UNIVERSITY OF READING
DEPARTMENT OF METEOROLOGY

Probabilistic rainfall forecast using a neighbourhood approach

By Thomas Loridan

A dissertation submitted in partial fulfilment of the requirement for the degree of Master of Science in Weather, Climate and Modelling.

August 2006

Abstract

Forecasts of rainfall from Numerical Weather Prediction (NWP) models are not yet as accurate as they are for other predicted fields like temperature or pressure. The high spatio-temporal variability of such a field as well as the errors implied by sub-grid scale processes limit considerably its predictability.

Probabilistic forecasts are a way to take this inherent uncertainty into account, and many methods have recently been developed to tackle this issue. Among these diverse approaches, a low-budget procedure using a spatio-temporal neighbourhood was introduced by Theis et al (2005). Its particularity is to use only information readily available from the Direct Model Output (DMO), and for this reason we chose to reproduce the methodology to post-process the Met Office Unified Model's outputs.

The variable selected to be post-processed was the rainfall rate, in order to verify the forecasts against radar observations. Once the initial neighbourhood method was correctly implemented and provided reasonable results, a few modifications concerning the distribution of weights in both space and time dimensions were tested. In order to consider different weather conditions, two distinct case studies were used: A convective episode and a frontal system.

The main result of the project was the case-dependent aspect of the procedure's performance, and the fact that a Gaussian distribution of weights in space could improve the subjective interpretation of the post-processed field.

Acknowledgements

I would like to thank my supervisor Dr. Robert Plant for his patience and help over the dissertation period and the whole year. I appreciated his technical advice as much as the personal guidance regarding my future projects. My thanks also go to all the friends I met here who made this year so enjoyable, and to my parents who supported me in my plan to come and study in the UK.

Contents

Chapter 1

Introduction

1.1 Forecasting rainfall

Out of all information given by a weather forecast, precipitation is certainly the one many users are most interested in. Motivations for this interest are diverse and involve great security and economic concerns. Even though it is not the main focus of this project it is important to realise how strongly our lives can be affected by the efficiency of such a forecast: Food supplies, especially in developing countries, depend on the productivity of local agriculture and this productivity can be highly improved by an accurate forecast of the rainfall. On this economic side we can also mention hydroelectricity management and the necessity of estimating energy production properly. Even more important are the issues where human lives are at risk, and flood forecasting is perhaps the area which would benefit the most from any progress in rainfall prediction: Warning systems are based on hydrological numerical models which are coupled with atmospheric ones to give a precipitation forecast over a drainage basin (Benoit and Pellerin 2000, Jasper and Kaufmann 2003). Therefore the whole procedure is extremely dependent on the accuracy of the initial rainfall prediction, and unfortunately the time scale for which this forecast remains reliable is often too short for preventive action to be taken in case of security threats (Krzysztofowicz et al. 1993).

This great need for accuracy does not match the actual efficiency of rainfall forecasts, and precipitation field is one of the least successfully simulated by Numerical Weather

Prediction (NWP) models. The Working Group on Numerical Experimentation of the World Meteorological organisation has underlined that although the major NWP models operationally used have made some important progress recently, they still experience some difficulties when it comes to producing some Quantitative Precipitation Forecasts (Erbert et al. 2003). The spatial and temporal scales considered as well as the diversity of physical processes involved (large scale ascent of moist air, convection, orographic lifting) are some of the many parameters which influence the level of predictability. Cassati, Ross and Stephenson (2004) have applied a new method called the intensity-scale approach to test the influence of precipitation intensity and spatial scale on the performance of the NIMROD operational system. Their conclusions were that poorer skills were obtained in small-scale events ($<$40km) involving localised intense precipitation.

Walser et al (2004) have applied an ensemble method (see section 2.1) to a mesoscale model, in order to investigate the predictability dependence of Quantitative Precipitation Forecasts (QPF) to some of these parameters. Using a cloud-resolving model with a mesh size of 3km they were able to show that even though modern efforts to decrease the grid length to a few kilometres lead to better representation of convective cells, it does not necessarily imply any predictability improvement. In fact the rapid growth of error due to the chaotic aspect of the atmosphere (Lorenz 1963), and more specifically here of its moist dynamic component lowers this predictability at small spatial scales. As previously parametrised sub-grid scale phenomenon become more realistically modelled in high resolution NWP models they tend to help the growth of small perturbations, leading to poorer predictability. Even though this decrease in predictability should be seen as a progress in the representation of the nonlinear behaviour of the atmospheric dynamics at these scales, it suggests that the accuracy of precipitation forecast will probably not be significantly improved in the years to come. Furthermore, they were able to show that even at larger scales (up to 100km) convective episodes still act to limit the predictability, meaning that realistic modelling of convection might require a different approach to the forecast interpretation at all scales.

1.2 Added value of probabilistic forecasts

Because of this unavoidable uncertainty in the NWP models' output, taking a deterministic approach to rainfall forecasting is not desirable and a probabilistic information can be useful in both the conception and the interpretation of these forecasts. Figure 1.1, obtained with the method explained in chapter 3, gives an example of such an output. The left hand part is the direct model output of rainfall rate obtained with the Met Office 4-km grid Unified Model, and the right hand one shows the probability of exceeding a threshold value of $5.10^{-4} kg.m^2.s^{-1}$ (corresponding to all coloured areas on the left-hand plot).

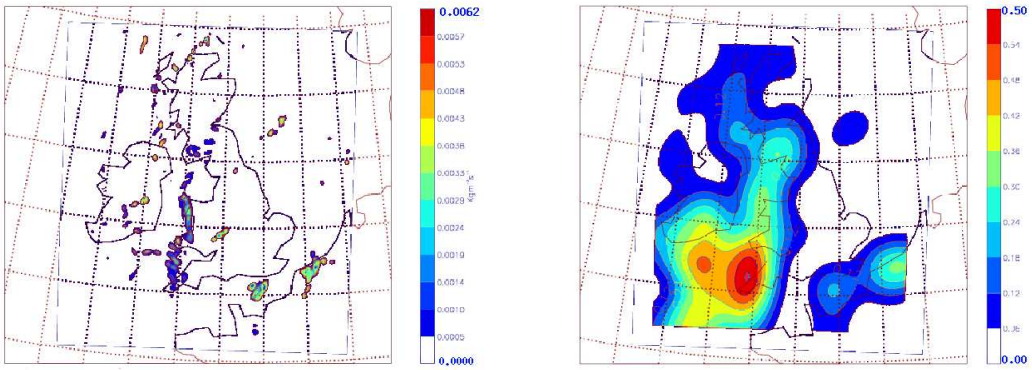

Figure 1.1: Direct model output for rainfall rate in $kg.m^{-2}.s^{-1}$ (left) and post processed output for probability of exceeding $5.10^{-4} kg.m^{-2}.s^{-1}$ (right).

This type of information provides the user with a measure of the likelihood of the prediction, giving a more realistic view of the model's abilities. Allan H. Murphy (1991) discusses the use of such probabilistic information as a link between the forecaster's judgement and the rational decision making process which follows. When looking at NWP outputs, operational forecasters try to estimate how confident they can be in the result, using their experience and knowledge as well as any recent observations they can gather (radar or satellite imagery, analysis charts) or other models' output. The main argument introduced in the discussion concerns the difficulty of translating the forecaster's subjective view of the situation in a quantitative way. The focus of the paper is on the

		Occurrence	
		NO	YES
Action	NO	0	L
	YES	C	$C + L - L_1$

Table 1.1: Expense matrix

forecast of rare events, and in this kind of situation the forecaster is only offered two possibilities: occurrence or non-occurrence. When an event has been forecast but did not occur it is referred to as a type 1 error, whereas the other situation where it was not forecast but actually happened is referred to as a type 2 error. It has been observed that when confronted to an ambiguous situation, the perspective of possible impacts on the user drives the forecaster to issue type 2 error forecasts more often. This common feeling that a type 2 error is more serious has lead to a significant level of over-forecasting and probabilistic forecast could help solve this issue by illustrating more realistically the forecaster's feeling.

Apart from the forecaster, users would also benefit widely from probabilistic information: Decision makers who base their actions on the likelihood of a weather event all have complex and different consequences to these actions, and a probability of occurrence would suit better this diversity. By giving a categorical yes/no forecast we force them all to the same decision no matter what their costs are: To differentiate this wide variety of users, the cost-loss relationship they face is often referred to. In a study by D.S.Richardson (1999) on the relative economic value of the ECMWF ensemble prediction system, this type of decision model is presented in the following way: Each preventive action is characterised by a cost C, the lost following a non protected action is noted L and the portion of the loss reduced when acting is L_1 (see expense matrix in table 1.1).

The aim of the decision maker is of course to minimise his expense, and if he is not given any hint (only climatological information available) the only reasonable choice is to either always or never act depending on the global cost of each option. If O represents the number of times the event occurs, always acting leads to an expense of $C+O(L-L_1)$ whereas never acting to OL. Therefore without any hint the forecaster should always act if $C+O(L-L_1) \leq OL$ and never act otherwise. Any forecast should offer a way to reduce

this expense E_{clim}. A common way to estimate the efficiency of a forecast is through the Relative Value RV which relates the reduction in expense provided by the forecast to a hypothetical perfect one (for which the expense would be $E_{perfect} = O(C + L - L_1)$):

$$RV = \frac{E_{clim} - E_{forecast}}{E_{clim} - E_{perfect}} \tag{1.1}$$

RV can be seen as a percentage of what the savings using a perfect forecast would be: If the use of a perfect forecast leads to an expense reduction of S, a forecast of Relative Value RV will save the user 100RV% of S.

Using this Relative Value as an index, D.S. Richardson has compared the skills of both deterministic and probabilistic forecasts. When considering deterministic systems, RV can be expressed as a function of the cost/loss ratio $\alpha = \frac{C}{L}$, providing a way to take into account different users' concerns. The main conclusion is that the forecast is not equally useful for all, and if some users with a cost/loss ratio of α between 0.1 and 0.5 do benefit from a positive value of RV, others with bigger α would not find any help in the forecast: The value of the prediction strongly depends on the user's α.

On the other hand, when given a probabilistic forecast, the user has the ability to choose a threshold value for his actions (act when the event is predicted with a probability of 70% for instance), and it is this ability that makes the forecast specific to each user's needs. As the threshold value varies, the relationship between RV and α is modified, and therefore decision makers can spot the threshold that would give them the optimum RV. For instance, it has been shown in the study that users with a value of $\alpha = 0.1$ (meaning important losses involved in case of occurrence) would benefit from a Relative Value of 0.4 if they act when the forecast probability is 10% or more but would not receive any useful information from it if they waited until the probability is 50% : Giving the same forecast to all users is not a reasonable option. Therefore, these two studies suggest that the use of probabilistic information in weather forecasting would not only

help the forecasters provide a realistic judgement of the situation, but also enable the users to optimise their decision making process.

The aim of this project is to forecast the probability of precipitation exceeding given threshold values, using NWP model's direct outputs. The way in which these probabilities should be estimated is still an open question, and the next chapter will focus on different recent methods designed for this purpose. The procedure as well as the data used for our different experiments will be presented in chapter 3, followed by results obtained during two different case studies: Chapter 4 will present a convective episode and chapter 5 a frontal system. Finally, the main conclusions regarding the project as well as the future possible lines of work will be discussed in chapter 6.

Chapter 2

Overview of different methods

In the previous chapter we exposed the difficulties inherent to Quantitative Precipitation Forecasting (QPF), and argued that probabilistic information offers some flexibility for both the forecaster and the user. Here we will review some methods which provide a probabilistic QPF without the help of a forecaster. Particular attention will be paid to the neighbourhood approach in section 2.3 since it will be the basis for the present project.

2.1 Ensemble forecasting

Due to recent progress in the field of high performance computing and the development of massively parallel machines, ensemble forecasting techniques have received particular interest in the recent years. The basis of such methods is to run multiple forecasts starting from slightly different initial conditions, called ensemble members. This approach offers a way to take into account different evolution scenarios and therefore to assess the uncertainty in the model forecast: If all runs converge to a similar result, the level of confidence for the prediction can be high, whereas if they tend to have very different behaviours it will be poorer. The spread of the ensemble members therefore represents a measure of the predictability. Since the atmospheric system has a tremendous number of degrees of freedom, the members can only partly represent all uncertainty possibilities and the choice of the perturbations in the initial condition is a key issue. Fields which are commonly perturbed are temperature, humidity and horizontal wind components, and

the actual perturbation procedure differs from one study to another. The experiment by Walser et al (2004) mentioned in introduction uses a lagged technique, in which the initialisation times of 6 different runs are separated by one hour. When all runs have been initialised, the deviation from the ensemble mean is amplified for the selected fields (temperature, wind, humidity) leading to different initial atmospheric states for the simulation. Other methods do not consider any time lag but focus more on the amplification procedure to realistically match observed analysis errors. A popular one introduced by Mullen and Baumhefner (1989) separates the large scale error from the small scale one in order to take into account the fact that the noise due to initial perturbation equals the signal for small scales (Du et al. 1997, Bright and Mullen 2002).

Although ensemble methods were originally designed to take into account the evolution of baroclinic perturbations in medium range forecasts (6 to 10 days), some recent studies have tackled the issue of mesoscale predictability of precipitation using similar approaches with limited area models (Walser et al 2004, Bright and Mullen 2002, Marsigli et al 2001, Du et al 1997). In this studies ensemble systems are a way to consider other growth mechanisms like those linked to moist physics and convection, as mentioned in section 1.1.

Du et al (1997) used a similar method to Mullen and Baumhefner (1989) to produce some probabilistic quantitative precipitation forecasts from a mesoscale model. They considered the 6 hours accumulated precipitation amounts and they divided the possible outputs into 5 categories ($p < 0.01$ inch ; 0.01 inch$< p < 0.10$ inch etc...). Then, for each grid point they computed a probability of being in each of the categories based on the population of ensemble members: They used 25 members so if there are 5 members in each category for instance, their probability would only be of 20% for each, meaning poor confidence. On the other hand if 20 of them are in the same category, it would be attributed a probability of 80% , expressing a high level of confidence in the forecast.

This ensemble prediction approach is probably the most commonly used to produce probabilistic outputs from NWP, however it requires an amount of computer power which is not always available and for this reason some alternative methods are sometimes preferred.

2.2 Statistical methods

Statistical methods refer to forecast techniques which draw a relationship between a set of variable to infer the value of one or a few others. The parameters used for the prediction are called predictors and the ones actually predicted the predictands. Applequist et al (2002) have applied and compared several different statistical models to predict the probability of precipitation exceeding a certain threshold after 24 hours (predictand). For each of the 5 models they used (linear regression, discriminant analysis, logistic regression, neural network and the classifier system), the procedure was to use a training data set to first select the best predictors and then to identify the coefficients or rules that would provide the best fit between predictors and predictand. The data used for the selection of the predictors was composed of synoptic and upper air analyses, as well as 6-hourly precipitation forecasts up to 24 hours. Altogether they considered more than 200 potential predictors including model variables as well as derived quantities like humidity or temperature advection, and binary variables for rain exceeding given thresholds. To rank and choose the predictors, the selection procedure uses the Brier Skill Score (BSS) which is based on the following Brier Score (BS) originally defined as twice this value by Brier (1950).

$$BS = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2 \tag{2.1}$$

Here, n represents the number of forecasts, y_k the forecast probability and o_k the actual observation ($o_k = 1$ when event observed and 0 otherwise). A good forecast method should therefore keep the BS score as low as possible, by having high probabilities when $o_k = 1$ and low ones otherwise. The BSS is then a measure of the forecast improvement over climatology, normalised by a hypothetical perfect forecast, so higher BSS corresponds to better forecasts:

$$BSS = \frac{BS_{forecast} - BS_{climatology}}{BS_{perfect} - BS_{climatology}} = 1 - \frac{BS_{forecast}}{BS_{climatology}} \tag{2.2}$$

Although the stepwise selection differs slightly from one model to another (fewer predictors were considered for the neural network and classifier system, and a different stopping rule for the linear regression), the algorithm is still the same: First choose a single predictor which provides the best BSS with the training data set, then look in the remaining ones for the combination that would best improve this BSS and so on until the BSS can not significantly be improved anymore. Depending on the statistical model, 2 to 6 predictors were chosen, and the ones selected with greatest frequency were the 24h accumulated precipitation forecast, binary variables for rain and relative humidity exceeding given thresholds and a layered averaged value of relative humidity after 12 hours.

