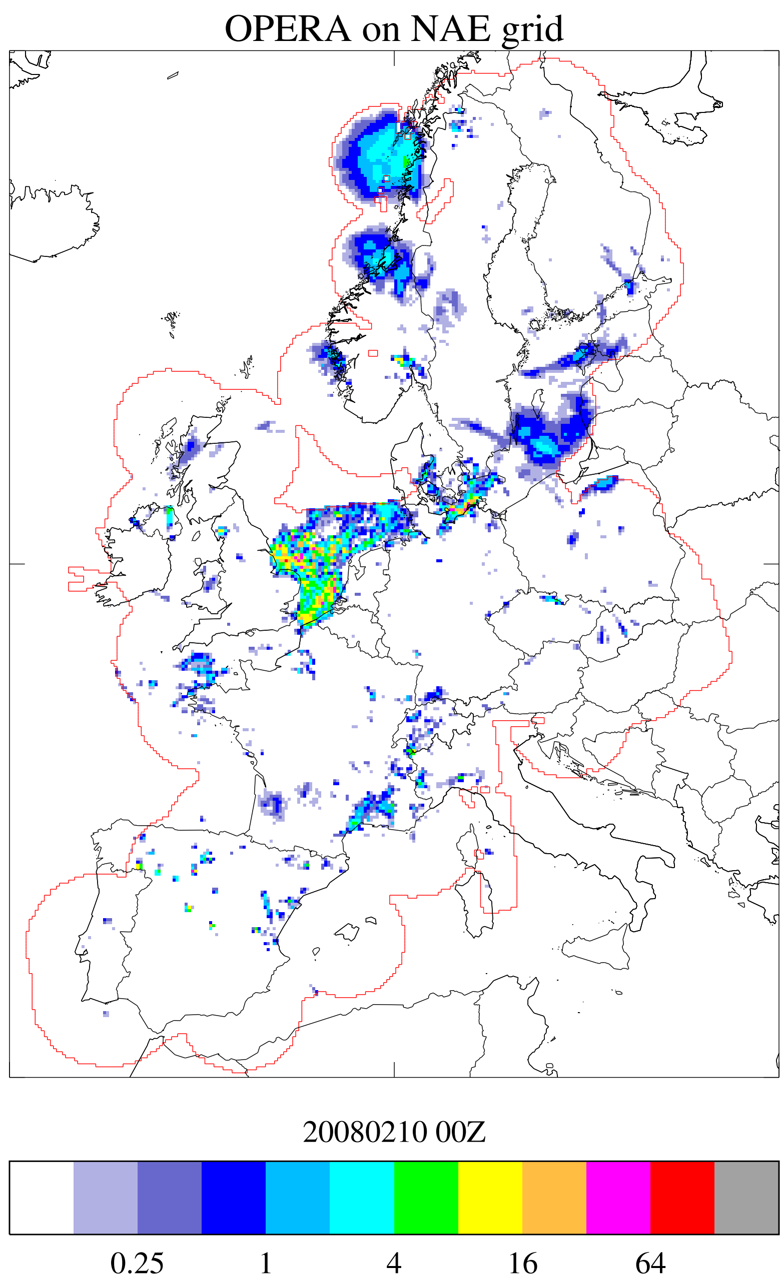

Figure 4: *OPERA accumulation for 10 February 2008, showing the anaprop over the North Sea, and interior of Spain. This is the OPERA field averaged onto the 12 km model grid.*

These scaled values can now be subtracted from each other to produce a type of anomaly difference which tells us something about features that are present in the radar data but are not in the model field. The difference is shown in Figure 6. The darker red edges show where there are range effects (i.e. the rain seems to fade away as the distance from the radar increases). One of the Norwegian radars also has a bright band problem (i.e. the radar beam intersects the melting level which acts to enhance the radar return and leads to an over-estimation if not corrected) where the area around the radar is blue, as this is where the radar has relatively speaking much more rain than the model.

The power of the diagnostic lies in the accumulation thereof over time. Problem areas stand out

