A Bayesian state space modelling approach to probabilistic quantitative precipitation forecasting

Dan Cornford

Neural Computing Research Group,

School of Engineering and Applied Science,

Aston University,

Birmingham B4 7ET, UK

`d.cornford@aston.ac.uk`

Tel. 0121 359 3611 x4667, Fax. 0121 333 4586

21/11/02

Abstract

The generation of very short range forecasts of precipitation in the 0–6 hours time window is traditionally referred to as nowcasting. Most existing nowcasting systems essentially extrapolate radar observations in some manner, however, very few systems account for the uncertainties involved. Thus deterministic forecast are produced, which have a limited use when decisions must be made, since they have no measure of confidence or spread of the forecast. This paper develops a Bayesian state space modelling framework for quantitative precipitation nowcasting which is probabilistic from conception. The model

1

treats the observations (radar) as noisy realisations of the underlying true precipitation process, recognising that this process can never be completely known, and thus must be represented probabilistically. In the model presented here the dynamics of the precipitation are dominated by advection, so this is a probabilistic extrapolation forecast. The model is designed in such a way as to minimise the computational burden, while maintaining a full, joint representation of the probability density function of the precipitation process. The update and evolution equations avoid the need to sample, thus only one model needs be run as opposed to the more traditional ensemble route. It is shown that the model works well on both simulated and real data, but that further work is required before the model can be used operationally.

1 Introduction

The provision of reliable forecasts of precipitation over the 0–6 hour period at high spatial resolution, often described as now-casting, is regarded as essential for flood forecasting by the Environment Agency (Golding, 2000). To adequately describe the spatial distribution and timing of precipitation over catchments with areas of the order of 10 km^2, prediction over a spatial scale of the order of 1 km is required, with a temporal scale of the order of 10 minutes. In the foreseeable future forecasts of precipitation at such high resolutions will require the application of probabilistic methods since precipitation generating processes are very sensitive to the correct specification of initial and boundary conditions, which is unlikely to be resolved in operational meso-scale models for some time, largely due to problems of data assimilation and parameterisations.

1.1 Current approaches to precipitation nowcasting

Operationally, precipitation forecasts are produced over a variety of time and space scales, using a variety of methods. We shall assume that nowcasting refers to the production of forecasts with lead times of $0-6$ hours (Golding, 2000), while short-term forecasting refers to lead times of $6-24$ hours (Collier, 2000). Beyond 24 hour lead times it is widely accepted that NWP based approaches are optimal for providing precipitation forecasts (Golding, 2000). In the range of $0-24$ hours a variety of methods are used to provide quantitative precipitation forecasts including:

- NWP based approaches (Kuligowski and Barros, 1999);

- model output statistics techniques (Antolik, 2000; Fox and Collier, 2000);

- purely statistical extrapolation methods, generally using radar data as a first guess (Grecu and Krajewski, 2000; Toth *et al.*, 2000; Mellor *et al.*, 2000);

- expert system based approaches which model precipitation cell evolution (Pierce *et al.*, 2000).

These different approaches have strengths and weaknesses for a range of forecast lead times, time scales and space scales. In the nowcasting range of $0-6$ hours NWP based forecasts are less often used because the assimilation and initialisation cycle of these models is of the order of 3 hours (Golding, 2000) and there are many unresolved issues surrounding the assimilation of precipitation data into these models (Zou and Kuo, 1996). Thus, to produce nowcasts of precipitation, various statistical methods have been developed, which take advantage of the availability of high resolution radar. In the UK

the radar system is only capable of providing information of instantaneous rainfall rates, so this is what is assumed to be available.

Wheater *et al.* (2000) review a series of elegant statistical spatio-temporal model for describing precipitation fields, based on a hierarchical decomposition of the precipitation into precipitation fields, bands and cells, each of which are modelled as point processes in space (precipitation fields), time (precipitation bands) or space and time (cells). Thus the models are defined as being continuous in space and time, and the actual precipitation field is described by a series of precipitation cells. The models can be fit to radar data, but this fitting is designed to to estimate the hyper-parameters of the models, such as the rates of the corresponding Poisson processes. Thus the models provide a statistical characterisation of the precipitation field but are not directly suitable for use in operational forecasting.