2.3 The neighbourhood approach

The method which will be referred to as the neighbourhood approach in the rest of the study was introduced by Theis et al (2005) in an attempt to provide some probabilistic QPF with a low-budget procedure. In effect, all the methods we considered so far offer some heavy constraints in one way or another: The subjective post processing of Direct Model Outputs (DMO) into a probabilistic forecast requires the experience of a qualified team (see section 1.2) and the use of ensemble methods is only possible with great computer power. Statistical methods might not require as much power, but the amount of data set needed to build the model is not always available either. Consequently, the main objective of this procedure is to use the information already available from the model output to obtain the QPF.

To illustrate their method, they post processed outputs of precipitation accumulation obtained with the German DWD lokal modell, which has a horizontal grid spacing of 7km and a domain of 325×325 grid points. To begin with, a neighbourhood is computed around each grid point in the spatial domain, as shown in figure 2.1 for the X-Y plane, and then a similar procedure selects points from previous and successive model runs in order to take into account the time dimension as well. However, the way in which this time procedure is done is not clear from the paper: Although it is clearly stated in the text that the size of the neighbourhood is held fixed at all lead times of the simulation, a

diagram similar to figure 2.1 but in the X-T plane suggests that the diameter is actually decreased when moving away in the time dimension. This contradiction will be further discussed in chapter 3 when trying to reproduce the method.

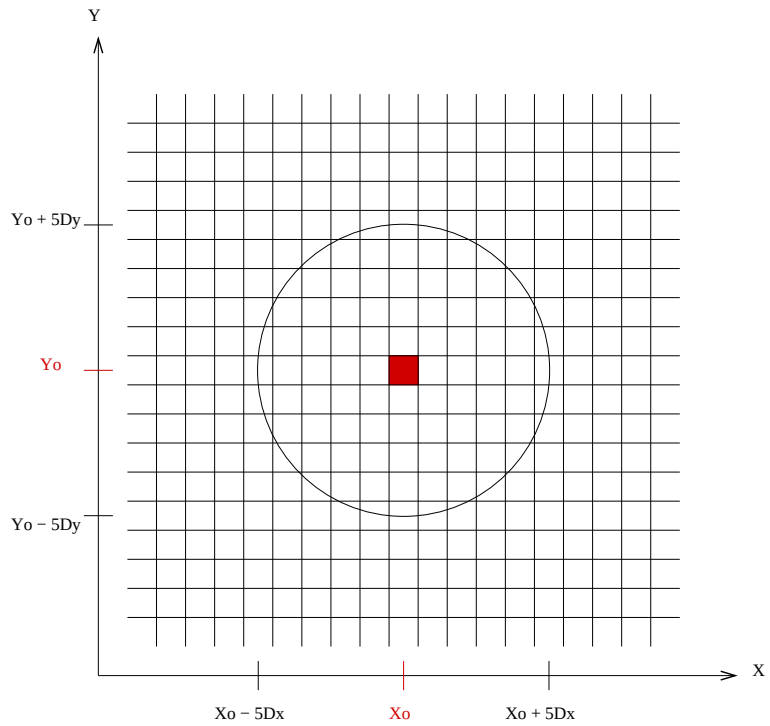

Figure 2.1: Example of a spatial neighbourhood with a diameter of 10 grid points (70km).

The next stage is then to go through all the grid points inside the neighbourhood (in both space and time) and count the ones for which the DMO exceeds a given threshold. The probability of exceedance is finally obtained by dividing this number by the total number of points inside the neighbourhood.

The verification procedure uses observed precipitation data from rain gauges to assess the skill of the method . The comparison is based on the Brier Skill Score (BSS) described in section 2.2 (see equations 2.1 and 2.2), but instead of having a climatological score as a reference in the BSS they chose to use the DMO score: This is obtained by transforming the deterministic output so that a probability of 100% is given when the forecast exceeds the selected threshold and 0% otherwise. In that way the BSS reflects the improvement of the method against the DMO. The influence of different parameters

such as the threshold value or the neighbourhood size were investigated, and all results lead to a positive value of the BSS (between 0.2 and 0.4), meaning that their method always outperforms the DMO. Perhaps the most interesting result of this study is the fact that of all neighbourhoods considered, the larger one (140km diameter) always leads to the best improvement. This suggests that some useful information on the predictability can still be found as far as 70 km away from the actual forecast location, and is to be related to the experiment by Walser et al (2004) which showed that predictability might be limited at scales up to 100km in convective situations.
The second part of the verification procedure uses the Relative Value (RV) (see equation 1.1) to represent the user's interest in the forecast. No matter what the cost/lost ratio is, the RV of the neighbourhood method is always higher than the DMO one. Furthermore, by post-processing the model output, the neighbourhood technique provides a widening of the curve representing the RV as a function of the cost/lost ratio: The post processed forecast would be beneficial to more users (see section 1.2).

In a study on the influence of model resolution on the skill of precipitation forecasts Roberts (2006) used a neighbourhood approach as a verification method. As shown in figure 2.2, he suggested that the range for the neighbourhood size should be bound. The lower limit representing the size for which the skill starts to be acceptable and the upper one the size at which the output has been smoothed out to much. In effect, when the neighbourhood size is increased too much, it leads to a smoothing of the forecast field which decreases the added value of a high resolution model. However, like the desired skill level, this upper level strongly depends on the user's needs: Flood forecasters for instance do benefit from a forecast over a wide area since they deal with river catchements.

In his discussion about the verification method N.Roberts also suggests that the averaging procedure giving the exceedance probability could perhaps be modified by using a Gaussian distribution of weights inside the neighbourhood, in order to give more importance to grid points close to the centre. Even though he believes that the method should stay as simple as possible and that the global improvement in the results would not justify the added complexity, some benefit could nonetheless be gained in the way

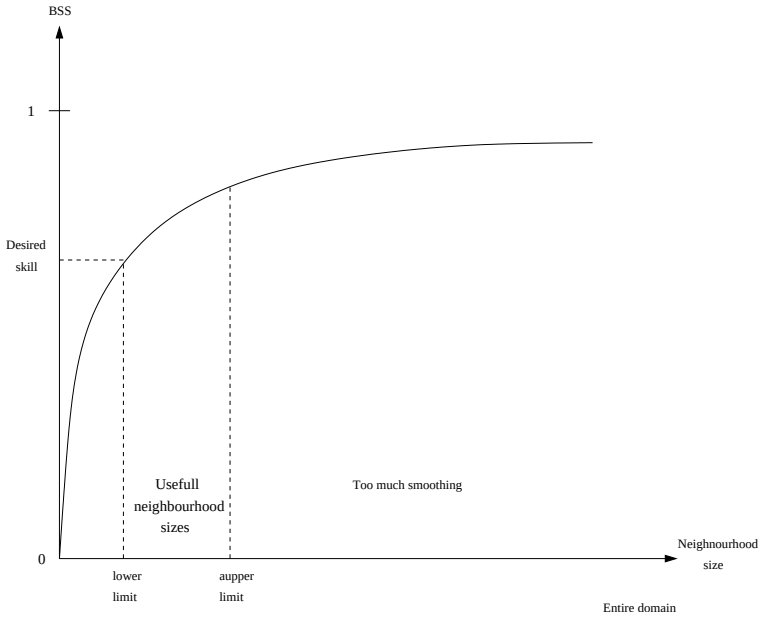

Figure 2.2: Skill of the forecast as a function of the neighbourhood size.

the probabilistic field is represented: Plots would show fewer discontinuities and a more circular shape around high probability areas.

After this quick overview of possible methods to obtain a probabilistic QPF, next chapter will now focus on the way to implement our own method, based on the neighbourhood approach by Theis et al (2005).

Chapter 3

Methodology

The aim of this chapter is to expose the procedure used to postprocess NWP outputs of rainfall rates into a probabilistic quantitative precipitation forecast. After a presentation of the model and verification data used, we will focus on the neighbourhood based method implemented as well as on different modification performed in the hope of improving the forecast skills.

3.1 Model and data

The NWP model used in the study is the non-hydrostatic version of the Met Office Unified Model (UM), with a grid length of 4km. The model domain includes the southern half of the UK as well as the coast of northern France in order to capture thunderstorms moving across the Channel. One particularity of this 4km grid model is that although the convective part of the rainfall is mainly explicitly resolved, a convection scheme is still included as a complement to the dynamic component. This version of the UM is still in an experimental stage, but has been running and archiving data for almost a year now. Corresponding to model runs at 03, 09, 15 and 21UTC every day, this archived data contains the analysis fields at the time of the run as well as forecasts with lead times of one and a half, three and four and a half hours ahead. In following sections of the report, this one and a half hour gap between these output times will be referred to as a time step (even though it does not correspond to the model's time step), and will be noted D_t. All model data used in case studies were obtained through the Joint Centre

for Mesoscale Meteorology (JCMM), thanks to Changgui Wang.

Composite radar data, easily available from the British Atmospheric Data Centre (BADC) were chosen to verify the method against observations. This data derived from composites of single radar sites is represented on a 5km grid, and therefore the direct comparison with model outputs is only possible after interpolating the data into the model's 4km grid. This part of the procedure was realized by Dr. R.Plant, with a linear interpolation between nearest points on the 5km radar grid.
Unlike rain gauges which represent the accumulation of rainfall over time, radar images give an instantaneous picture of the precipitation field and are a tool to measure its intensity distribution at a given time. Consequently, the selected forecast field has to represent similar characteristics, and the rainfall rate seems to be the most appropriate quantity: Both its large scale and convective components were summed to obtain the variable used in the post-processing procedure.

One particularity of this variable choice is that it might lead to some poorer skill for the DMO. In effect, when rainfall accumulation is used like in most of the studies referenced in chapter 2, a time averaging is implicitly involved and small delays in the model's forecast (mis-timing) do not penalise the calculated skill if they remain in the accumulation period. In the present study however, the model output is considered as an instantaneous image and compared to observed radar data, therefore similar mis-timings will be considered as model errors and the DMO skill will suffer from it. Figure 3.1 gives an example of a situation where the model correctly identified the overall pattern of the precipitation field but failed to represent the detailed structure accurately.

Because of this additional difficulty in forecasting rainfall rates, the use of a probabilistic post-processing method should lead to an even better improvement than when accumulations are considered.

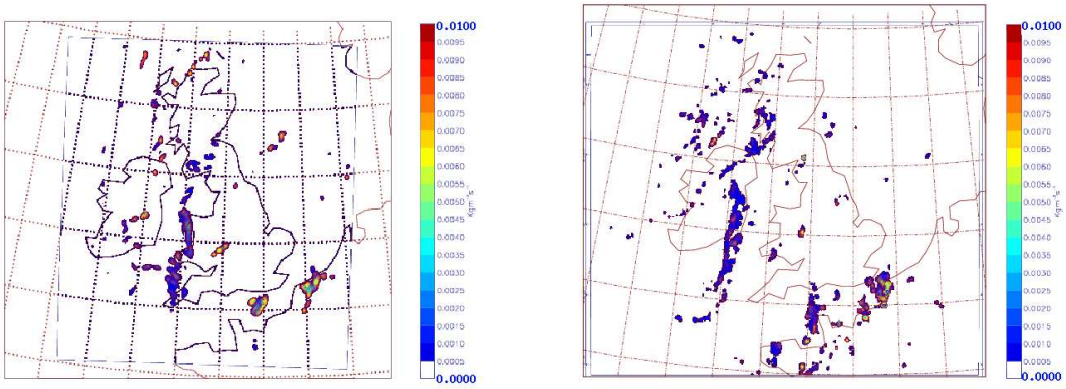

Figure 3.1: Direct Model Output of rainfall rates (left) and corresponding radar image (right).

3.2 Initial procedure

Since the basis of the method is the neighbourhood approach (Theis et al 2005) the obvious first step of the project was to implement and test it on the Unified Model's outputs. Details of the selected case studies will be given in chapters 4 and 5, and here we will focus on the technical aspect of the procedure.

As explained in section 2.3, the probability of exceeding threshold value Thr at grid point (x_0, y_0, t_0) is obtained by first building a neighbourhood around it in both space and time dimensions , and then dividing the number of points exceeding Thr by the total number of points inside the neighbourhood. The description of the algorithm computing it will be described step by step, starting with spatial considerations first and then extending the concept with a time neighbourhood.

3.2.1 Spatial neighbourhood

For every grid point (x_0, y_0) in the domain, a circular neighbourhood similar to figure 2.1 needs to be built. This is done by comparing the distance between (x_0, y_0) and its surrounding points, to the neighbourhood radius. The algorithm responsible for this part of the procedure is represented in a simplified way in table 3.1 , where α represents the

do $i_0 = \alpha + 1, N_x - \alpha$ do $j_0 = \alpha + 1, N_y - \alpha$	Loop over the whole domain except the border for which no probability is computed.
do $i = i_0 - \alpha, i_0 + \alpha$ do $j = j_0 - \alpha, j_0 + \alpha$	Loop around surrounding points to build the neighbourhood.
Compute distance between (i_0, j_0) and (i, j) If distance $< \alpha$ then (i, j) is in the neighbourhood Check if DMO(i,j) exceeds threshold	From Pythagoras theorem, distance is: $\sqrt{(i - i_0)^2 + (j - j_0)^2}$
end do end do compute probability for (i_0, j_0)	probability for (i_0, j_0) is: $\frac{\text{points exceeding threshold}}{\text{number of points in neighbourhood}}$
end do end do	

Table 3.1: Algorithm used to build a neighbourhood and compute the probability of exceedance with the initial method.

circle radius, N_x and N_y the domain size, Δ the grid length and DMO(i,j) the model rainfall rate for grid point $(x = i\Delta, y = j\Delta)$. Since no complete neighbourhood can be built for points at the edges of the domain, the probability was fixed to zero on a band of length α along the border. It is also important to notice that from now on all distances will be expressed in number of grid points, meaning that the corresponding length in kilometres is obtained by multiplying by $\Delta = 4km$. Similarly the notation (i,j) will be often used to designate grid point $(x, y) = (i\Delta, j\Delta)$.

3.2.2 In time

As discussed earlier in section 2.3, the way in which the time dimension is added to the spatial neighbourhood is not really clear from the paper by Theis et al (2005). The main contradiction comes from a diagram showing an elliptic shape in the X-T plane like the one on figure 3.2, which suggests that the spatial radius of the neighbourhood decreases as we move away from time t_0, whereas in the text it is actually stated that this radius is kept fixed at all lead times (corresponding to the blue rectangle on the X-T

plane). When considering the problem from a physical point of view, modifying the spatial size of the neighbourhood would make more sense if the radius increased with time instead of decreasing: Some information might be spread out into a wider area around point (x_0, y_0). Therefore the procedure will keep a fixed neighbourhood size at all times, as expressed by the blue rectangle on figure 3.2.

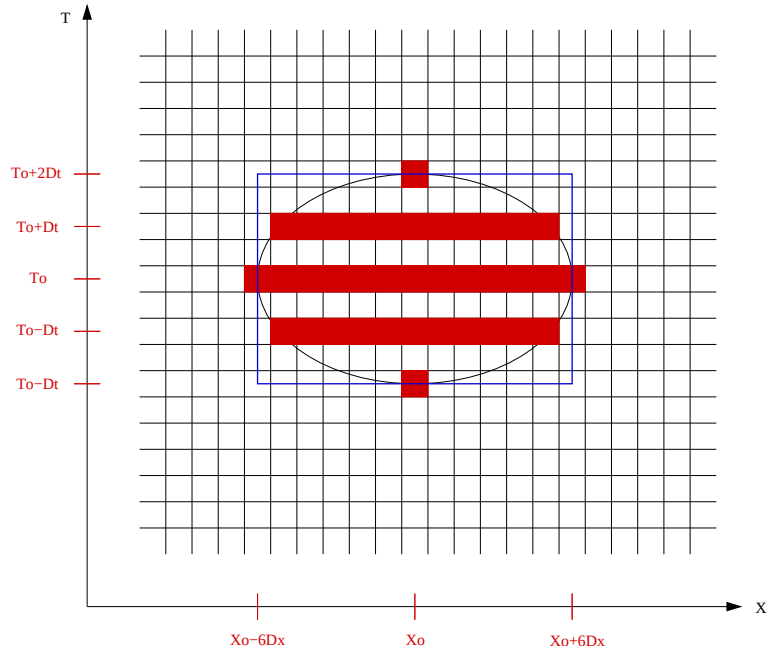

Figure 3.2: Shape of the neighbourhood in the X-T plane.