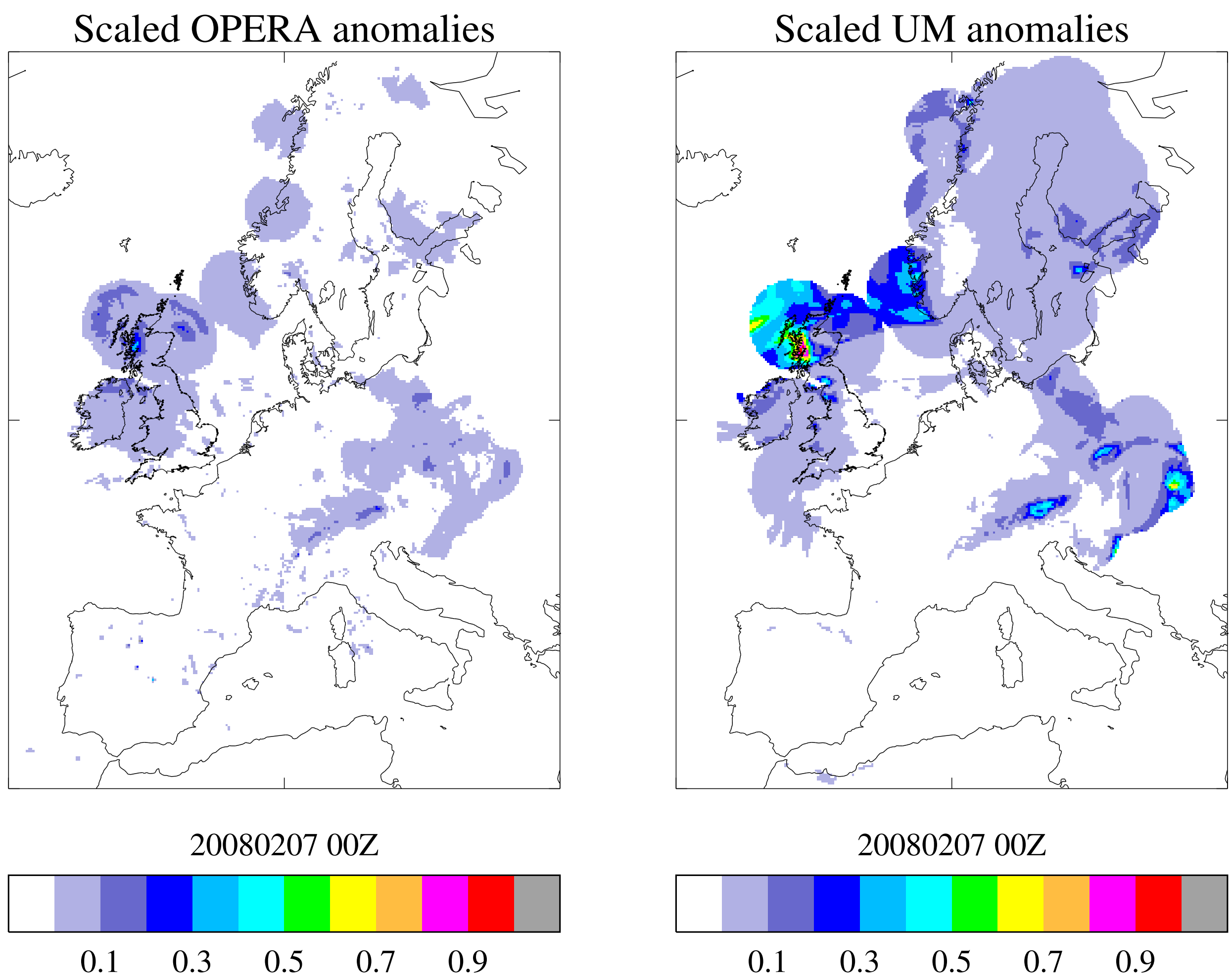

Figure 5: *Scaled model and radar accumulations (with respect to the maximum) as shown in Fig. 3.*

in bright red or blue. Two periods have been analysed between February and July 2008. Unfortunately complete data continuity could not be achieved during this time. This analysis was instrumental in identifying various problems, highlighting the usefulness of the approach and the analysis itself.

The first 35 days spanning the period between February 7th and March 12th represented the best data set, spanning the winter period. From March 13th there was 7 week break in the data due to inadvertently aggressive housekeeping which meant that the 90% data availability criterium was not met. This was fixed and from May 8th there are data but, as we subsequently discovered, without the UK composite (Irish radars are present) and Hungarian sites. A bug in the BUFR code supplied by the Technical University of Graz in Austria was found to be responsible. This was fixed and as of June 19th the data appear to be in good order.

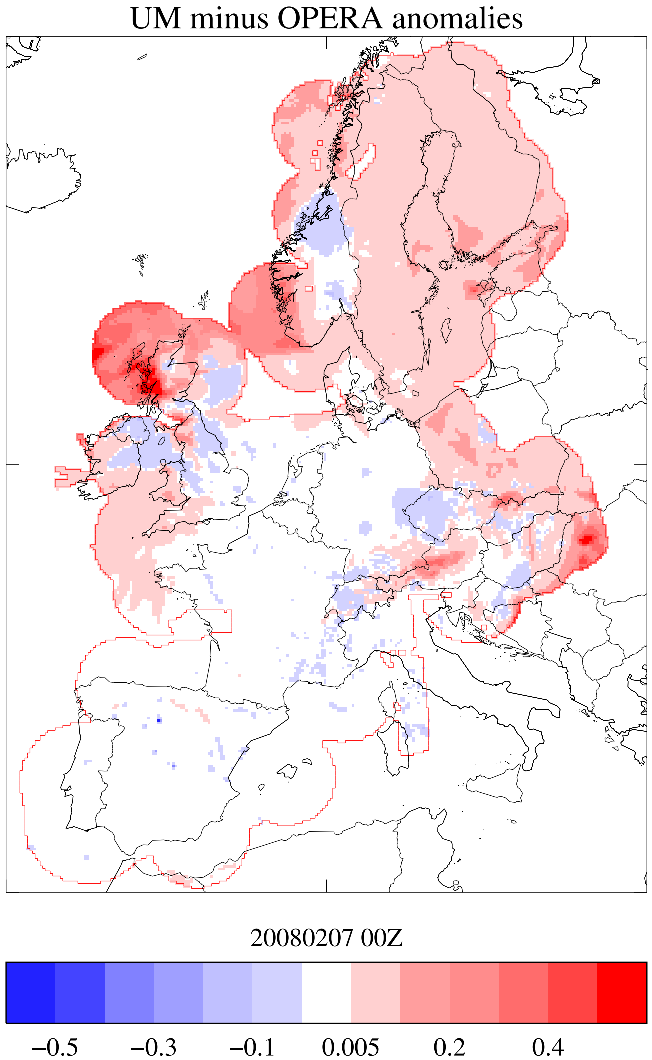

Figure 6: *Difference of scaled accumulations in Fig. 5.*

The analysis is therefore based on the initial data collected during February and March, and then the latter half of June and July. The accumulated 32- and 42-day anomalies are shown in Figures 7. It is regretable that the opportunity of tracking the continuous progression from season-to-season has been lost. We can therefore only present results for late winter and mid-summer. There is a great deal of detail to be analysed in these figures, as discussed in subsequent sub-sections.

3.3.1 Range effects

The red "rings" along the edges of the radar mask are the manifestation of range effects. The radar totals decrease with increasing distance from the radar, but the model totals do not. As you may see in Fig. 7(a) these are more pronounced in the winter months but could be prevalent all year around over cooler seas and in colder high-latitude climates, as the northern Norwegian radars in Fig. 7(b) suggest.

3.3.2 Seasonal and weather regime effects

The uniform redness over the Nordic and Baltic countries in Fig. 7(a) suggests a cold-season bias which is at its strongest over the cold seas off the Norwegian coast. Precipitation is predominantly snow, which