In a similar vein Mellor *et al.* (2000) describe a statistical model for the production of a space-time precipitation field, based on their so called Modified Turning Bands method. This uses triangular prisms to define the location of the regions with potential to generate precipitation, and these are modulated by a paired sinusoidal function to produce temporal pulsing in the resulting precipitation cell generation potential field. Cells are generated from this potential as a Poisson process (the potential defines the rate of the process). Each precipitation cell is described by an inverted parabola, and has its own characteristic lifetime, width and maximal intensity. The model requires manual intervention to be fit to data and is tuned to frontal systems with a linear structure. By generating realisations of the Poisson processes ensembles can be generated, but the rather ad-hoc nature of the fitting of the model to data means that the operational use of this system is currently limited.

In the GANDOLF system, used operationally in the Met Office (Pierce *et al.*, 2000), convection is represented using an object-oriented approach whereby each convective cell is identified and evolved through a life-cycle model, which includes initiation of daughter cells. The system uses data from a variety of sources including the Met Office mesoscale model, METEOSAT images and the radar network. While the GANDOLF system has been shown to produce reasonable convective precipitation forecasts (Pierce *et al.*, 2000), there are several areas which could be improved. One of the most significant problems is that one single deterministic forecast is produced which, for such a non-linear system, is unlikely to characterise the true distribution of precipitation at a forecast lead time of say 3 hours.

Georgakakos (2000) developed a hybrid model, which has aspects of the model output statistics approach, in that NWP derived fields are used extensively, for example to predict advection vectors, but also directly uses radar data (both near ground precipitation rates and vertically integrated water content) to produce a nowcast. The paper is novel in that a method is developed which also incorporates model uncertainty, through the specification of both observation and model errors. In that respect is it similar to the work described in this paper, however Georgakakos (2000) assumes a discrete state space (that is a grid based model, with numerical integration in the standard way) and thus can only represent the joint covariance matrix of the precipitation field in neighbouring cells, which is rather restrictive.

While there have been some interesting research developments in statistical modelling for precipitation forecasting, there remain many unanswered questions, in terms of what framework is most appropriate and how model uncertainty and observation uncertainty can be combined to capture the full probability distribution function of precipitation through space and time. In

the next section, one possible approach is reviewed.

1.2 Bayesian state space modelling

To progress to the development of probabilistic models for precipitation, which can then be fed into a Bayesian hydrologic forecasting system such as presented in Krzysztofowicz (1999), we essentially have three options:

1. use model output statistics techniques to post-process deterministic NWP (Antolik, 2000);

2. use Monte Carlo (ensemble) methods with statistical or NWP models (Mellor *et al.*, 2000);

3. directly represent and evolve the probability density function of the precipitation process (this paper, and to some extent (Georgakakos, 2000)).

The first method, using model output statistics has a disadvantage, that the probability density function (pdf) of the precipitation process will be complex and state dependent, which may make parameterisation very difficult. In addition, many users (for instance those running distributed hydrological models) will require the joint pdf for precipitation over some spatial domain, so that the spatial structure of the precipitation is preserved, which will be very difficult to achieve.

The second option, based on sampling methods uses a large number of samples to characterise the pdf. The problem with these methods is that to correctly characterise the joint pdf of precipitation over a spatial domain may require many thousands of samples. Current operational ensemble prediction systems use around 100 samples, and are computationally very expensive, since each member must be integrated separately (Molteni *et al.*,

1996). However ensembles maintain an advantage in that any model can readily be run as an ensemble, if the computational expense can be coped with. In addition it may be possible to post-process output to produce better approximations to the true forecast pdf (although this will again be very difficult for joint pdfs) (Mylne *et al.*, 2002).