From a practical point of view this will require the model rainfall rate data to be stored in a three dimensional array DMO(i,j,n), where n represents the time index.
The procedure to check if the value exceeds the threshold will then become:

do $n = n_0 - \beta, n_0 + \beta$

check if DMO(i,j,n) exceeds threshold

end do

where β refers to the time radius and n is the time index. As before, the notation (i,j,n) will be often used to refer to point $(i\Delta, j\Delta, nD_t)$.

This initial version of the code, representing the original method by Theis et al, was tested on different case studies and gave satisfying results and skill scores (see chapters 4 and 5). It will therefore be used as a reference against which all following modifications will be compared.

3.3 Additional experiments

The aim of this section is to present two modifications to the method which were implemented in the hope of improving the value of the post-processed output. The original motivation comes from the idea that all points should not be given the same importance in the averaging procedure leading to the probability estimation, and the distances in both time and space should be taken into account.

3.3.1 Spatial weighting distribution

In the reference method, the averaging procedure is done considering that any point inside the neighbourhood accounts for an equal fraction of the probability. Whether the threshold is exceeded at a grid point directly adjacent to (x_0, y_0) or at a distance of α does not make any difference, whereas exceedance at a distance of $\alpha + 1$ is not taken into account. This characteristic is responsible for some discontinuities in the post-processed field, and as suggested by N.Roberts (2006), a Gaussian distribution of the weights could help improve this aspect of the procedure. Whether or not it will also improve the forecast skill is not clear, and the different studies in following chapters will try to investigate this question.

The Gaussian function is of the form $g(x) = exp(-\frac{x^2}{c})$, where the parameter c defines the sharpness of the curve. Figure 3.3 shows both the square function used in the reference method (in blue) and the Gaussian (in red) as a function of the distance to point (x_0, y_0). This example corresponds to a neighbourhood of radius 20 grid points and illustrates the way in which a Gaussian approach should improve the discontinuity problem: Points at distances α and $\alpha + 1$ from the centre are now attributed very similar

weights.

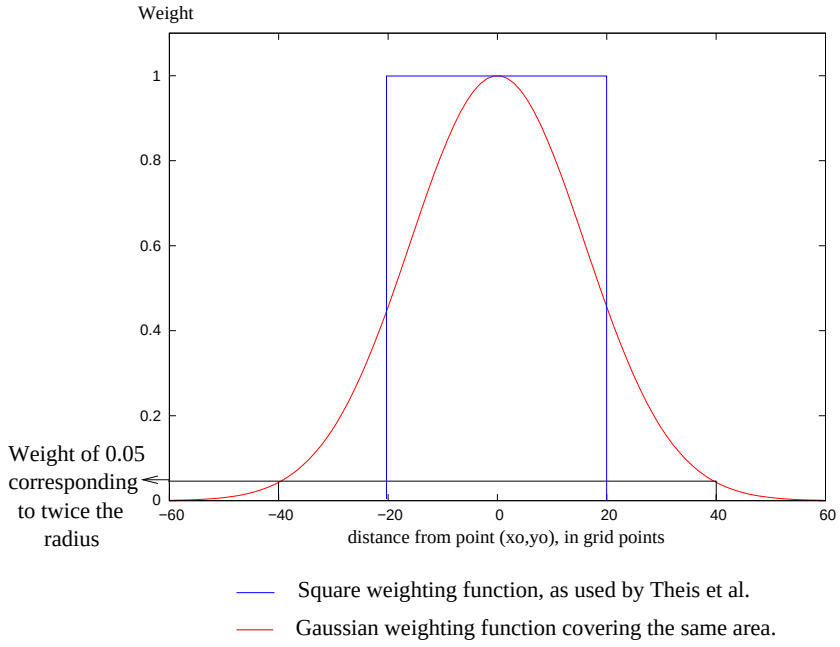

Figure 3.3: Example of Square and Gaussian weighting functions for a neighbourhood radius of $\alpha = 20$ grid points. The horizontal black line indicates the weight for a radius of 2α.

To stay consistent with the reference method and enable direct comparison, the coefficient c should be chosen carefully: Both function must represent the same total amount of weights, meaning that the areas below the curves should be equal. By expressing this later condition as follows, it is possible to link c to the radius α.

$$\int_{-\infty}^{+\infty} exp(-\frac{x^2}{c})dx = \int_{-\alpha}^{+\alpha} 1dx$$

Which, following a few lines of calculation can be written in the form:

$$c = 4\frac{\alpha^2}{\pi} \tag{3.1}$$

Even though this Gaussian distribution of weights in space does not correspond to a neighbourhood anymore, its parameterisation is still linked to the circle of radius α providing the same amount of weights. All comparisons of the two methods in chapters 4 and 5 will be based on this equation 3.1, and a Gaussian method of parameter c will be referred to through the corresponding α.

The algorithm for this modified method is slightly different from that in table 3.1 since all points surrounding (i_0, j_0) are now attributed a weight depending on their distance. In order to save some running time, points situated at a distance of 2α or more will not be dealt with (they represent a weight of less than 0.05 as shown by the black horizontal line on figure 3.3). The probability at grid point (i_0, j_0) is now obtained by:

$$prob(i_0, j_0) = \frac{\sum_{i,j} W(i,j) \times EXC(i,j)}{\sum_{i,j} W(i,j)} \tag{3.2}$$

Where $\sum_{i,j}$ corresponds to the sum over points at a distance smaller than 2α from (i_0, j_0), $W(i,j)$ is the Gaussian weight at grid point (i,j), and EXC(i,j) is a binary array representing the exceedance: EXC(i,j)= 1 when threshold is exceeded for point (i,j), 0 otherwise. An obvious downside of this Gaussian method is the extra running time involved (due to the computation of the weights and the extended spatial loop required), especially since the particularity of the neighbourhood approach was to be a low-cost procedure.

3.3.2 Time weighting distribution

Following on the idea that all grid points in the spatio-temporal domain might not be of equal importance in the estimation of a probability for point (x_0, y_0, t_0), a different approach to the distribution of weights in time should be tested. Let us consider a time radius of $\beta = 3$, like in the largest setting used by Theis et al. It seems reasonable to consider that an output field 3 time steps away from the actual time of the forecast t_0 contains less information than the forecast field itself, and therefore it should be given less weight. To do so, all points (x, y, t) have their weights reduced accordingly to the

distance between t and t_0. Figure 3.4 presents a weighting function linearly decreasing with time, which reduces the Gaussian weights by a factor R(t). A grid point (i,j) in the field at $t_o + 3D_t$ would for instance have its weight reduced from $W(i,j)$ to $R_3 W(i,j)$ (where R_3 is the notation for $R(t_o + 3D_t)$).

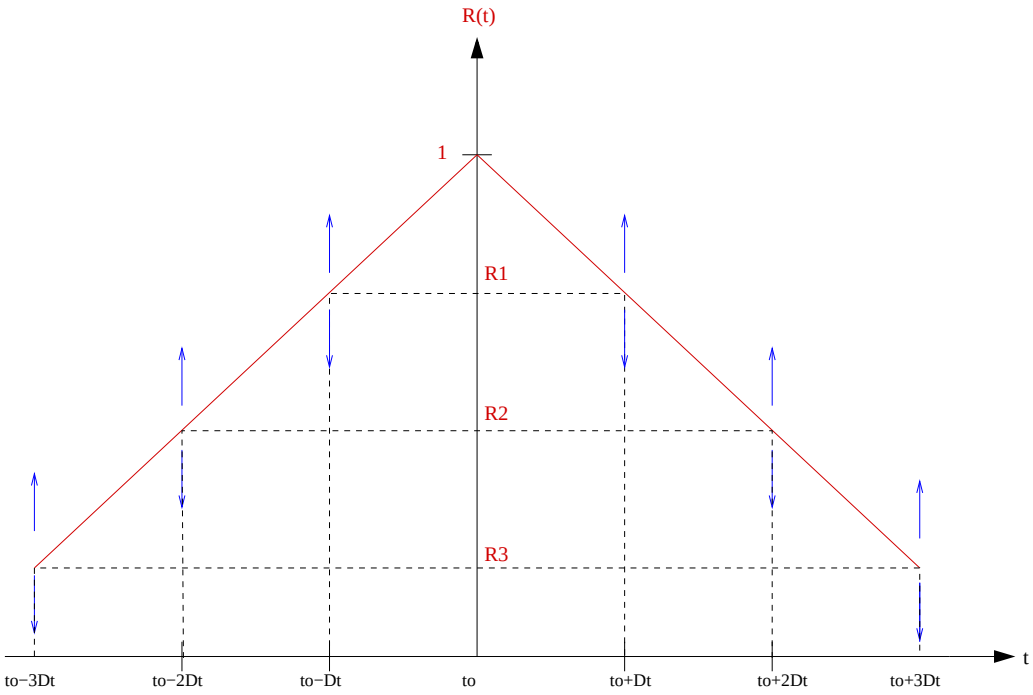

Figure 3.4: Reduction function using 3 time steps (in red). Blue arrows represent the different possibilities of parameterisations.

Therefore, the new probability will now be:

$$prob(i_0, j_0, n_0) = \frac{\sum_{i,j,n} R_n W(i,j,n) \times EXC(i,j)}{\sum_{i,j,n} R_n W(i,j,n)} \tag{3.3}$$

where the time index n ranges from $n_0 - \beta$ to $n_0 + \beta$.

Linear functions were chosen as a first guess to test the influence of this modification because they are easy to parametrise (just one parameter, the slope). As illustrated by the blue arrows on figure 3.4, there is of course a wide range of possible parameterisations for this reduction function and one of the focusses of following chapters will be to try and identify an optimum one.

The main objectives of this section were first to explain the theoretical basis of the procedure used in chapters 4 and 5 and secondly to quickly introduce the notations and algorithm used in the fortran program given in appendix to help its understanding. Several questions were raised about possible influences and benefits of the modifications:

- How will the choice of rainfall rates as a variable to post-process influence our results?
- Will the use of a Gaussian weighting function help solve the discontinuity problem, and if it does what will be the impact on the forecast skill score?
- What are the effects of the time averaging procedure on this skill score?

Case study of chapter 4 will now try to provide us with some first answers for a convective episode, and chapter 5 will extend these answers to a different situation involving a frontal system.

Chapter 4

Case study number 1: Convective situation.

This chapter presents the first case study on which the concepts described previously were tested and compared to the reference neighbourhood method. This comparison will be made step by step, starting from a single spatial field and progressively adding adjacent ones in time. The situation exposed was selected for its convective aspect, partly in order to test the potential of the Gaussian method to decrease the discontinuity issues noted in the method by Theis et al (2005).

4.1 Initial experiment: 18^{th} of May 2006 at 1800UTC.

As explained in section 3.1, the Unified Model's data are archived every day at 03, 09, 15 and 21UTC and are available through the Joint Centre for Mesoscale Meteorology (JCMM). This first experiment was built to represent a situation where a forecast at 1800UTC has to be issued from the model run at 1500UTC. Figure 4.1 shows the model's forecast at 1800UTC and the corresponding radar observations. This direct comparison of the two outputs shows that the model has correctly identified the two major regions of precipitation, namely the small band over Wales and Cornwall and the area of intense rain over northern France and Belgium. However, the size of the area covered as well as the amount of isolated convective cells were considerably underestimated. The other available forecasts at 1630UTC and 1930UTC are represented in figure 4.2 to show the

evolution of the situation.
Unlike following experiment which uses analysis field as part of the procedure (see section 4.2), this first case represents a real forecast situation where no observed information are used other than from the model run at 1500UTC.

4.1.1 Only space: First results

In order to test and validate the methods on a simple case, this first experiment only includes one model field, the one at forecast time t_0 =1800UTC. The fact that there is no time influence on the estimation of the probabilities makes the interpretation of the results easier and helps spotting possible errors in the code. Furthermore, this direct comparison of the two spatial weighting procedures offers an opportunity to tackle the issue about discontinuity in the post-processed field formulated in section 3.3.1, and to compare the first skill scores obtained.

Figure 4.3 presents outputs obtained with the Gaussian method (right hand side plots) and the square one (left hand side plots) for two different thresholds: Top plots show the probability of having rain rates over $10^{-4} kg.m^{-2}.s^{-1}$, corresponding to almost all precipitation in the area, and bottom ones over $10^{-3} kg.m^{-2}.s^{-1}$. The use of different thresholds helps quantify the intensity of the expected precipitation and these results show that although there is a high probability of having some rainfall in the western side of the domain, most intense event could occur in the south-eastern part.

One first obvious and encouraging result when looking at these plots is that the Gaussian approach has turned the contours from a rectangular to a circular shape, which gives the field a more continuous aspect, as we hoped. By looking more carefully at the areas of higher probabilities, it can also be noticed that the Gaussian outputs provide bigger values over more localised regions: The south-eastern corner of the bottom plots is a good example of such a situation since the reference method predicts a lower probability over an extended region. From a forecaster's point of view it can be argued that a more specific prediction is easier to use, and should therefore be prefered if the skills are similar.

Following on this idea, the Brier Skill Scores of both methods have been computed,

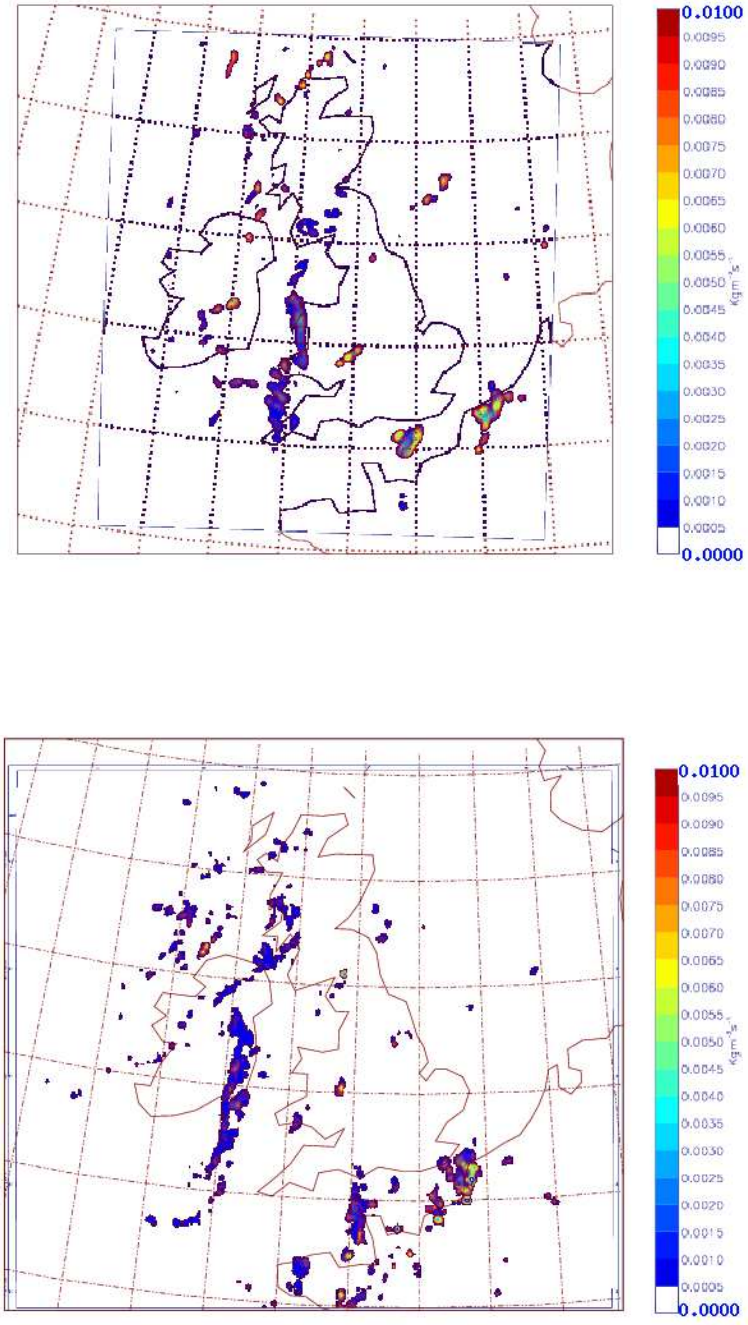

Figure 4.1: Direct Model Output of rainfall rates at 1800UTC on the 18^{th} of May (top) and corresponding radar observations (bottom). Rates are given in $kg.m^2.s^{-1}$.