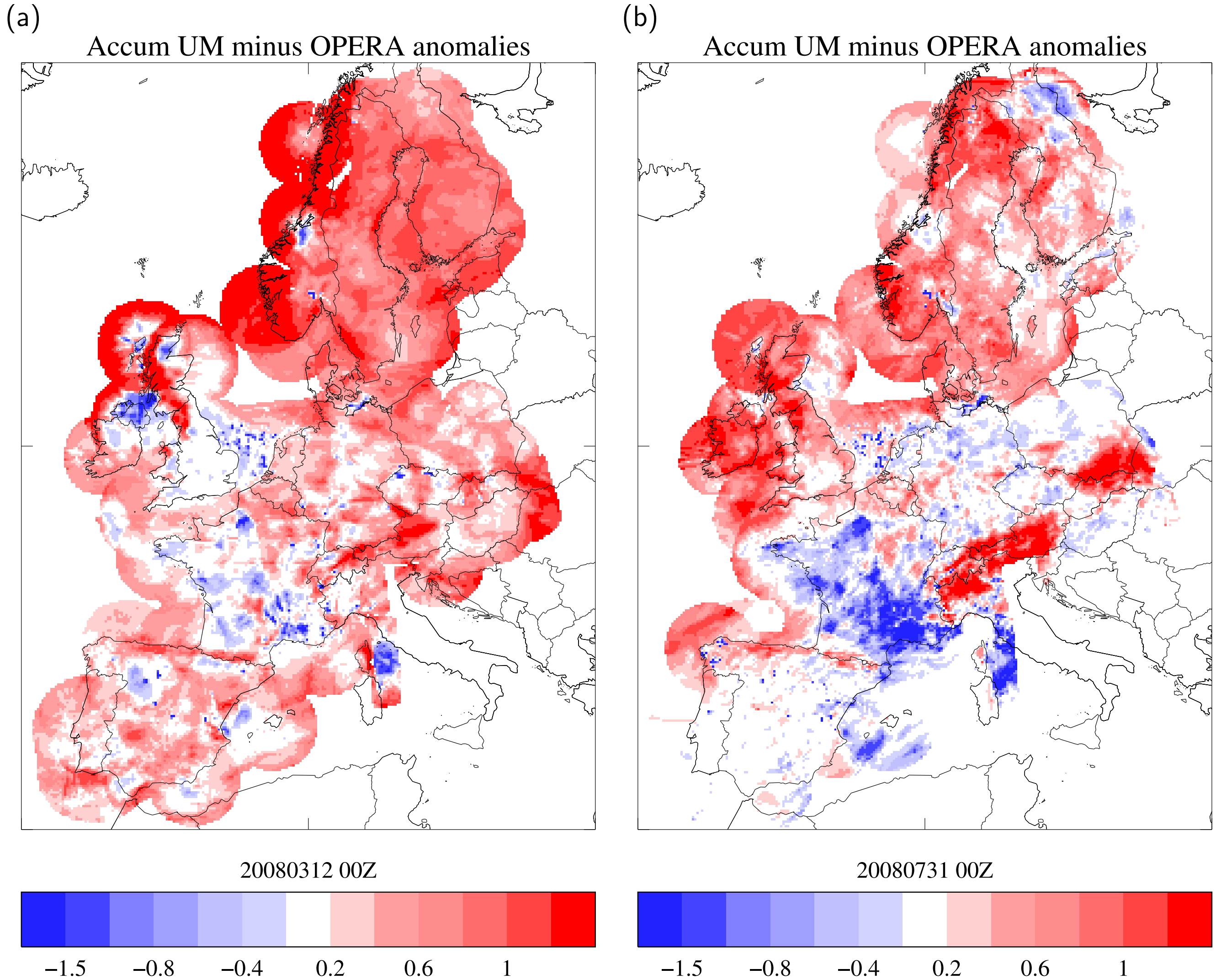

Figure 7: *Accumulated anomalies based on the conventional composite for (a) 7 February to 12 March, and (b) 20 June to 31 July.*

by definition yields much smaller reflectivities and therefore precipitation rates. We do not receive corrected data from Norway or any of the BALTRAD countries so the impact seen here is as bad as it gets.

Whilst it is claimed that Spanish radars tend to under-estimate rainfall in the summer, Fig. 7 indicates no systematic trends in the summer months over the Iberian peninsula.

Southern France in particular turned very blue in the summer, i.e. the OPERA composite has much larger anomalies compared to the model). There could be several reasons for this observed behaviour in Fig. 7(b), but given the dominant convective regime it may well be deficiencies in the model rainfall peaks (leading to smaller anomalies) as well as too-large radar-rainfall peaks due to the contamination of the estimates due to the presence of hail or the bright band in the stratiform parts of decaying thunderstorms and mesoscale convective systems (MCS).

3.3.3 Orographic effects

The strong red signal coincident with the Norwegian coast, over Highland Scotland and Cumbria and to a lesser extent along the north Spanish coast, Pyrenees, and Alps (in the winter, stronger in the summer) are at least in part due to the fact that the model gets "over-excited" when producing rainfall over high ground or rugged coasts. Added to this is the reduced anomalies produced by the OPERA composite given the problems that radars have in mountainous terrain. The strong anomaly over the Alps in Fig. 7 is very striking. Even more striking is the fact that the anomalies are picking up the broad valley between the Swiss Jura and the Alps (narrow band of blue).

3.3.4 Vertical profile and bright band effects

Although the physical radar locations have not been plotted in Fig. 7, the circular blue regions that appear to be centrally located within a circular radar area are the bright band signal. Again this is more obvious in the winter months when the freezing level is low. French and Spanish radar data in particular do not appear to be corrected for vertical profile/bright band effects. At least some of the blue (OPERA) anomalies over France in Fig. 7(b) are also due to stratiform rainfall.

3.3.5 Anomalous propagation

The Dutch radars overlooking the North Sea are plagued by anomalous propagation *a lot of the time*, as shown by the blue speckles in Fig. 7, and this seems to be independent of the season. Spain also suffers from anomalous propagation, which is possibly worse in the summer, but is observed all year around (Bech and Haase, 2008).

Based on the above, it is clear that there are similarities in the patterns when comparing winter and summer. These similarities are the true indicators of radar quality and underlying corrections that have or have not been applied. Clearly no region is without a problem of some description or another.

3.4 Conventional vs maximum composites

In addition to the conventional composite, a "max" composite is also produced by the pilot data hub. This is based on using the maximum value in the vertical at a given location (where available). In addition the maximum value in a radar overlap is then also used. This has also been tested and the results shown in Fig. 8. The maximum composites were not available during the winter period.

From Fig. 8 it is clear that using the maximum composite acts to swing the anomalies quite strongly in areas where maximum reflectivity data are available. Over Spain for example there is no change, when compared to Fig. 7(b).

The radar anomalies over Germany and Switzerland are the most dramatic, changing from very bland (near-zero) values to very negative values, suggesting that the radar anomalies are now much enhanced. The same is true over E France. The redness over the Alps has been confined to south of the Alps, suggesting here that in the presence of a southerly flow the model still acts to over-estimate upslope precipitation and enhance the anomalies through the maximum.