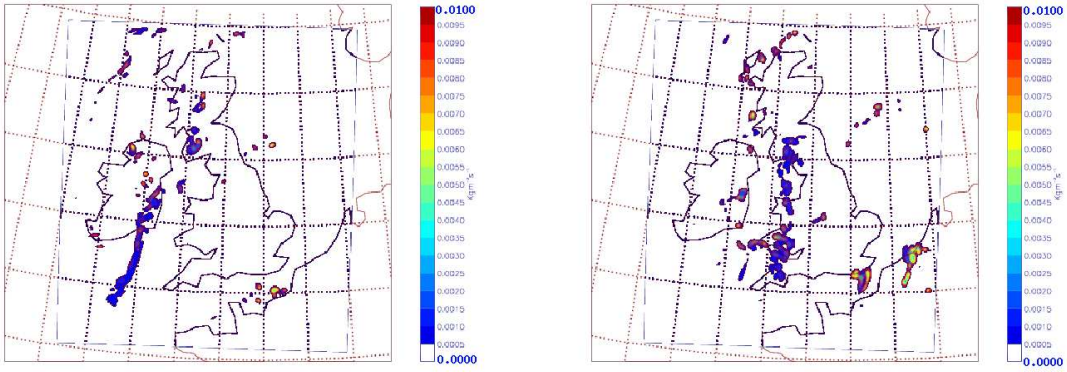

Figure 4.2: Direct Model Output of rainfall rates at 1630UTC (left) and 1930UTC (right) on the 18^{th} of May.

and the results are presented in figure 4.4. In order to be consistent with the study by Theis et al the very same BSS was implemented, using the DMO score as a reference instead of the climatological one (see equation 2.2 for details). As discussed earlier, this score provides a measure of the improvement obtained when using the post-processing method. It is expressed as a percentage of what the maximum improvement would be if a perfect forecast could be used.

A first interesting result is that the scores obtained range between 0.4 and 0.72 which is considerably higher than in the study by Theis et al. This significant improvement over the DMO can be explained by the choice of rainfall rates instead of accumulation as the variable to post-process (see section 3.1). The graph showing the evolution of the BSS as a function of the threshold value (left hand side) was realised with a neighbourhood radius of 15 grid points. It provides a way to test the model on two different types of competences: The first one is the ability to detect the areas of precipitation without considering the intensity values, and the second one the ability to correctly position more intense individual cells. When looking at small thresholds, almost all the precipitation in the domain is considered, and therefore the focus is on the identification of the precipitation area. When higher thresholds are selected, most light rain areas are removed and the difficulty becomes to accurately position the remaining localised cells. The BSS increases with the threshold value, meaning that when very low thresholds are

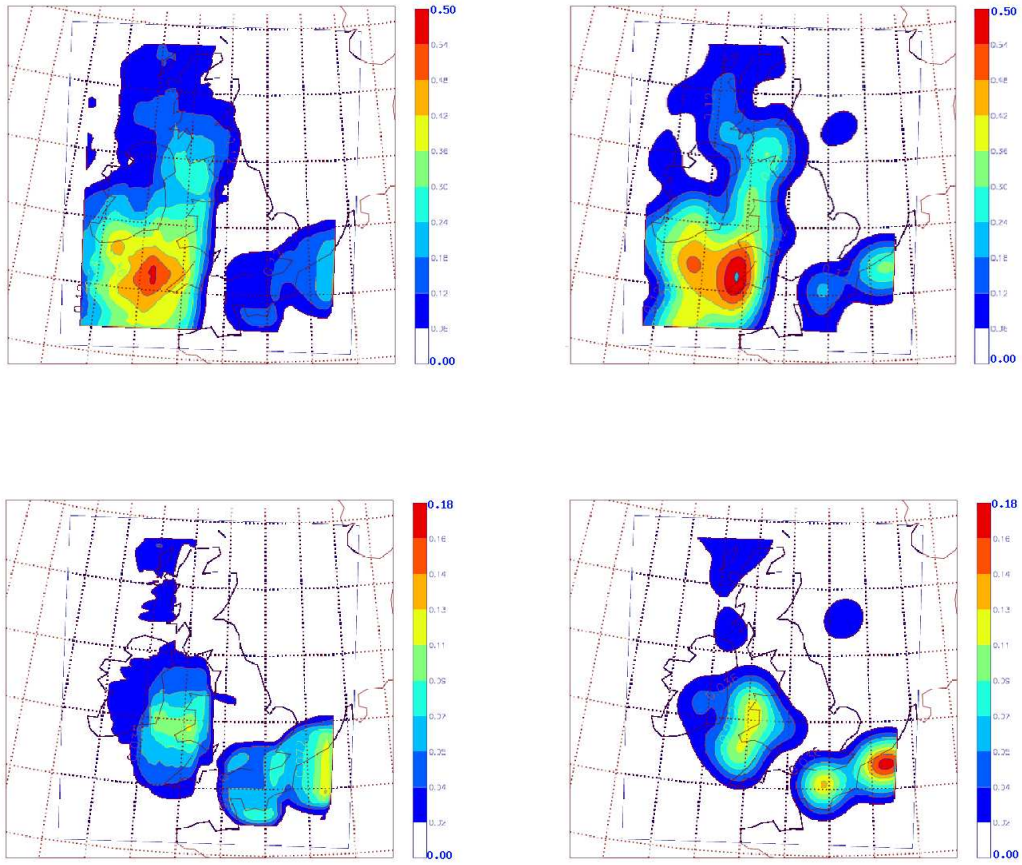

Figure 4.3: Exceedance probability with a threshold of $10^{-4}kg.m^{-2}.s^{-1}$ (top) and $10^{-3}kg.m^{-2}.s^{-1}$ (bottom) obtained with a neighbourhood radius of 20 grid points. Fields on the right hand side were obtained using a Gaussian weighting function, and the others with a square one.

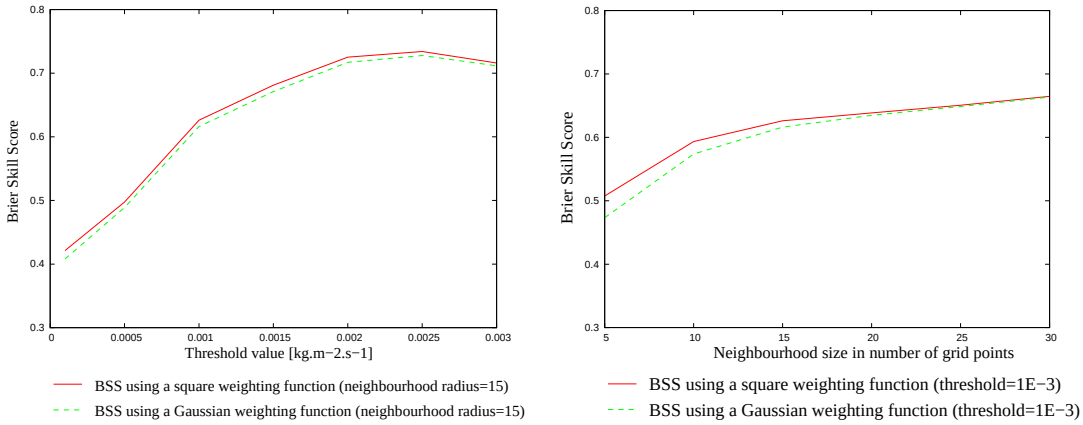

Figure 4.4: Brier Skill scores of the two methods (Gaussian in green and Square in red) as a function of threshold value (left) and neighbourhood size (right).

considered ($10^{-4}kg.m^{-2}.s^{-1}$ or $5.10^{-3}kg.m^{-2}.s^{-1}$) the model does recognise the overall pattern of precipitation correctly. Once these light rain areas are removed due to higher thresholds, the misplacement errors start to be identified and the improvement provided by the smoothing is getting higher.

Right hand side of the figure shows the evolution of the BSS with the neighbourhood radius α for a threshold value of $10^{-3}kg.m^{-2}.s^{-1}$. As expressed by N.Roberts (2006), the BSS keeps on increasing with α, but the slope is bigger for α below 15 than after. The outputs corresponding to these different sizes are shown in figure 4.5 where the threshold was fixed to $10^{-3}kg.m^{-2}.s^{-1}$ but values of α vary between 10 and 25 grid points (40 to 100kms). The smoothing phenomenon discussed by N.Roberts is clearly noticeable, with high probability values falling from 0.4 to 0.15 as α increases. We already argued with figure 2.2 of section 2.3 that the optimum value of α depends on the user's needs, and since here most of the improvement in the BSS is obtained before $\alpha = 15$ (figure 4.4), this neighbourhood size of 15 will be selected as a standard for the following experiments. Furthermore, this corresponds to a diameter of 120kms, which is in the same range of sizes that what was used by Theis et al.

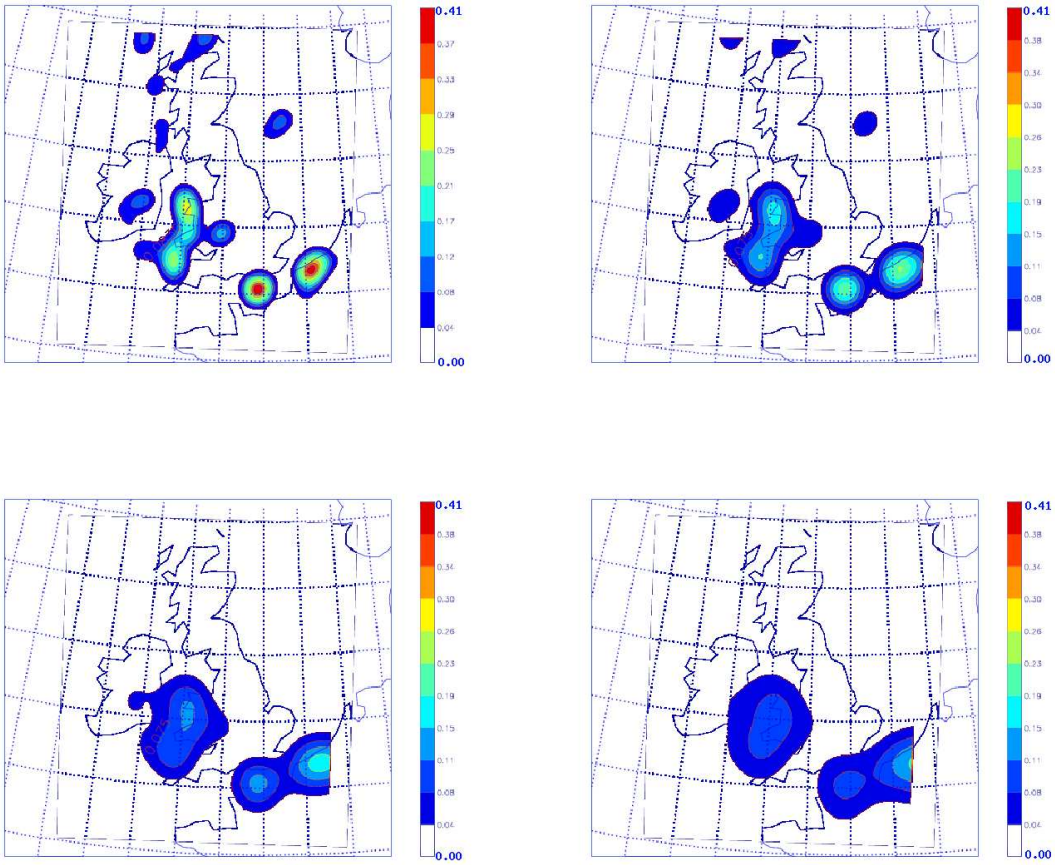

Figure 4.5: Probability of exceeding $10^{-3} kg.m^{-2}.s^{-1}$ obtained using the Gaussian method with different neighbourhood sizes: Radius is 10 gird points for top left plot, 15 for top right, 20 for bottom left and 25 for bottom right.

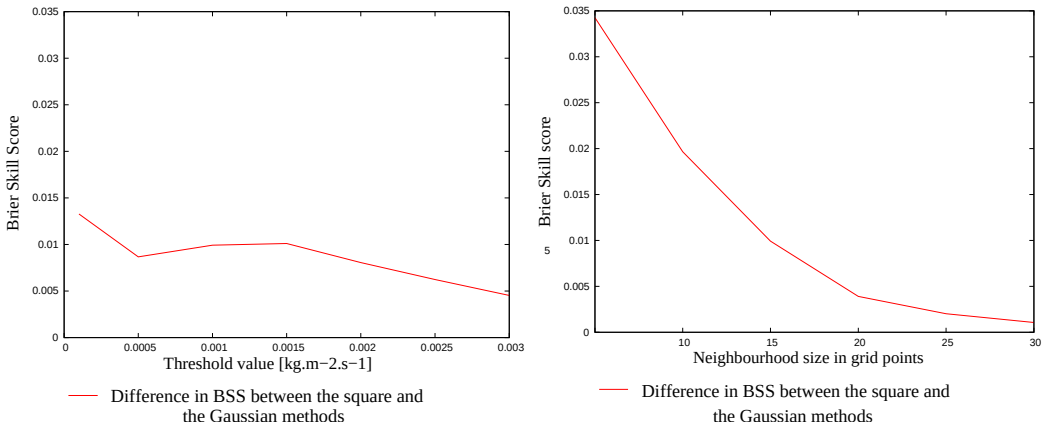

Figure 4.6: Difference in BSS between the two method (Square-Gaussian) as a function of the threshold value (left) and the neighbourhood radius (right). Left hand plot was obtained with a fixed radius of 15 grid points and right hand one with a threshold of $10^{-3} kg.m^{-2}.s^{-1}$.

The final aspect to discuss in figure 4.4 is the relative behaviour of the two curves: The reference method always outperforms the Gaussian one, but the difference does not seem significant. To further investigate this point the difference between the two BSS was plotted on figure 4.6 as a function of both threshold (left) and neighbourhood radius (right). The left hand plot, obtained for $\alpha = 15$, shows a rather steady difference of about 0.01, meaning that the reference method increases the BSS one percent more than the Gaussian method without real influence of the threshold value selected. The right hand plot however, shows an exponential decrease of this difference with α and to pass below this 1% over performance of the reference method α needs to be of 15 grid points at least.

This first experiment involving only the spatial dimension has provided us with some interesting information about the two methods: If the Gaussian approach helps solving the continuity issue and offers a more subjectively useful post-processed field, it scores slightly lower than the reference one for all thresholds and all neighbourhood sizes tested. However, providing a correct choice of parameters, this under performance is only around 1% of the BSS. This latter remark highlights the second important result of this section: Parameters are to be chosen carefully. The neighbourhood size is crucial since it influences the BSS and the level of smoothing in the post-processed field, and the threshold value determines the type of event we are focussing on (localised intense rain or global pattern

of precipitation). The following experiments will be performed using a Gaussian spatial weighting function with α =15 and a threshold of $10^{-3} kg.m^{-2}.s^{-1}$. This setting offers a good compromise between the aspect of the output field (more continuous and with limited smoothing) and the BSS value.

4.1.2 Space and Time

The objective of this second step of the experiment is to test the influence of the time dimension on the probability estimation, and the corresponding skill score. To do so, model outputs at 1630UTC (t_0-1) and 1930UTC (t_0+1) were added to the procedure as explained in section 3.2.2 (no reduction function). The addition of a time parameter leads to some further smoothing in the post-processed field, as shown by figure 4.7 where both the outputs with (left) and without considering the time dimension (right) are presented. If we now look at the effect on the BSS of this additional time dimension, it is noticeable from figure 4.8 that once again the smoothing is associated with an increase in the skill scores. The improvement is characterised by the gap between the green curve showing the BSS for the Gaussian method with time dimension involved and the red one showing similar BSS without time averaging. While this improvement seems to be roughly independent of the selected threshold (right hand plot), it does vary with the neighbourhood size (left hand plot): At small α the increase is considerably higher (up to 8% increase for $\alpha = 5$). This can explained by the fact that at large values of α much smoothing has been done by the spatial averaging and therefore the additional time averaging does not raise the BSS as much as it does for small values.

After these first few experiments on the influence of the different parameters involved in the procedure one main conclusion seems to arise: The increase in the Brier Skill Score obtained is often linked to a smoothing of the post-processed field. The level of smoothing acceptable depends on each user's need, but it is probably the main parameter to take into account when using the method.

Finally, this first case study offered an opportunity to test the effect of the reduction function mentioned in section 3.3.2 in a very simple way: Since we only have two extra time fields available (t_0-1 and t_0+1) the reduction is given by the parameter R_1 of figure

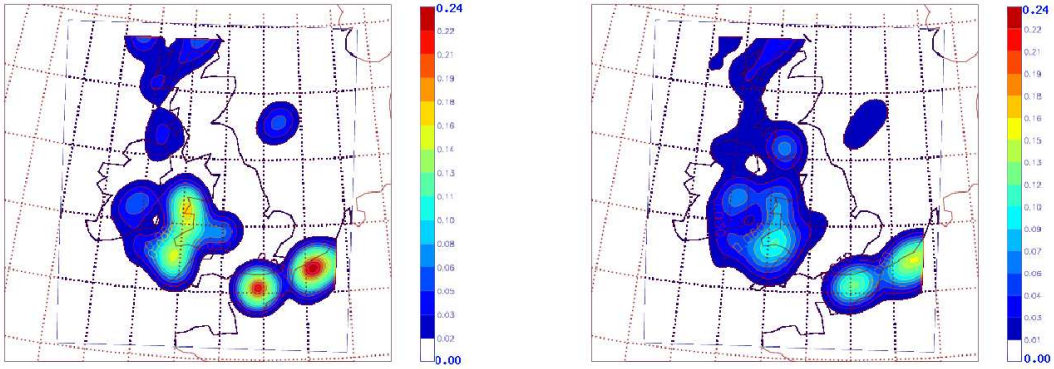

Figure 4.7: Exceedance probability for a threshold of $10^{-3}kg.m^{-2}.s^{-1}$ with α=15. Left hand plot was obtained with just the spatial dimension and right hand one with both time and space.

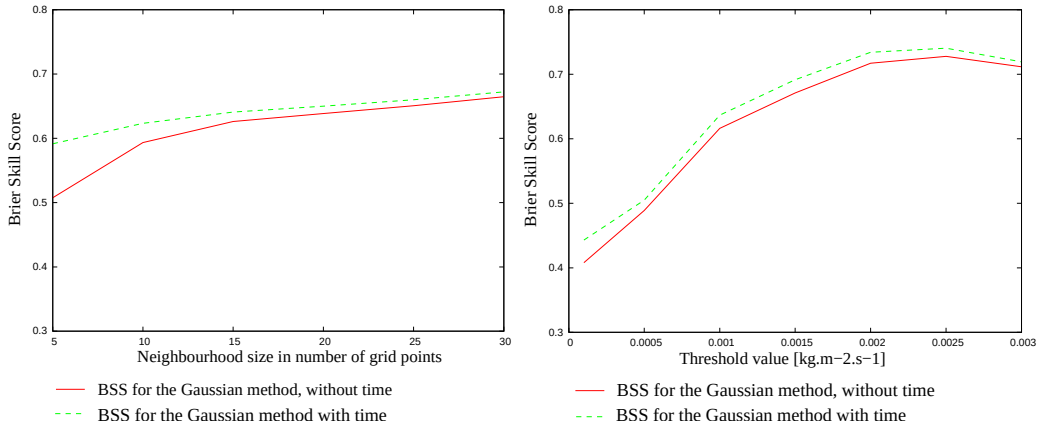

Figure 4.8: BSS for the Gaussian function with (in green) and without time dimension (in red) as a function of α (left) and threshold value (right).

3.4. The plot on figure 4.9 shows the evolution of the BSS as a function of this reduction parameter, for the standard settings discussed earlier. The lower limit of the x-axis, representing R_1=0 corresponds to the procedure without time averaging and the upper one (R_1=1) to the procedure including adjacent outputs fully. Since the BSS increases and reaches its best value for R_1=1 (2% improvement), there is no benefit from this reduction function.

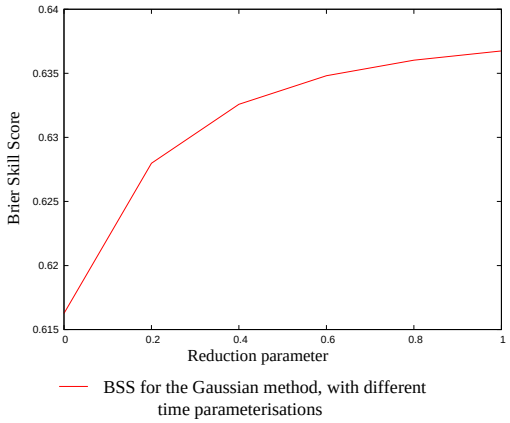

Figure 4.9: BSS with different time reductions, for α=15 and a threshold of $10^{-3} kg.m^{-2}.s^{-1}$.

The setting of this first test case does not enable further investigation on the influence of a second time step in the procedure since time t_0+2 corresponds to the analysis file at 2100UTC, which contains some observations. Using this field would not make much sense in the context of a real forecast situation and therefore this first case was slightly modified into a second case study to provide some further focus on the time dependence of the BSS.

4.2 Extended case: 18^{th} of May 2006 at 1500UTC.

This second case is in fact only an extension of the first one since it concerns the same weather situation. The main difference is that we will now post-process the analysis field at t_0=1500UTC. The motivation for this choice is to be able to use up to 6 output fields adjacent in time (3 after t_0 and 3 before) without including an analysis field in one of these forecast files. Figure 4.10 shows the content of the 2 files used and helps understand the set up of this experiment. This approach is slightly unusual since it represents a situation in which the forecast is issued from the analysis field, meaning that some observation have been included in the process. However, if the post-processing procedure improves the skills obtained with the analysis field it can be assumed that it would also improve the ones obtained with a forecast file (and probably by a larger amount).

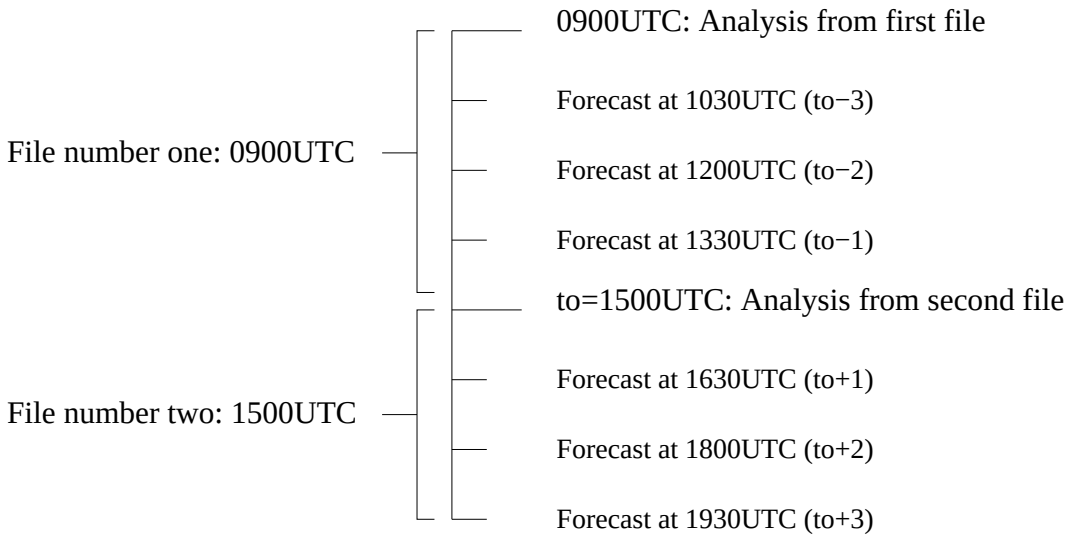

Figure 4.10: Composition of the 2 files used for the extended experiment.

The analysis field before post-processing is shown in figure 4.11 with the radar observations at that same time, and the complete sequence of files used in this section are shown in figure 4.12 on following page.

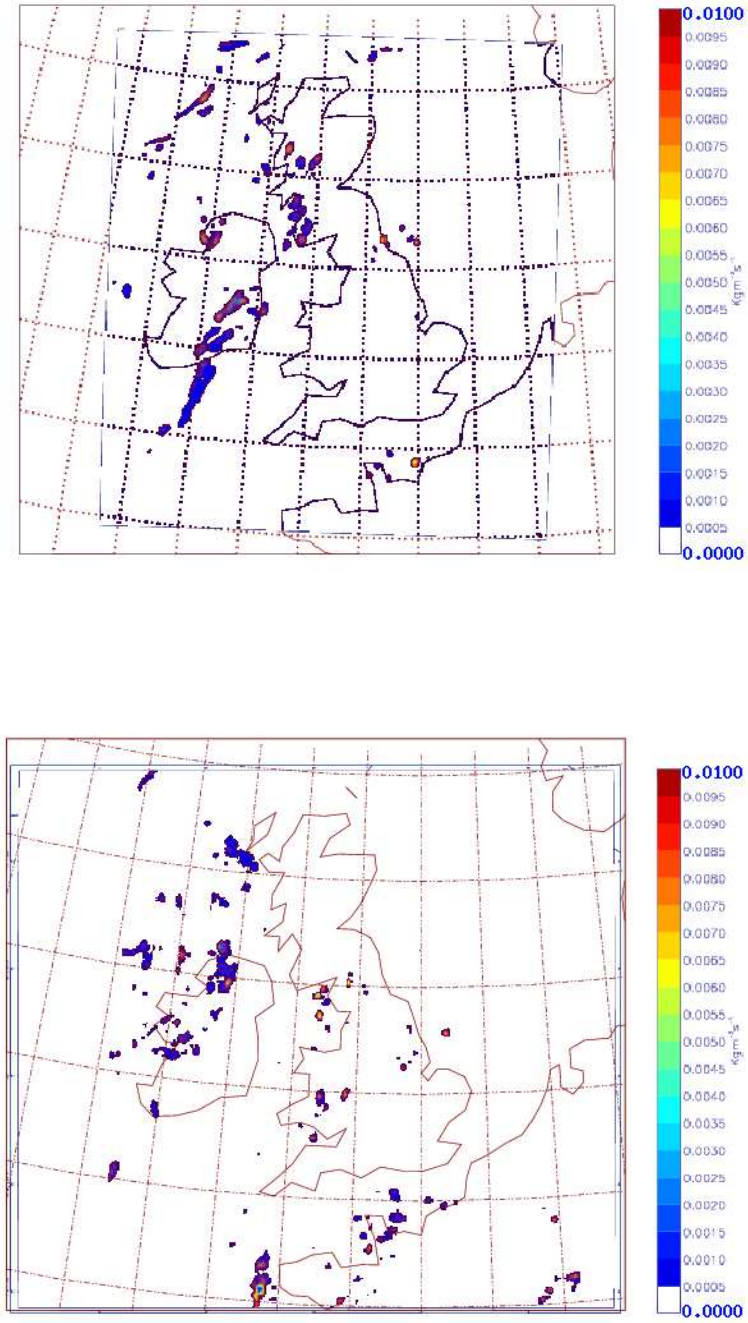

Figure 4.11: Output from Analysis file at 1500UTC (top) and corresponding radar observation (bottom).

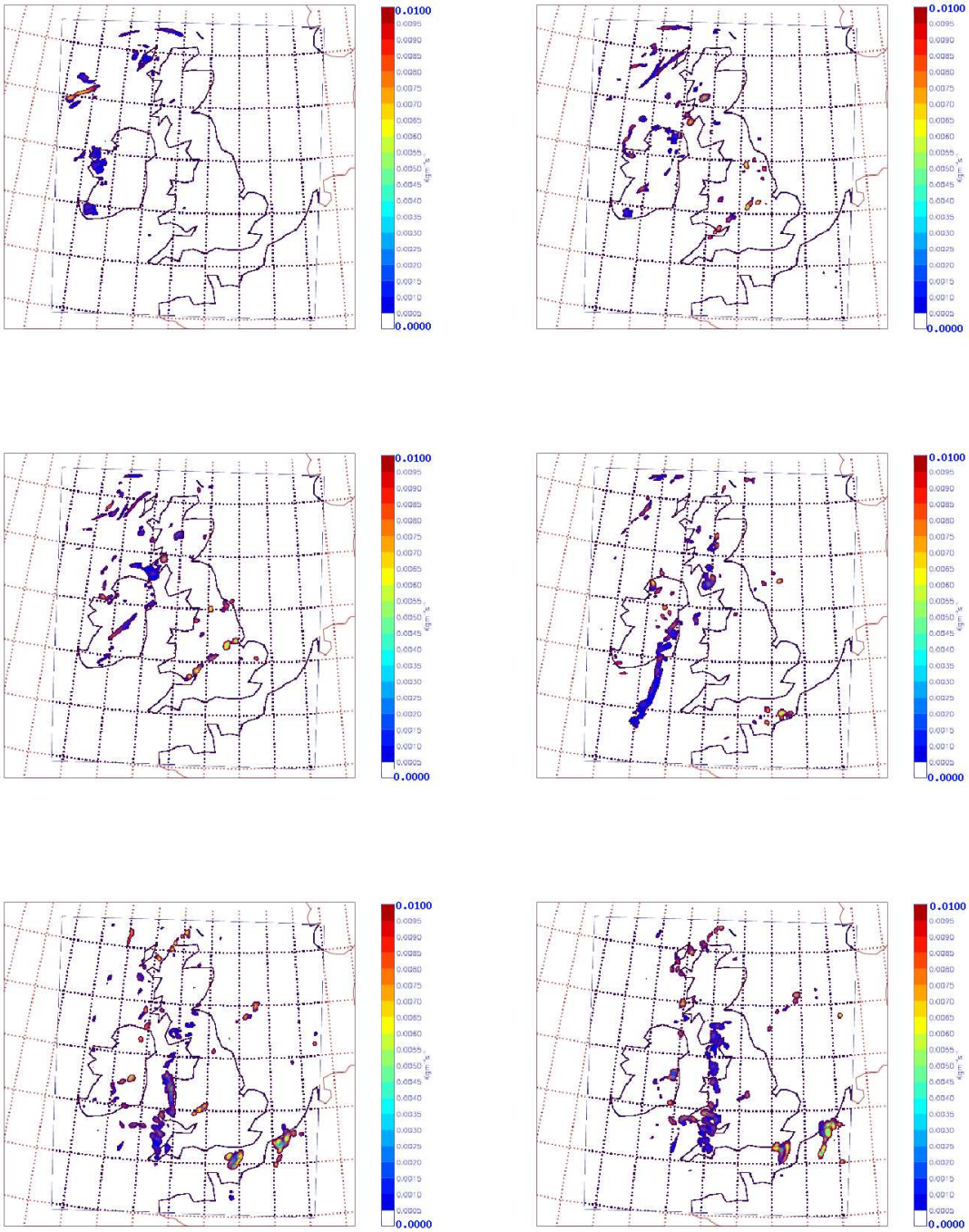

Figure 4.12: Output fields from the different files considered in the experiment, from 1030UTC (top left) to 1930UTC (bottom right).

To test the effect of the time range on the BSS, both the Gaussian and the square methods were ran using successively 0,1,2 and 3 time steps with no reduction at all. The results are presented in figure 4.13, and it is surprising to see that the use of a second and a third time step does not contribute to raise the BSS further: If the fields immediately after and before t_0 still provide some useful information (mainly on mis-timing errors), files further away are too different to be used. Another interesting result is that even though analysis data were used, the Brier Skill Scores obtained are still initially as high as in case number 1 (around 0.6) but the main difference is found on the improvement after one time step: Only 0.6% for the Gaussian here compared to 2% with the same settings in previous experiment (see figure 4.9).

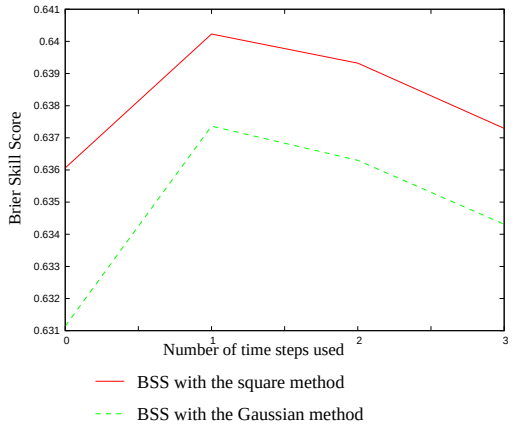

Figure 4.13: BSS for the Gaussian method (green) and the square one (red) as a function of the number of time steps used.