Some of the redness over Scandinavia has disappeared, thus reducing the anomaly differences to be near zero when the maximum composite is used. Most of England and Wales, which has near-zero

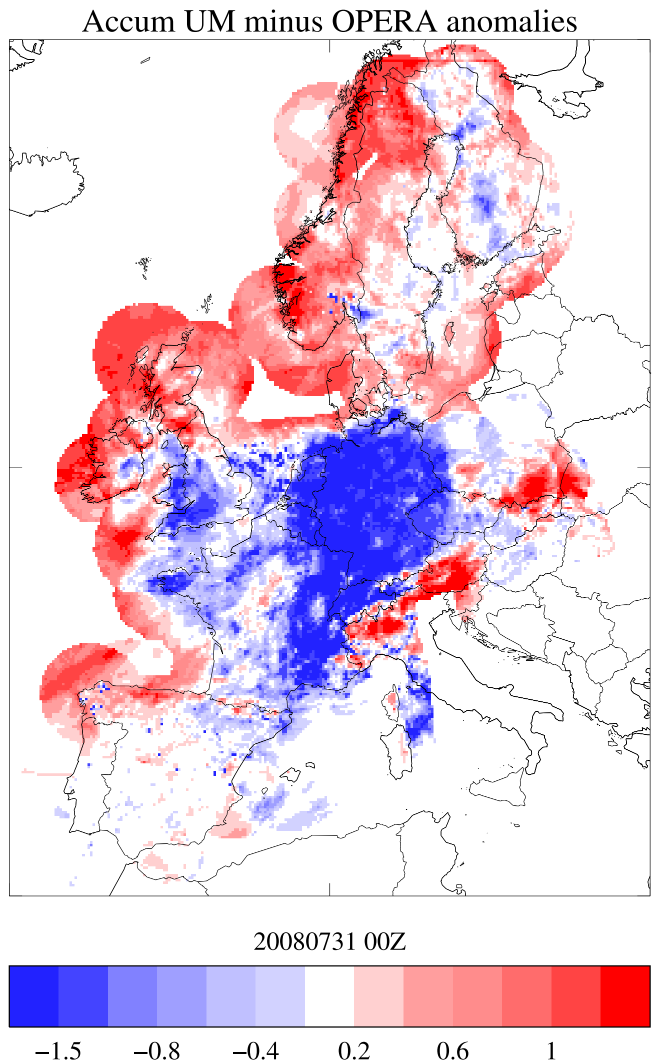

Figure 8: *Accumulated anomalies based on the maximum composite (not available for the first period).*

anomaly differences in Fig. 7(b) is now blue, i.e. radar anomalies are dominating the accumulation. Overall the only area where the maximum composite appears to match the model anomalies better is over the Baltic countries, where data are not being corrected for range effects and precipitation type.

3.5 Anomalies using the UK operational composite

An important part of the methodology (and the comparison) is to see how the UK radar composite compares to the OPERA results over the UK. We know from the verification scores plotted in Fig. 2 that the bias is generally greater than 1 (i.e. the model predicts too many occurrences of a given event), and this is true for all thresholds. So, it is therefore unsurprising that in Fig. 9 the anomalies are positive (i.e. the peaks in rainfall are too large when compared to the radar *except* over orography where the radar enhancement is clear. This is particularly evident during the February-March period. There are some similarities in the pattern for the summer period, although the large blue anomalies over orography are less evident and the positive anomalies are reduced over southern central England. Notice also that the maximum range between the summer and winter months is different, especially noticable in Scotland. On balance there are no detectable systematic radar problems in the Nimrod composite, other than the orographic signal. Mittermaier (2008) has investigated this and found evidence of a positive-negative dipole in anomalies, especially along the Welsh coast suggesting that the model placement of rainfall under strong upslope rain events is slightly out of phase with the orography.

These results contrast somewhat with the positive anomalies found over these parts in the OPERA composite. There are two potential reasons. The maximum value used for scaling is specific to the domain, i.e. these values will be different for the UK-only and OPERA composites. Secondly the radar data frequency will have a strong impact on orographic rainfall, as discussed in Section 3.1.