This situation where the BSS starts decreasing with the use of a second time step in the procedure is an interesting case to test the influence of the reduction function with different parameters. In order to spot an optimum parameterisation, a similar method to the one used by Applequist et al (2002) in their search of best predictands was used (see section 2.2): First run was made with only one time step but different values of R_1 (similar to figure 4.9) in order to select the best parameter. Then a second run used two time steps to try to identify the parameter R_2 which would give the best BSS when associated with R_1, and finally the same method was used with a third run to get R_3. The evolution of the BSS during these three runs is presented by figure 4.14, and the optimum parameter are R_1=1, R_2=0.2, and R_3=0. This parameterisation only improves the

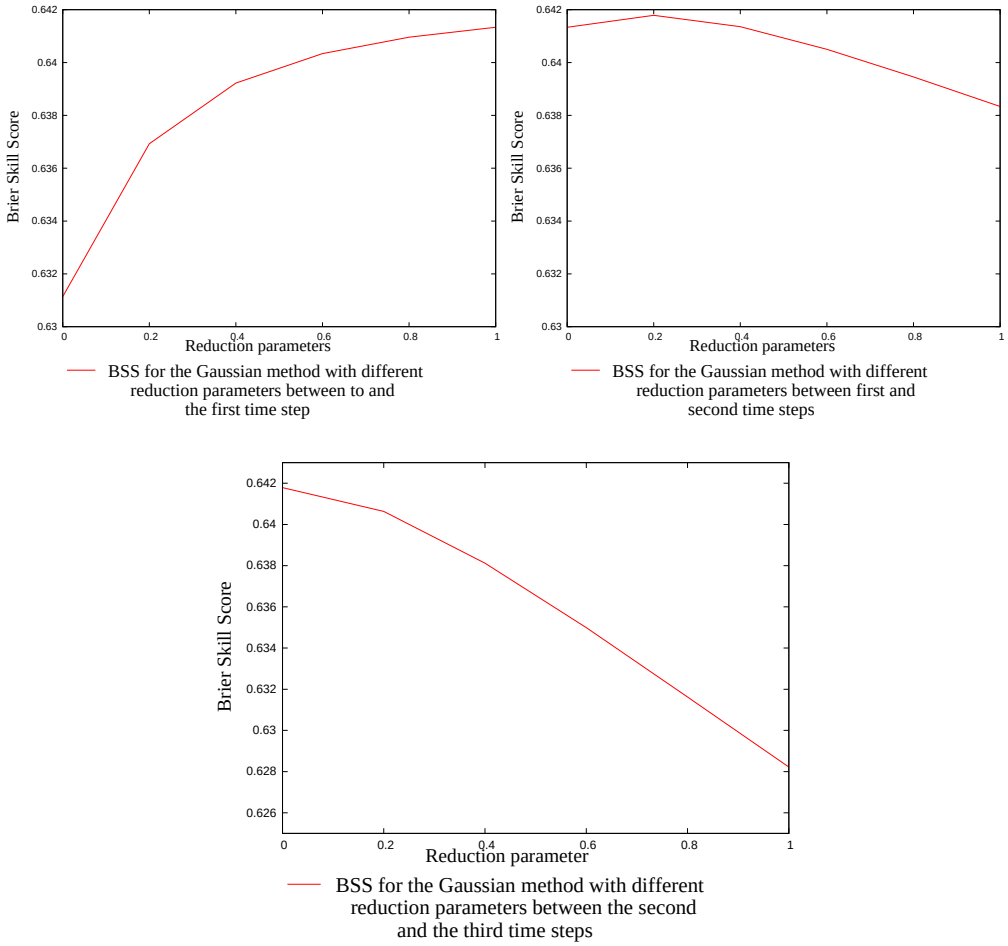

Figure 4.14: Evolution of the BSS for the Gaussian method with different reduction parameters. Top left plot shows the selection of parameter R_1, top right one the selection of parameter R_2 and bottom plot the selection of R_3.

normal case with one time step by less than 0.1% and therefore can not be considered as useful here, especially when the additional cost of finding the optimum parameters is considered. However the decrease in BSS following the use of a second time step was turned into an increase just by modifying the weights attributed to the fields at t_0+2 and t_0-2, showing that the reduction function could perhaps be useful in some different cases.

This extended case illustrates the fact that the way to handle the time dimension is not straight forward: If the addition of one time parameter has proved to be beneficial for the BSS, the second and third ones only decreased the performance. Furthermore, the identification of an optimum parameterisation for the reduction function has shown that even though the reduction method is not worth using here, the idea of decreasing the weights accordingly to the distance in time could help improve the forecast.

Chapter 5

Case study number 2: Frontal system

The purpose of this new case study is to repeat the whole analysis with a frontal system, in order to extend our conclusions to a wider range of weather conditions. As previously, a first case corresponding to a real forecast situation will be presented and then an extended version of it using the analysis will provide some additional data to increase the time radius in the procedure.

5.1 Initial experiment: 07^{th} of May 2006 at 0600UTC.

The forecast field which will be post-processed is presented in figure 5.1 with the corresponding radar observation at 0600UTC, and the outputs from the two other forecast files used in this section (0430UTC and 0730UTC) are given in figure 5.2. This stationary front laying on the east coast of the UK presents a totally different pattern of precipitation than the previous case, with weaker values of rain rates but a higher density of cells in the area covered. The comparison shows that the model has accurately positioned the front, but underestimated its extent to the northern part of the domain. Values of rainfall rates are similar in the major part of the area covered, but the regions of more intense precipitation (in yellow) are slightly misplaced.

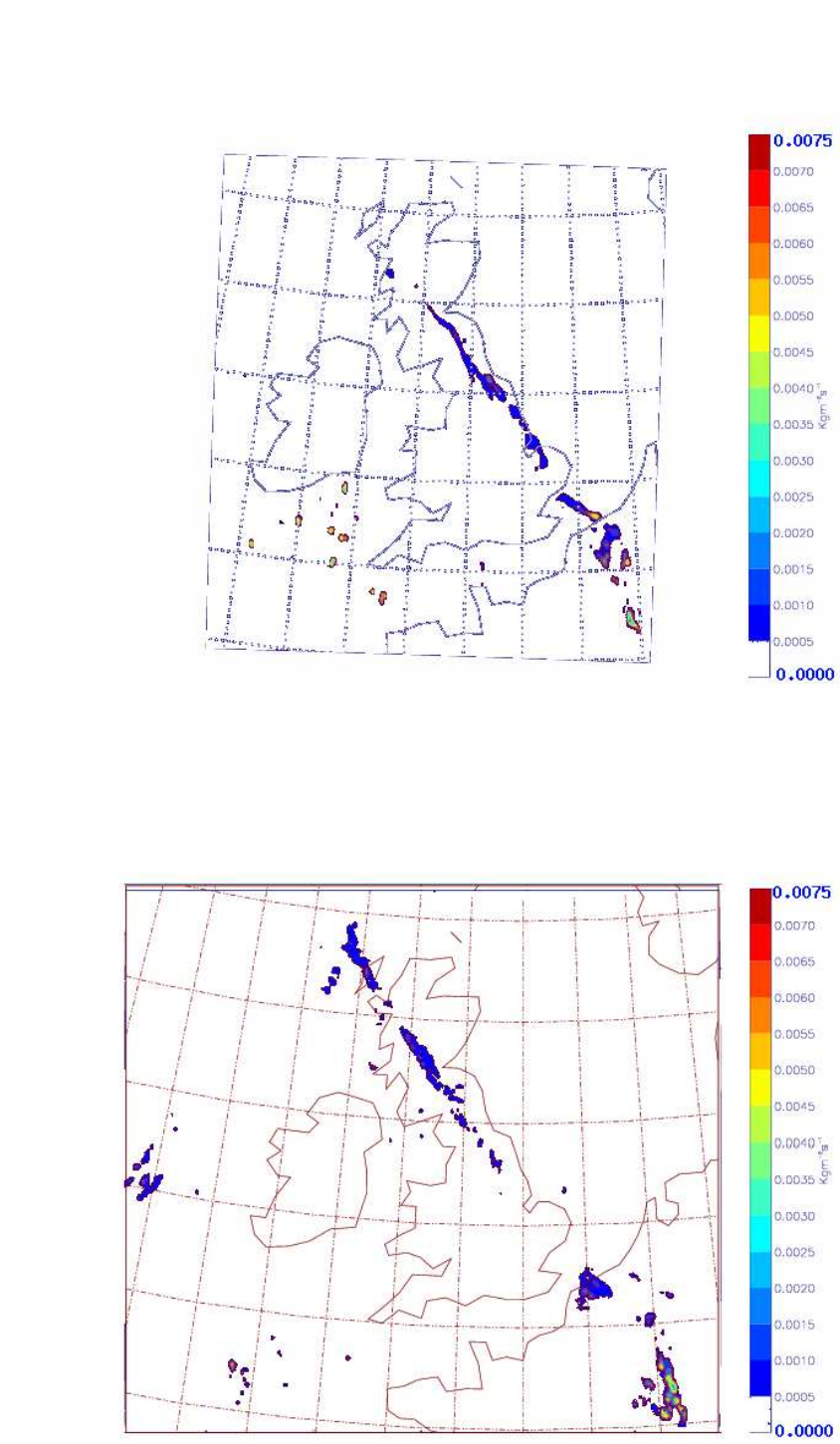

Figure 5.1: Forecast field of rainfall rate (top) and corresponding radar observation (bottom) on the 07^{th} of May 2006 at 0600UTC.

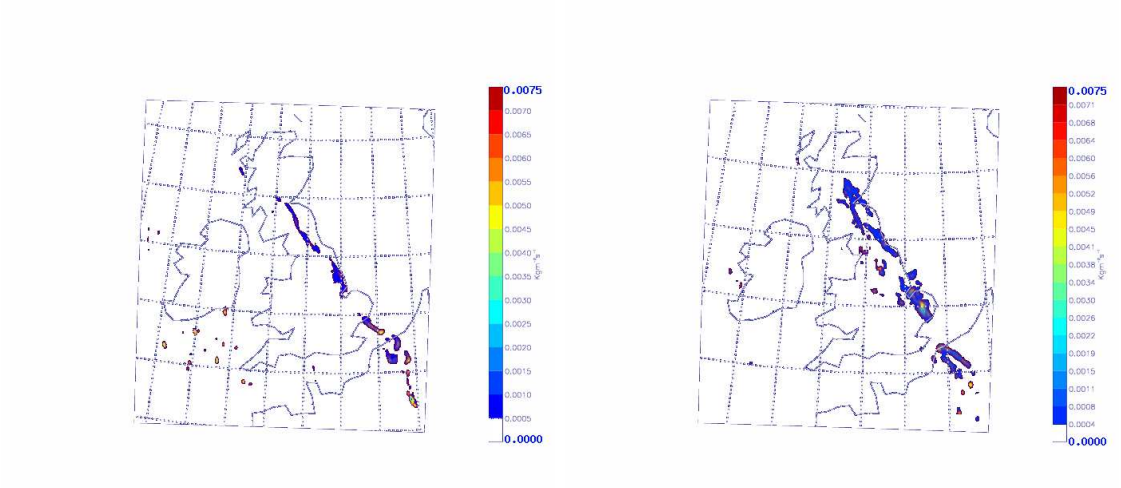

Figure 5.2: Forecast fields of rainfall rates at 0430UTC (left) and 0730UTC (right).

First post-processed fields obtained with the spatial dimension only, are presented in figure 5.3 for both the Gaussian (right) and the square methods (left). The neighbourhood radius is still fixed to 15 grid points as a standard value and plots show results for threshold values of $10^{-4}kg.m^{-2}.s^{-1}$ (top) and $10^{-3}kg.m^{-2}.s^{-1}$ (bottom). As in case number one, the square method tends to decrease the highest probability values, but the discontinuity problem is not as important anymore since the overall pattern of the front does offer some lines of abrupt changes: The advantage of the Gaussian output for a forecaster is less obvious here than for a convective situation.

The evolution of the Brier Skill Score of the two methods with the threshold value and the neighbourhood size is presented by figure 5.4 (top). The actual values of the BSS are very similar to case number one, and the shape of the curve only presents small differences: The plot of the BSS as a function of the neighbourhood radius for instance, does not converge to a maximum value at high radius like it did previously, and there is almost no increase between α =10 and α =15. Once again this highlights the idea that parameters should be chosen accordingly to each user's needs: If the user can deal with a high level of smoothing in the post-processed field, large values of α should be selected since they provide some additional increase in the BSS. On the other hand, if the smoothing offers some difficulties in the subjective interpretation of the forecast, a value of α =10 is probably the best choice since an increase to 15 grid points does not

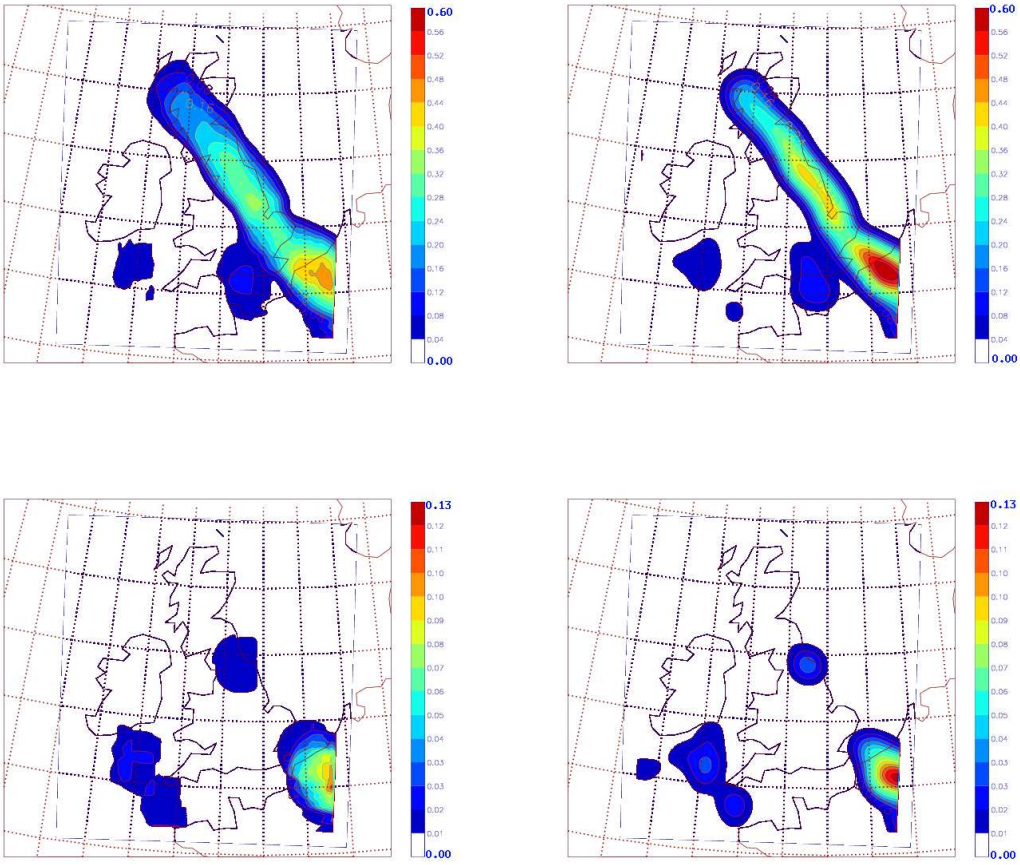

Figure 5.3: Probability of exceeding a rainfall rate of $10^{-4}kg.m^{-2}.s^{-1}$ (top) and $10^{-3}kg.m^{-2}.s^{-1}$ (bottom). Left hand plots were obtained using the square function and right hand ones the Gaussian function.

improve the BSS considerably.

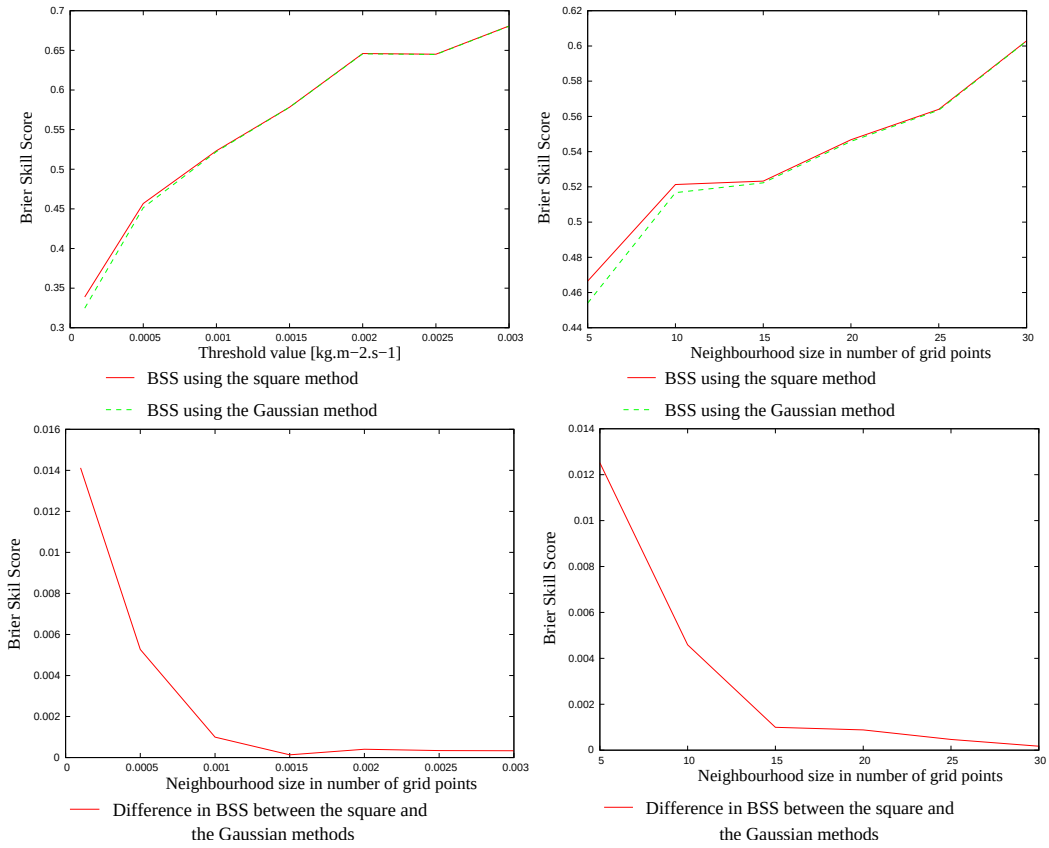

Figure 5.4: BSS and error as a function of the threshold value (left) and the neighbourhood size (right).