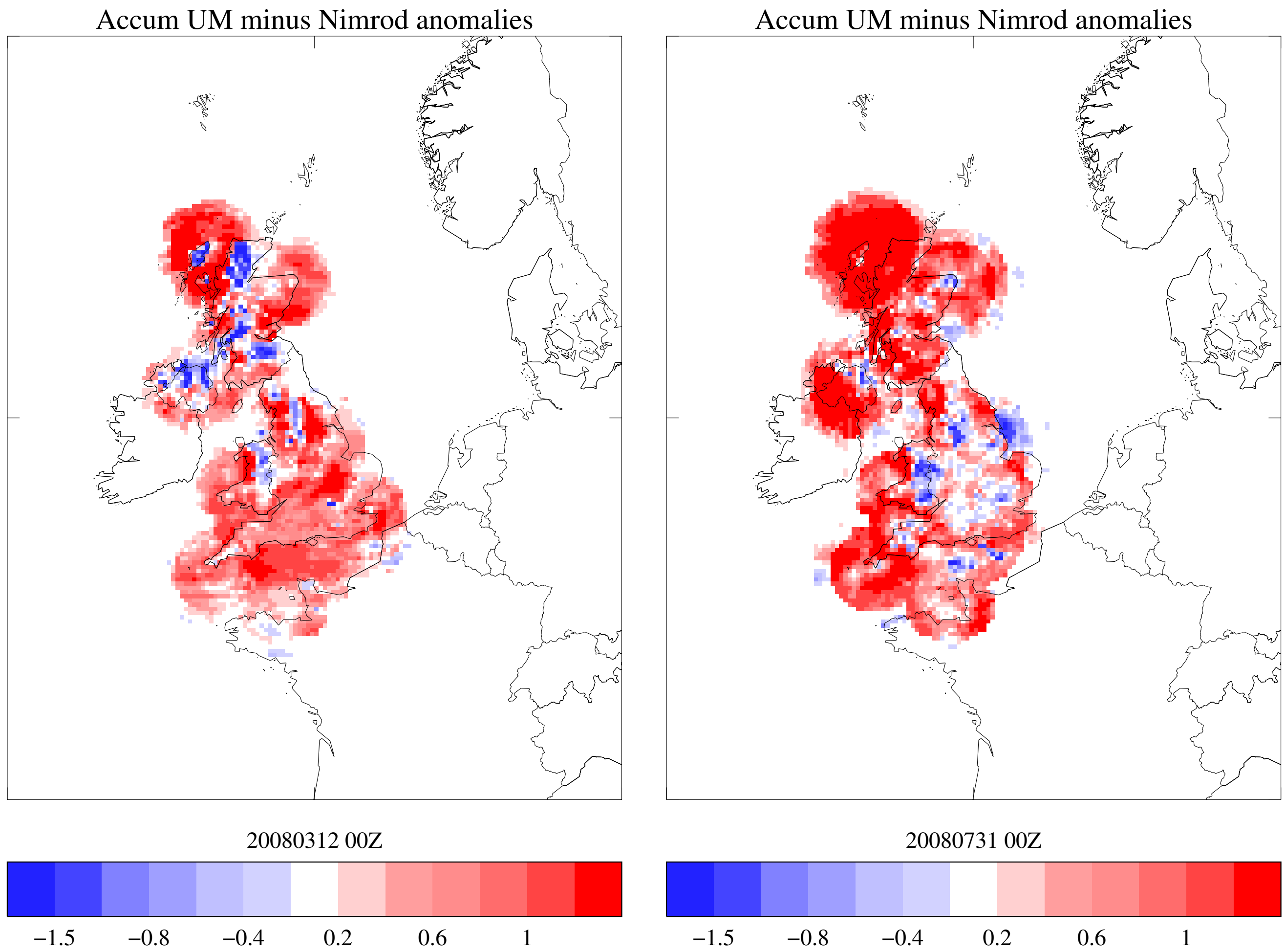

Figure 9: *Accumulated anomalies based on the UK Nimrod composite.*

3.6 Model error against gauges

One of the key questions that gets raised is "what about the model accuracy?". This is addressed in Fig. 10 which compares the long-term model *rmse* error between the gauges and the daily model accumulations. It should be pointed out that gauges are not necessarily ideal when considering convective rainfall, therefore some of the results in the Mediterranean region could well be unrepresentative, but overall errors are very widely less than 4 mm.d^{-1}. The problem areas for the model are tied to the high ground and the coast where errors increase sharply. Note that this is a very crude comparison as this simply uses the nearest model grid point to a given gauge location.

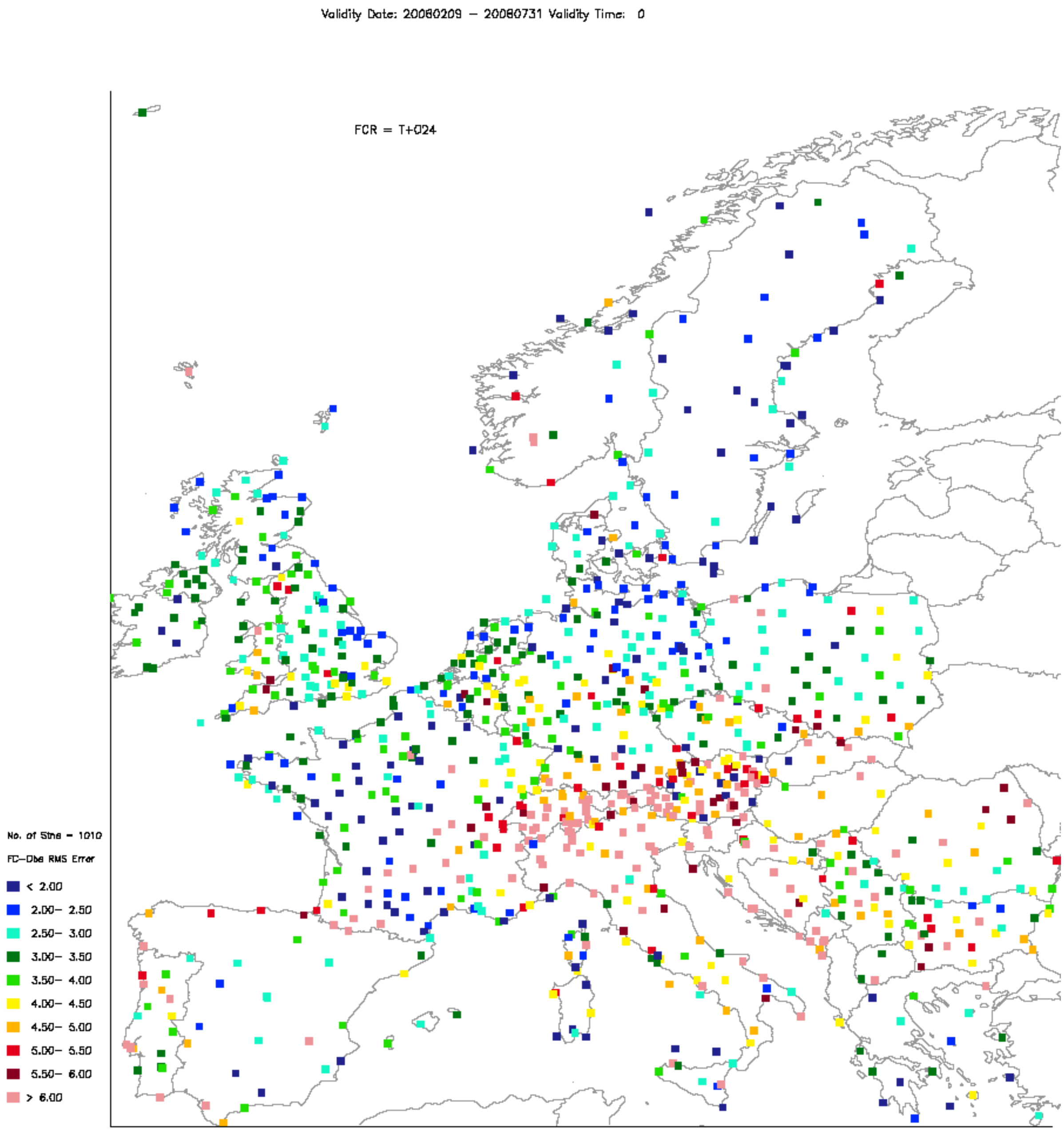

Figure 10: *Model-gauge comparison across the OPERA composite region.*

4 Data assimilation tests

Whereas the quality assessment (as described in the previous sections) used daily accumulations from the operational 12 km North Atlantic European (NAE) model 00Z run, the assimilation studies were performed using the 24 km grid-length version of the NAE. The data assimilation scheme is 4DVAR (Rawlins *et al.*, 2007) for most observation types, together with a nudging scheme for cloud and rainfall data. The nudging scheme allows the model to assimilate hourly instantaneous surface precipitation rate estimates from a radar composite via Latent Heat Nudging (Macpherson, 2001). The weight given to rainfall rates during assimilation depends on distance from the nearest radar and beam height above freezing level.