The plots of the difference in BSS between the two methods (bottom) both show an exponential decrease with values of less than 1% almost immediately reached, meaning that the methods score similarly in the range of values used for our experiments.

This first part of the case has not lead to any major differences in the interpretation of the Brier Skill Score, nor in the relative behaviour of the two methods tested. However, the benefit of using a Gaussian method is not obvious since the discontinuities in the post-processed field are not a real issue anymore. The extra cost (in running time) involved when using the Gaussian approach is not justified by any significant improvement in this study and the normal square method should therefore be preferred.

5.2 Extended case: 07^{th} of May 2006 at 0300UTC.

The experiment will now be slightly modified in the same way that in section 4.2, in order to increase the number of forecast files in the procedure. Once again, the situation we are tying to avoid is the one where some analysis data at a time greater than t_0 would be used in the procedure. The field to be post-processed is presented by figure 5.5 with the radar observation at the same time (0300UTC on the 7^{th} of May). The sequence on the following page (figure 5.6) corresponds to all the forecast fields available for the experiment, and shows the evolution of the situation between the 6^{th} at 2100UTC and the 7^{th} at 0730UTC.

This slow moving frontal system offers a new characteristic to test since the very same part of the domain is affected during the whole sequence. Simulations using 0,1,2 and 3 time steps were ran for the two methods, and the results are presented by figure 5.7. The improvement in BSS after one time step is slightly less than it was in the convective case, but the major result here is that the use of a second step still increases its value. This can probably be explained by the particularly slow motion of the front: While most of the convective cells had moved or vanished 3 hours after the forecast time (or were not yet present 3 hours before), the frontal system considered here is still in the same area of the domain. Hence, some useful information about mis-timing errors can be obtained from fields at $t_0 + 2$ and $t_0 - 2$. The search for the optimum parameterisation of the reduction function was ran in the same way that in section 4.2 and the results are shown in figure 5.8. The parameters providing the higher BSS are R_1 =1, R_2 =1 and R_3 =0 meaning that the optimum choice corresponds to the one using two time steps with 100% of their outputs.

The influence of adjacent fields in time on the averaging procedure is not clear but seems to be case dependent: The number of fields to be used should be determined by the type of weather event considered, and more particularly by the time scale over which it might stay in the area of study.

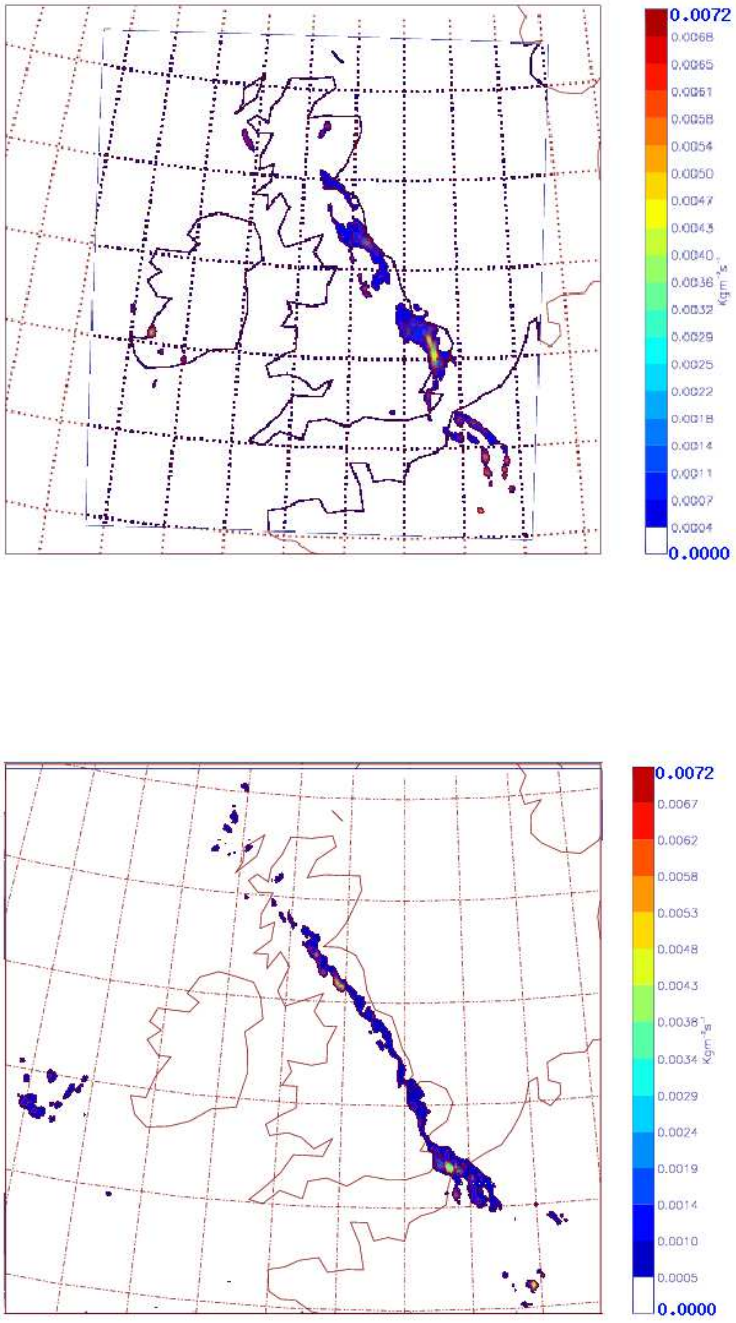

Figure 5.5: Analysis field of rainfall rate (top) and corresponding radar observation (bottom) on the 07^{th} of May 2006 at 0300UTC.

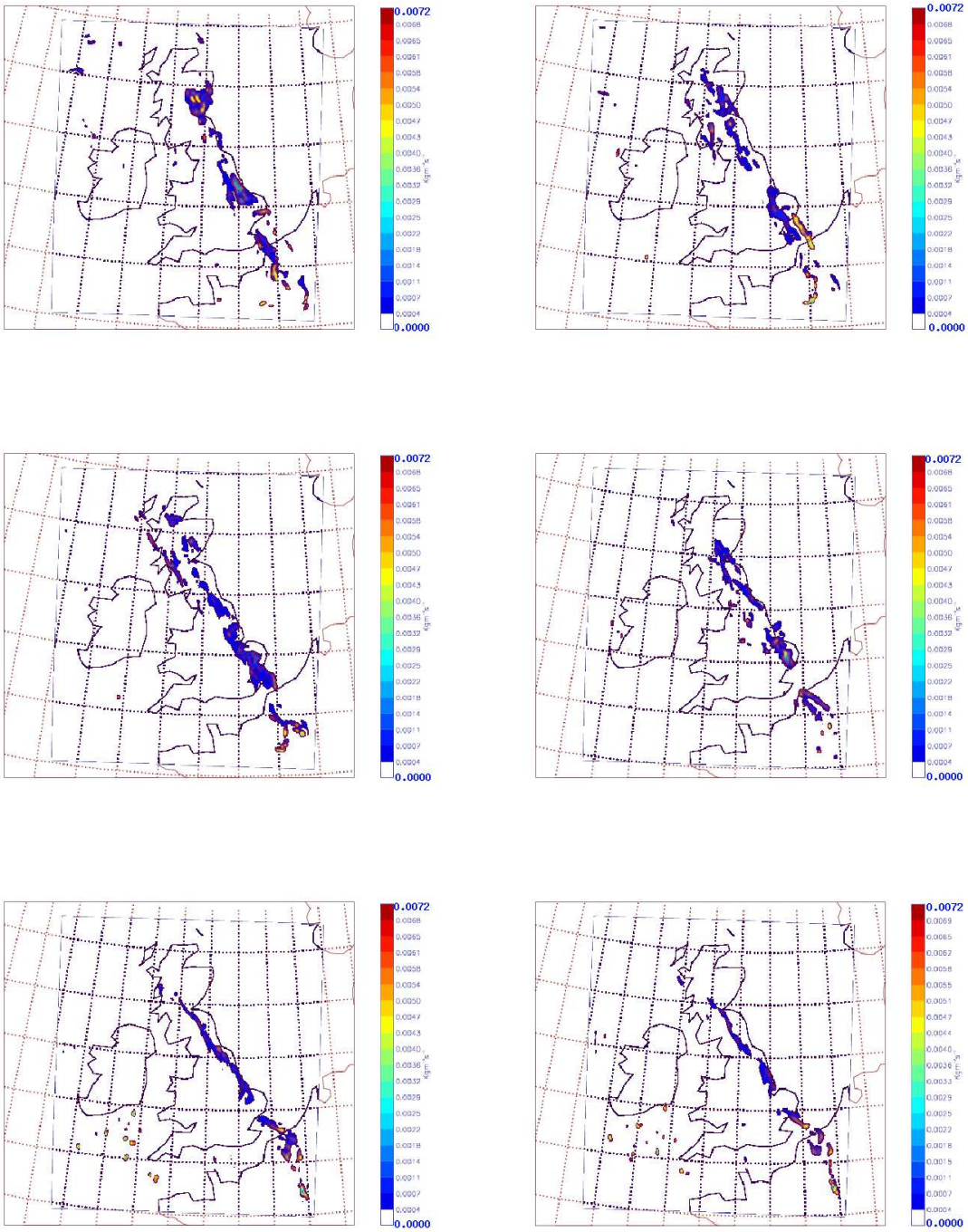

Figure 5.6: Successive model forecasts of rainfall rates between the 6^{th} at 2100UTC (top left) and the 7^{th} at 0730UTC (bottom right).

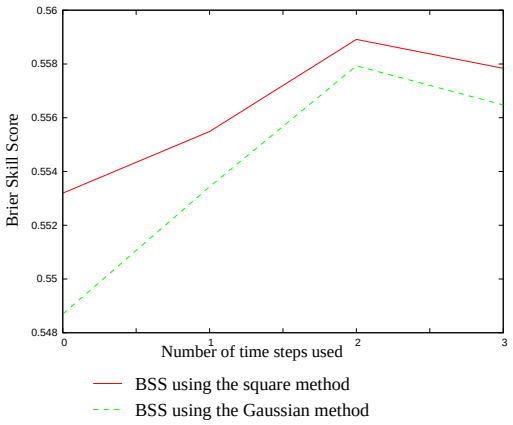

Figure 5.7: BSS for the Gaussian method (green) and the square one (red) as a function of the number of time steps used.

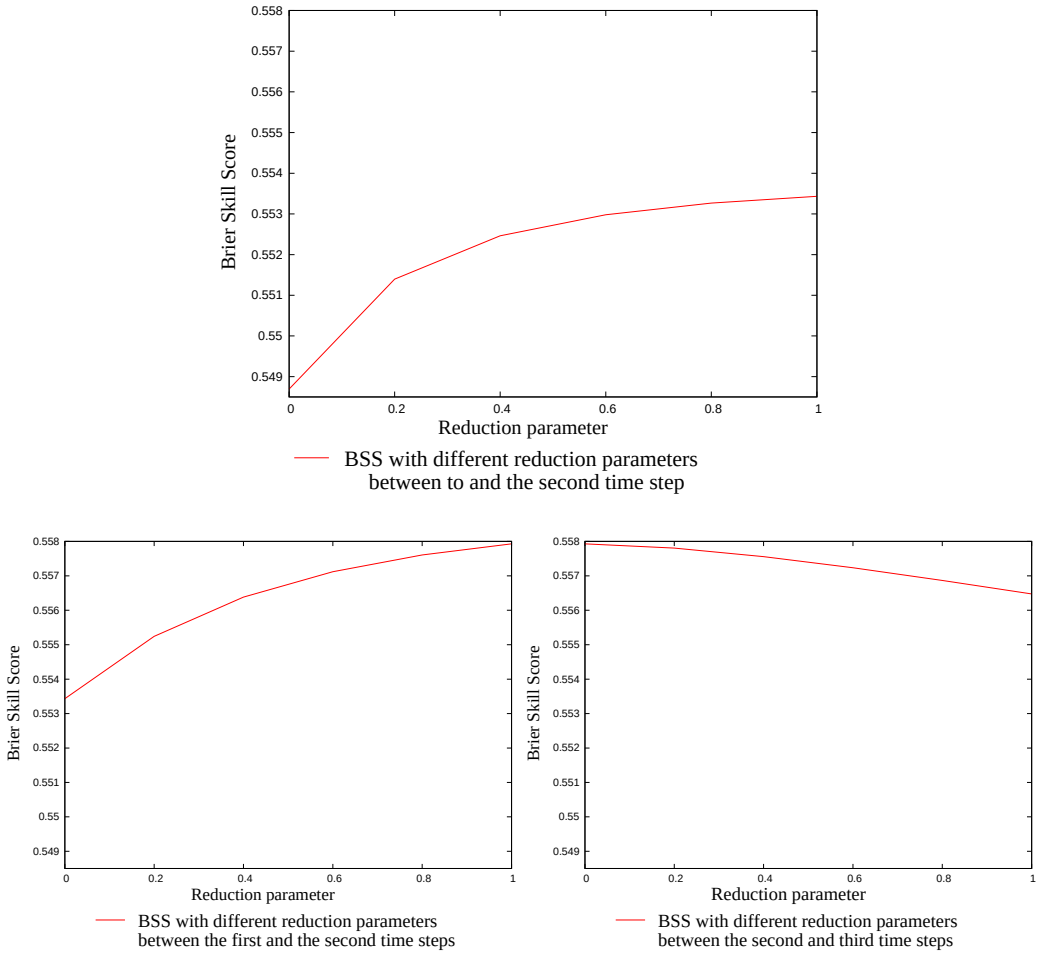

Figure 5.8: Evolution of the BSS for the Gaussian method with different reduction parameters. Top plot shows the selection of parameter R_1, bottom left one the selection of parameter R_2 and bottom right plot the selection of R_3.

Chapter 6

Conclusions and further work

6.1 Conclusions

Precipitation forecasts suffer from an important predictability problem, and probabilistic information are a way to take this unavoidable uncertainty into account. Among the different methods leading to a probabilistic forecast, the neighbourhood approach (Theis et al, 2005) was selected for its low-budget characteristic. While building our own procedure to test a similar approach on the Met Office Unified Model, a few questions arose and were formulated at the end of chapter 3. Using the results gathered during the experiments of chapters 4 and 5, we will now try to provide some answers:

- How will the choice of rainfall rates as a variable to post-process influence our results?

  One particularity of this study was to use the model rain fall rate instead of the more commonly used rainfall accumulation (Theis et al. 2005, Applequist et al. 2002). The extra variability involved due to the absence of time averaging (as opposed to accumulations) has lead to larger values of the Brier Skill Scores: The use of a neighbourhood post-processing method is even more beneficial to rainfall rate's outputs.

- Will the use of a Gaussian weighting function help solve the discontinuity problem, and if it does what will be the impact on the forecast skill score?

  The main motivation for the use of a Gaussian weighting function was to reduce the

discontinuities in the output field which were produced by the initial neighbourhood method. The convective case presented in chapter 4 provided us with a good example of a situation where the Gaussian method did help solve this issue, but the second case considered, involving a frontal system, failed to offer similar conclusions. As far as the skill score is concerned, a slight under-performance of the Gaussian approach was observed in all situations where the two methods were compared. However this difference in Brier Skill Scores was smaller than 1% for the parameter values selected during the experiments.