4.1 Experiments

Three experiments were run for extended periods with full forecast verification to examine different issues. A control NORAIN experiment runs without rainfall assimilation of any kind. A second UKOPER experiment runs with the UK Met Office operational radar composite with partial coverage over Europe facilitated by bilateral agreements with several National Met Services (Fig. 11(a)). The OPERA experiment runs with a composite giving maximum European coverage and derived from a mix of single site data and national composites (11(b)) Relative to the UKOPER area, it adds Scandinavia, Iberia, and a large section of mainland Europe approximately east of a line from Denmark to Italy.

(a) (b)

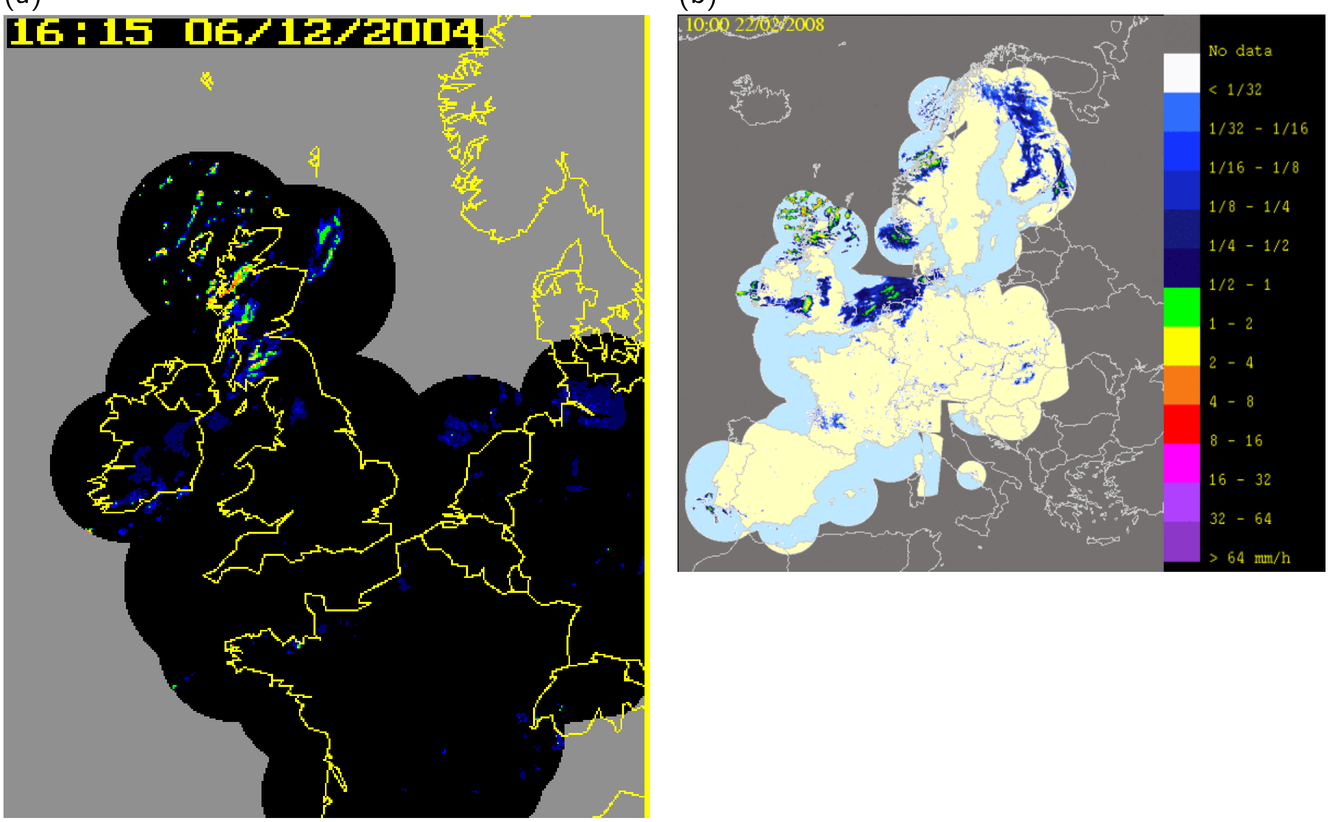

Figure 11: *Examples of the 15-min (a) European Nimrod rainfall rate composite and (b) OPERA rainfall rate composite to show coverage.*

4.2 Results

Comparison of UKOPER with NORAIN results gives an indication of benefit of the current operational assimilation capability when fed by a limited composite derived from harmonious processing of single site data. The OPERA experiment contains potentially beneficial signals from the extended data coverage, but also possible problems associated with the inhomogeneous nature of the radar processing applied within the various inputs to the composite.

The first experimental period ran from March 1st to March 26th. The second began on April 15th and was stopped on May 26th. Two 48-hour forecasts are run daily from 00 and 12 UTC. The four 6-hour forecasts run as part of the daily data assimilation cycle are also verified. Six-hourly rainfall accumulations are verified against SYNOP reports, as are other near-surface variables.

In order to maximize the signal from differences between our 3 different assimilation experiments, we focus here on results verified at forecast range t+6h. We have twice as many forecasts verified at this forecast range, compared to later ranges, thus improving statistics. We also expect the impact of radar assimilation will usually be smaller at later forecast times. Looking at the impact early in the forecast also gives us a better chance of correlating verification signals in smaller regions with changes in radar coverage in those same regions. Since the observation type we assimilate is rainfall rate, we focus on skill of rainfall forecasts, and in particular the Equitable Threat Score (ETS, Eq. 2) for the 6-hour accumulation to t+6h. Since the LHN technique is more likely to give greater impact in cases with intense rainfall, we restrict ourselves here to the highest rainfall amount verified with a threshold of > 4 mm in 6 hours. Verification regions used were the whole NAE area (enclosing all of Fig. 11(b), a UKMES area like that in Fig. 11(a) and various combinations of WMO block numbers giving rise to areas best labelled as: UK, Scandinavia, France, Germany, Iberia, Central Europe, Eastern Europe. Central Europe includes WMO blocks 6 and 10, while Eastern Europe covers blocks 11–16. Subsequent tables show ETS values for each experiment and verification area, for both experimental periods.

Table 2: *ETS scores for t+6h rainfall accumulation for the threshold >4mm/6-hours, for the first period March 1st to 26th, 2008.*

	NORAIN	UKOPER	OPERA
NAE	0.352	0.358	0.348
UKMES	0.363	0.375	0.348
UK and Ireland	0.336	0.344	0.308
France	0.333	0.351	0.337
Germany	0.320	0.344	0.326
Scandinavia	0.348	0.358	0.353
Iberia	0.357	0.347	0.378
Central Europe	0.348	0.361	0.340
Eastern Europe	0.360	0.358	0.361

For the two large areas (NAE and UKMES) and for the UK and Ireland, the results in both periods show a positive benefit from radar assimilation when we compare UKOPER with NORAIN. In these areas and in both periods, the OPERA experiment verifies worse than either UKOPER or NORAIN. If these results are typical, this might imply that the quality of the OPERA composite is as yet limited by the differences in processing between it and the UKOPER processing. For example, no VPR correction scheme can be run on much of the single site data received by the data hub, since it comes in Constant

Altitude Plan Position Indicator (CAPPI) form and the scan elevation angle is unknown. Also, the spurious echo removal scheme functions only on the UKMES area, and not on those parts of the larger NAE area outside it. It was, at first, slightly surprising that the OPERA composite apparently performed less well over the UK, since that is a region where the UKOPER and OPERA processing should be more similar. In early June the verification analysis established that there was no UK composite data in the OPERA composite from March 13th, so in actual fact the full impact over the UK can not be quantified.

The OPERA run does show benefit relative to NORAIN in France and Germany for March 2008, but does not match UKOPER. During April and May, OPERA verifies worse than UKOPER over both regions and worse than NORAIN over Germany.

Table 3: *ETS scores for t+6 rainfall accumulation for the threshold >4mm/6-hours, for the second period April 15th to May 26th 2008.*

	NORAIN	UKOPER	OPERA
NAE	0.301	0.312	0.291
UKMES	0.314	0.336	0.292
UK and Ireland	0.284	0.321	0.213
France	0.305	0.336	0.326
Germany	0.346	0.349	0.334
Scandinavia	0.333	0.307	0.271
Iberia	0.319	0.327	0.315
Central Europe	0.365	0.355	0.351
Eastern Europe	0.263	0.273	0.256

In regions where the OPERA composite has greater coverage than the UKOPER one, we see some evidence of benefit in 6-hour rainfall forecasts. For Iberia during March, the OPERA experiment verifies better than either NORAIN or UKOPER. The same is marginally true for Eastern Europe. However, the OPERA run does not improve on UKOPER for Scandinavia. Also, the OPERA results for April-May are worse than NORAIN for Scandinavia, Iberia and Eastern Europe, the 3 areas with the most extra coverage. This again may point to a need to optimize the radar processing within the data hub for data received from these regions.

Analysis of the frequency bias statistic revealed that the OPERA run produces fewer events of heavy rain than UKOPER. For the April-May period, the UKOPER bias was 0.99 over the NAE area, while the OPERA run gave 0.90. This signal probably comes from the simple averaging method in the OPERA compositing technique, in which inclusion of distant radars, which might miss rain at long range, will tend to reduce rates in the composite. The UKOPER composite, on the other hand, selects the single radar with lowest beam height.

A third period was run from 28th May to 11th June to allow comparison of the conventional OPERA composite with the max reflectivity one. In terms of ETS at t+6h, the max reflectivity run performed better, giving a value of 0.247, compared with 0.227 for the conventional composite, over the NAE area. This apparent improvement was obtained, however, at the expense of increasing the frequency bias for rain events with accumulation > 4mm/6-hours from 1.12 (conventional composite) to 1.36 (max composite). The results for max reflectivity were also worse for longer lead times, smaller rain thresholds and other model variables such as *pmsl*. Surface pressure forecasts were significantly worse, as seen in Fig. 12. The impact of the conventional OPERA composite is neutral on the *rmse*.