- What are the effects of the time averaging procedure on this skill score?
  The dependence of the BSS on the number of time steps used was perhaps the most unpredictable aspect of the study: While the convective case showed that only fields situated one and a half hours apart from time t_0 were improving the BSS, the frontal situation still benefited from fields up to three hours away. The use of a reduction function to decrease the weights attributed to fields away from t_0 did not offer any significant improvement, however it turned a decrease in BSS into a small increase in the convective case and therefore the method could perhaps be useful in some other situations: If the model output times were of 2 hours or more for instance, the reduction function would provide a way to build some intermediate time fields.

The best way to summarise these observations is to consider what would be the optimum procedure, which of course depends strongly on the weather event considered. In a convective situation, the Gaussian approach offers some subjective advantages to a forecaster using the post-processed output, and since its under-performance in terms of BSS is very small, it should be preferred to the initial method. The length of the time radius should not be larger then 2 hours, in order to stick to the life time of the convective cells considered.

For a frontal system however, the benefits of the Gaussian weighting function are less obvious, and since it involves some extra costs in running time, the initial square function should be selected instead. On the other hand, the larger time scale over which the front remains in the area of study suggests the use of a wider time radius (3 to 4 hours).

6.2 Further work

An interesting final aim for the project would be to have a completely automatic procedure post-processing all model runs. Since our conclusions suggest a case dependent method, an important step would be to select a criteria for an automatic identification of convective episodes. If the 12km version of the model was used, the convective scheme would provide a good criteria, but since the 4km version mainly deals with convection in its dynamic component, a different approach needs to be taken. One possibility could be to identify a threshold value for the vertical velocity, and start using the convection procedure when the velocity given by the model is higher than this threshold.

Another interesting area of work that has not been considered in this study is the modification of neighbourhood shape in space. It would be interesting for instance to produce an elliptic shape extending in the direction of the wind, to focus on points from which the air is advected. Such a method capable of parameterising the shape would also be extremely beneficial for flood forecasting purposes since the neighbourhood could then be adapted to river catchements boundaries: Main concern for flood forecaster is to estimate the accumulation of rain inside the catchements and the circular neighbourhood method we have been using in this study would lead to some important errors at the boundaries.

Bibliography

[1] Applequist, S., Gahrs, G.E., Pfeffer, R.L., Niu, X.F. (2002). Comparison of methodologies for probabilistic quantitative precipitation forecasting. Weather and Forecasting. **17**, 783 – 799

[2] Benoit, R., and Pellerin, P. (2000). Toward the use of coupled atmospheric and hydrologic models at regional scale. Monthly Weather Review. **128**, 1681 – 1706

[3] Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. Monthly Weather Review. **78**, 1 – 3

[4] Bright, D.R., and Mullen, S.L. (2002).Short-range ensemble forecasts of precipitation during the southwest moonsoon. Weather and Forecasting. **17**, 1080 – 1099

[5] Cassati, B., Ross, G., and Stephenson, D., B (2004). A new intensity-scale approach for the verification of spatial precipitation forecasts. Meteorol. Appl. **11**, 141 – 154

[6] Du, J., and Mullen, L. (1997). Short range ensemble forecasting of quantitative precipitation. Monthly Weather Review. **125**, 2427 – 2459

[7] Erbert, E.E., Damrath, U., Wergen, W., Baldwin, M.E., (2003). The WGNE assessment of short-term quantitative precipitation forecasts. Bulletin of the American Meteorological Society. **84**, 481 – 492

[8] Jasper, K., and Kaufmann, P. (2003).Coupled runoff simulations as validation tools for atmospheric models at the regional scale. Q.J.R.Meteorol. Soc. **129**, 673 – 692

[9] Krzysztofowicz, R. (1993). Probabilistic quantitative precipitation forecast for river basins. Weather and Forecasting. **8**, 424 – 439

[10] Lorenz, E. N. (1963). Deterministic nonperiodic flow. J. Atmos. Scie. **20**, 130 – 141

[11] Marsigli, C., Montani, A., Nerozzi, F., Paccagnella, T., Tibaldi, S., Molteni, F. (2001). A strategy for high-resolution ensemble prediction.2: Limited-area experiments in four alpine flood events. Q.J.R.Meteorol.Soc. **127**, 2095 – 2115

[12] Mullen, S.L., and Baumhefner, D.P. (1989). The impact of initial condition uncertainty on numerical simulations of large-scale explosive cyclogenesis. Monthly Weather Review. **20**, 130 – 141

[13] Murphy, A.H. (1991). Probabilities, odds, and forecasts of rare events. Weather and Forecasting. **6**, 302 – 307

[14] Richardson, D.S. (2000). Skill and relative economic value of the ECMWF ensemble prediction system. Q.J.R.Meteorol.Soc. **126**, 649 – 667

[15] Roberts, N. (2006). Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Submitted in 2006.

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

[17] Walser, A., Luthi, D., and Schar, C. (2004). Predictability of precipitation in a cloud-resolving model. Monthly Weather Review. **132**, 560 – 577

Appendix: Fortran program

```
!Program  to post-process direct model outputs
!using a neighbourhood approach
!
program neighb

implicit none

character(80)                          :: OBS1='rad050703Z.pp'
character(80)                          :: FILE1='20060506HS.21Z.pp'
character(80)                          :: FILE2='20060507HS.03Z.pp'
character(80)                          :: raintype
integer, parameter                     :: DX=4000
real                                   :: threshold, fact
integer                                :: i,i0,j,j0,n,n0
integer                                :: nlow,nup, tlevel
integer                                :: ERROR,NX,NY, alpha, beta
integer                                :: EXC, EVENT, EXCDMO
integer, dimension(45)                 :: IHEAD, IHEAD2, IBADHEAD
logical                                :: END1
real                                   :: dist, radius, W1, SUMW1, W2, SUMW2, size
real                                   :: PI, BS1, BS2, BSREF, BSS1, BSS2
real                                   :: constant, constant2, W1BAK, W2BAK
real, dimension(19)                    :: RHEAD, RHEAD2, RBADHEAD
real, dimension(7)                     :: THR
real, dimension(6)                     :: CONST
integer, dimension(6)                  :: NSIZE
real, dimension(:,:,:), allocatable :: LSRR, CORR, DATA
real, dimension(:,:), allocatable   :: DUMMY, PROB1, PROB2, OBS, RADAR

!************************************************************************
```

```
PI=4*ATAN(1.0)

print*,'#########'
print*, 'START'
print*,'#########'

! Open files
      OPEN(UNIT=51,FILE='BSBDAY',FORM='FORMATTED',STATUS='UNKNOWN')
      OPEN(UNIT=61,FILE=OBS1,FORM='UNFORMATTED',STATUS='OLD')
      OPEN(UNIT=81,FILE=FILE1,FORM='UNFORMATTED',STATUS='OLD')
      OPEN(UNIT=82,FILE=FILE2,FORM='UNFORMATTED',STATUS='OLD')
      OPEN(UNIT=91,FILE='rates.pp',FORM='UNFORMATTED',STATUS='UNKNOWN')
      OPEN(UNIT=93,FILE='DMOBDAY.pp',FORM='UNFORMATTED',STATUS='UNKNOWN')
      OPEN(UNIT=94,FILE='pbsquare.pp',FORM='UNFORMATTED',STATUS='UNKNOWN')
      OPEN(UNIT=95,FILE='pbgauss.pp',FORM='UNFORMATTED',STATUS='UNKNOWN')

! Allocate arrays, read and write

READ(81)IHEAD,RHEAD
      NX=IHEAD(19)
      NY=IHEAD(18)
      ALLOCATE(LSRR(NX,NY,4))
      ALLOCATE(CORR(NX,NY,4))
      ALLOCATE(DUMMY(NX,NY))
      ALLOCATE(DATA(NX,NY,4))
      ALLOCATE(PROB1(NX,NY))
      ALLOCATE(PROB2(NX,NY))
      ALLOCATE(RADAR(NX,NY))
REWIND(81)

READ(61)IBADHEAD,RBADHEAD
```

```
READ(61)radar(1:NX,1:NY)

!start reading first file
  Do i=1,16
       READ(81)IHEAD,RHEAD
        WRITE(92,*) i
       WRITE(92,*)IHEAD,RHEAD
       if(IHEAD(42) .eq. 4203) then
          WRITE(91)IHEAD,RHEAD
           if(IHEAD(14) .eq. 0) then
            READ(81)LSRR(1:NX,1:NY,1)
            WRITE(91)LSRR(1:NX,1:NY,1)
           elseif(IHEAD(14) .eq. 1) then
            READ(81)LSRR(1:NX,1:NY,2)
            WRITE(91)LSRR(1:NX,1:NY,2)
           elseif(IHEAD(14) .eq. 4) then
            READ(81)LSRR(1:NX,1:NY,3)
            WRITE(91)LSRR(1:NX,1:NY,3)
           elseif(IHEAD(14) .eq. 5) then
            READ(81)LSRR(1:NX,1:NY,4)
            WRITE(91)LSRR(1:NX,1:NY,4)
           endif
       elseif(IHEAD(42) .eq. 5205) then
           WRITE(91)IHEAD,RHEAD
           if(IHEAD(14) .eq. 0) then
            READ(81)CORR(1:NX,1:NY,1)
            WRITE(91)CORR(1:NX,1:NY,1)
           elseif(IHEAD(14) .eq. 1) then
            READ(81)CORR(1:NX,1:NY,2)
            WRITE(91)CORR(1:NX,1:NY,2)
           elseif(IHEAD(14) .eq. 4) then
```

```
            READ(81)CORR(1:NX,1:NY,3)
            WRITE(91)CORR(1:NX,1:NY,3)
           elseif(IHEAD(14) .eq. 5) then
            READ(81)CORR(1:NX,1:NY,4)
            WRITE(91)CORR(1:NX,1:NY,4)
           endif
      else
           READ(81)DUMMY(1:NX,1:NY)
      endif
  enddo

!start reading second file

 Do i=1,16
      READ(82)IHEAD,RHEAD
      if(IHEAD(42) .eq. 4203) then
         WRITE(91)IHEAD,RHEAD
          if(IHEAD(14) .eq. 0) then
           READ(82)LSRR(1:NX,1:NY,5)
           WRITE(91)LSRR(1:NX,1:NY,5)
          elseif(IHEAD(14) .eq. 1) then
           READ(82)LSRR(1:NX,1:NY,6)
           WRITE(91)LSRR(1:NX,1:NY,6)
          elseif(IHEAD(14) .eq. 4) then
           READ(82)LSRR(1:NX,1:NY,7)
           WRITE(91)LSRR(1:NX,1:NY,7)
          elseif(IHEAD(14) .eq. 5) then
           READ(82)LSRR(1:NX,1:NY,8)
           WRITE(91)LSRR(1:NX,1:NY,8)
          endif
      elseif(IHEAD(42) .eq. 5205) then
```

```
          WRITE(91)IHEAD,RHEAD
          if(IHEAD(14) .eq. 0) then
          READ(82)CORR(1:NX,1:NY,5)
           WRITE(91)CORR(1:NX,1:NY,5)
          elseif(IHEAD(14) .eq. 1) then
           READ(82)CORR(1:NX,1:NY,6)
           WRITE(91)CORR(1:NX,1:NY,6)
         elseif(IHEAD(14) .eq. 4) then
           READ(82)CORR(1:NX,1:NY,7)
           WRITE(91)CORR(1:NX,1:NY,7)
          elseif(IHEAD(14) .eq. 5) then
           READ(82)CORR(1:NX,1:NY,8)
           WRITE(91)CORR(1:NX,1:NY,8)
          endif
      else
           READ(82)DUMMY(1:NX,1:NY)
      endif
enddo

!*************************************************************************
!Compute a neighbourhood around grid point (i0,j0,n0)

raintype='SUM'
nlow=3
n0=5
nup=7
!*************************************************************************

if (raintype .eq. 'CON') then
 do i=nlow,nup
  DATA(:,:,i)=CORR(:,:,i)
```

```
 enddo
elseif (raintype .eq. 'LS') then
 do i=nlow,nup
  DATA(:,:,i)=LSRR(:,:,i)
 enddo
elseif (raintype .eq. 'SUM') then
 do i=nlow,nup
  DATA(:,:,i)=CORR(:,:,i)+LSRR(:,:,i)
 enddo
endif

!*********************************************************************
THR=(/1.0E-4, 5.0E-4, 10.0E-4, 15.0E-4, 20.0E-4, 25.0E-4, 30.0E-4/)
CONST=(/ 0.0, 0.2, 0.4, 0.6, 0.8, 1.0/)
NSIZE=(/ 5, 10, 15, 20, 25, 30 /)
!*********************************************************************
do tlevel=1,6

  !threshold=THR(tlevel)
  threshold= 1.0E-3
  constant2=CONST(tlevel)
  !alpha=NSIZE(tlevel)
  alpha=15

print*,' '
print*,'*******************************************************'
  print*, ' '
  print*, 'threshold is ',threshold
  print*, 'spatial radius is ',alpha
  print*, 'time constant is ',constant2
```

```
  constant=(4*(alpha)**2)/PI

  print*, 'Space constant is ',constant

PROB1=0
PROB2=0
BS1=0
BS2=0
BSREF=0
size=0

do i0=alpha+1,NX-alpha
  do j0=alpha+1,NY-alpha

    size=size+1
    SUMW1=0.
    SUMW2=0.
    EXC=0
    W1=0.
    W2=0.

       do i=alpha+1,NX-alpha
         do j=alpha+1,NY-alpha

          dist=((i-i0)**2+(j-j0)**2)**0.5
          IF(dist.gt.2*alpha) CYCLE
          If(dist .le. alpha) W1=1.
          W1BAK=W1

          W2=exp((-dist**2)/constant)
          W2BAK=W2
```

```
      do n=nlow,nup

                  W1=W1BAK

                  W2=W2BAK
                  !W2=(1-(abs(n-n0)*constant2))*W2

                  If(abs(n-n0).eq.1) W2=W2
                  If(abs(n-n0).eq.2) W2=W2
                  If(abs(n-n0).eq.3) W2=W2*constant2

                If(DATA(i,j,n) .ge. threshold) then
                 EXC=1
                else
                 EXC=0
                endif

               PROB1(i0,j0)=PROB1(i0,j0)+W1*EXC
               SUMW1=SUMW1+W1

               PROB2(i0,j0)=PROB2(i0,j0)+W2*EXC
               SUMW2=SUMW2+W2

        enddo

   enddo
enddo

 PROB1(i0,j0)=PROB1(i0,j0)/SUMW1
 PROB2(i0,j0)=PROB2(i0,j0)/SUMW2
```

```
      EVENT=0
      if (RADAR(i0,j0) .ge. threshold) EVENT=1
      EXCDMO=0
      if (DATA(i0,j0,n0) .ge. threshold) EXCDMO=1

      BS1=BS1+(PROB1(i0,j0)-EVENT)**2
      BS2=BS2+(PROB2(i0,j0)-EVENT)**2
      BSREF=BSREF+(EXCDMO-EVENT)**2

  enddo
enddo

BS1=BS1/size
BS2=BS2/size
BSREF=BSREF/size
BSS1=1-BS1/BSREF
BSS2=1-BS2/BSREF

!write direct model output for DATA
!do n=nlow,nup
! write(93)IHEAD,RHEAD
! write(93)DATA(1:NX,1:NY,n0)
!write probabilities
! write(94)IHEAD,RHEAD
! write(94)PROB1(1:NX,1:NY)
! write(95)IHEAD,RHEAD
! write(95)PROB2(1:NX,1:NY)
!*******************************************************************************

WRITE(51,*)alpha, BS1, BS2, BSREF, BSS1, BSS2, BSS1-BSS2
```

```
print*,  'Brier score (theis et al):',BS1
print*,  'Brier score (gaussian):',BS2
print*,  'reference score:',BSREF
print*,  'Brier skill score (theis et al):',BSS1
print*,  'Brier skill score (gaussian):',BSS2
print*,  'difference BSS1-BSS2: ', BSS1-BSS2

!enddo ! tlevel

!*******************************************************************************
      CLOSE(51)
      CLOSE(81)
      CLOSE(82)
      CLOSE(91)
      CLOSE(92)
      CLOSE(93)
      CLOSE(94)

      WRITE(*,*) 'neighb v1.1 COMPLETED SUCCESSFULLY'

      WRITE(*,*) 'EXECUTION TERMINATED'
!
      END PROGRAM
```