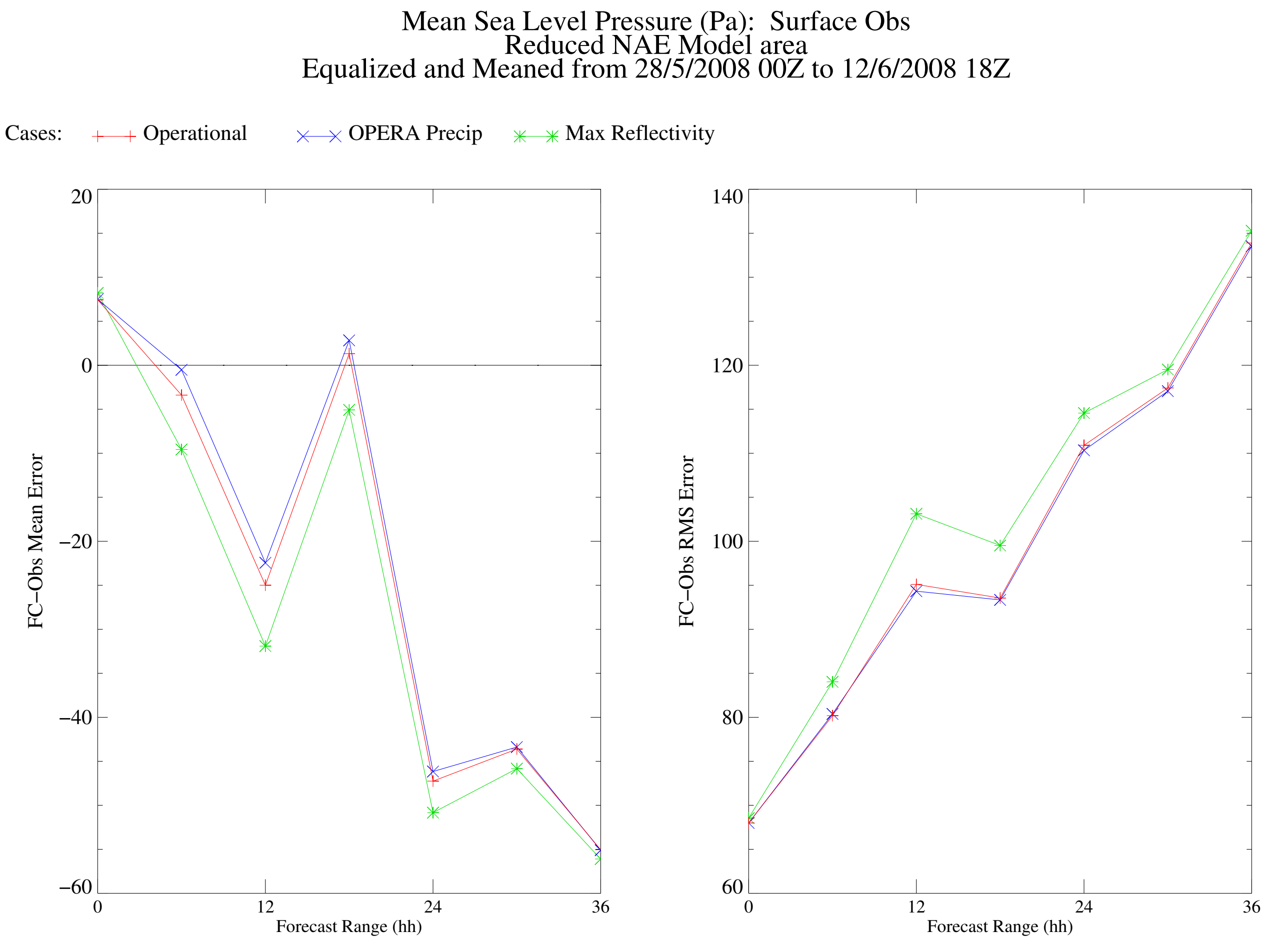

Figure 12: *PMSL differences (in Pa) between the two different OPERA composites.*

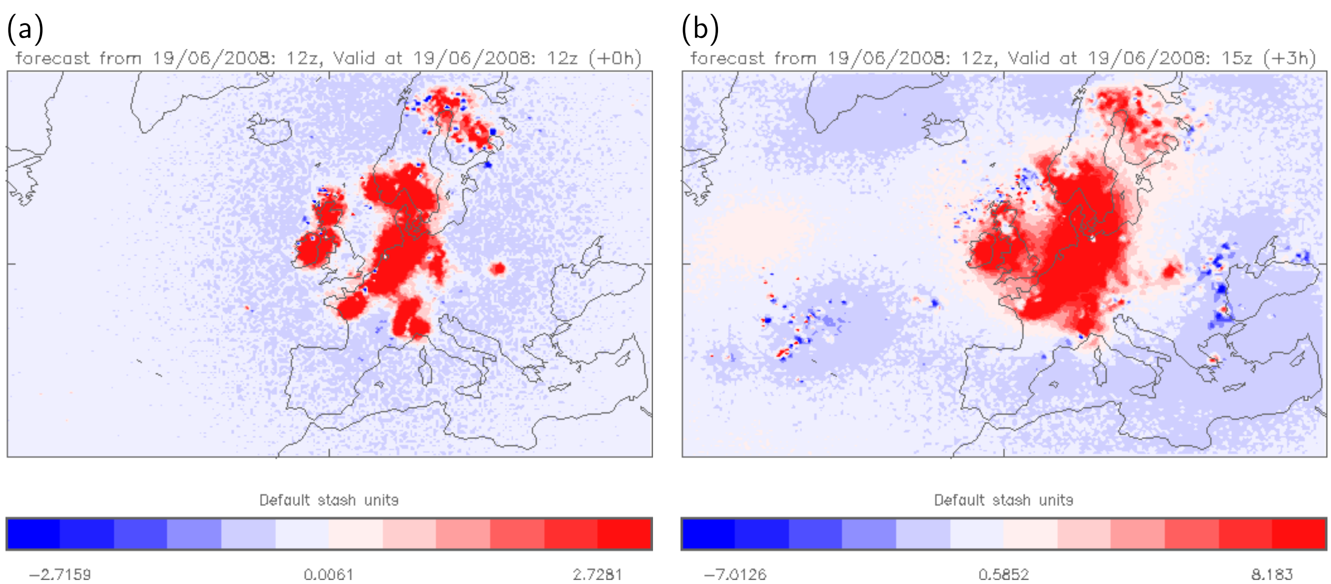

Figure 13: *Increment differences (OPERA - OPERA Max) for* pmsl *at (a) t+0h and (b) t+3h.*

Figure 13 shows the *pmsl* increment differences (OPERA minus OPERA Max) at t+0h and t+3h. The OPERA Max forecast has a negative *pmsl* bias (red) with respect to the standard OPERA run. Again the units are Pa.

5 Summary

The objective of this study was to highlight data quality issues and suggest improvements, and make a recommendation to what extent these composites can be used for verification of model forecasts and data assimilation.

The study showed that the inherent continuity in the model forecast accumulations can be used to identify the full range of potential radar errors, and these have a detrimental impact on the overall quality of the composite.

Some model behavioural issues have also already become clearer, e.g. slight misplacement of precipitation over high ground, and over-estimation of rainfall over high ground.

Preliminary results from data assimilation trials confirm earlier studies pointing to some benefit for short-period rainfall forecasts from the assimilation of a radar composite into the NWP model. The OPERA composite in general performs less well than the UKOPER one and in some cases less well than the NORAIN run without radar. Assimilation of a composite based on maximum reflectivity leads to forecasts with too much heavy rain and less accurate surface pressure. More extensive runs are required to establish the reliability of the signals present in this preliminary sample of results.

We therefore recommend that:

- Data availability be monitored carefully.
- Change management must be much more structured, along with testing.
- Single site data are sent to the data hub (this has been agreed for the operational hub), with an improved and harmonised approach to data quality and correction.

The full composite shows much promise and potential for the future. To make this a reality will require considerable investment, during the operational implementation phase, and beyond.

References

Bech, J. and Haase, G. (2008). Do anomalous propagation conditions occur frequently with precipitation? In *Proceedings of Fifth European Conference on Radar in Meteorology and Hydrology*.

Jolliffe, I. T. and Stephenson, D. B., editors (2003). *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley and Sons.

Macpherson, B. (2001). Operational experience with assimilation of rainfall data in the Met Office mesoscale model. *Meteor. Atmos. Phys.*, **76**, 3–8.

Mason, I. (2003). Binary events. In I. T. Jolliffe and D. B. Stephenson, editors, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, chapter 3. John Wiley and Sons.

Meischner, P., editor (2004). *Weather radar, principles and advanced applications*. Physics of Earth and Space environments. Springer.

Mittermaier, M. (2008). Using model precipitation forecasts for monitoring radar data quality. *In preparation*.

Rawlins, R., Ballard, S., Bovis, K., Clayton, A., Li, D., Inverarity, G., Lorenc, A., and Payne, T. (2007). The Met Office global 4-dimensional variational data assimilation scheme. *Quart. Jour. Roy. Meteorol. Soc.*, **133**, 347–362.

Rossa, A. (2006). Use of radar observations in hydrological and NWP models. Technical Report EUR 21954, COST.