The third option is pursued in this paper, that is a model will be constructed which retains a probabilistic representation of the evolution of the precipitation field. Thus the model must propagate the full joint pdf of the precipitation field, and update this as observations become available. State space models provide a suitable framework, these being most readily presented in their natural Bayesian context. Any dynamical system can be represented in the state space modelling framework, which provides formalism for the updating and propagation of the pdf of the state of a system. In the linear, Gaussian case this is generally referred to as the Kalman filter. For non-linear and non-Gaussian systems there are a large variety of extensions to the Kalman filter (Chatfield, 1996), such as the extended Kalman filter, the ensemble Kalman filter (Evensen, 2001) and particle filter (Doucet *et al.*, 2001) based methods.

State space models assume that the state of the system at time t, denoted $\boldsymbol{x}_t$, is not directly observed (it is a latent variable). The state is taken to evolve in time according to the (Markovian) state evolution equation:

$$\boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t) + \boldsymbol{\eta}_t \ ,$$

where $f()$ is the state evolution (or system) model which maps the state at time t to the new value at time $t + 1$ and $\boldsymbol{\eta}_t$ is the system noise. This system noise represents model error, and can often be difficult to determine. The observations at time t, $\boldsymbol{y}_t$, are related to the state by the observation

equation:

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t) + \boldsymbol{\epsilon}_t ,$$

where $h()$ is the observation (or forward / sensor) model which maps the state variable to the observables. $\boldsymbol{\epsilon}_t$ is the observation error and reflects the uncertainty in the observations which may arise from several sources, including incomplete knowledge of the observation process, such as arises when using radar images, as well as the intrinsic measurement uncertainty due to the observation system.

This framework is completely general, since there have been no restrictions imposed on f, h, $\boldsymbol{\eta}$ or $\boldsymbol{\epsilon}$. In the standard Kalman filter f and h are no longer functions but linear operators (matrices) and $\boldsymbol{\eta}$ and $\boldsymbol{\epsilon}$ are assumed Gaussian, as is the initial distribution of the of the state $\boldsymbol{x}_0$. Thus for all future times the state will remain Gaussian distributed.

In this work the aim is to track the pdf of the state given all the previous observations, denoted by $\mathrm{p}(\boldsymbol{x}_t \mid \boldsymbol{D}_t) = \mathrm{p}(\boldsymbol{x}_t \mid \boldsymbol{y}_t, \boldsymbol{y}_{t-1}, ..., \boldsymbol{y}_1)$.

[Figure 1 about here.]

The observation process is written as $\mathrm{p}(\boldsymbol{y}_t \mid \boldsymbol{x}_t, h)$, the conditional probability of the observations at time t given our estimate of the state at time t and the observation model, h. This reflects the assumption that the observation process in often not completely understood and that it has errors associated with it. The evolution of the system is given by $\mathrm{p}(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_t, f)$, the conditional probability of the state at time $t+1$ given our estimate of the state at time t and the system model, f. In section 2 these are shown for the precipitation model.

The Bayesian interpretation the state evolution (forecast) step becomes:

$$\mathrm{p}(\boldsymbol{x}_t \mid \boldsymbol{D}_{t-1}) = \int \mathrm{p}(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1})\mathrm{p}(\boldsymbol{x}_{t-1} \mid \boldsymbol{D}_{t-1})\mathrm{d}\boldsymbol{x}_{t-1} .$$

The state update (assimilation) step can then be written as:

$$p(\boldsymbol{x}_t \mid \boldsymbol{D}_t) = \frac{p(\boldsymbol{y}_t \mid \boldsymbol{x}_t)p(\boldsymbol{x}_t \mid \boldsymbol{D}_{t-1})}{\int p(\boldsymbol{y}_t \mid \boldsymbol{x}_t)p(\boldsymbol{x}_t \mid \boldsymbol{D}_{t-1})\mathrm{d}\boldsymbol{x}_t} \ ,$$

which can be seen as the application of Bayes theorem to the inverse problem of estimating the state given a set of observations, where the prior distribution, $p(\boldsymbol{x}_t \mid \boldsymbol{D}_{t-1})$ comes from the state evolution step. These equations are then applied sequentially, to propagate the pdf of the state forward in time and update it given some observations, as illustrated in Figure 1.

This general framework is very flexible and can adapted to almost every modelling situation. In section 2 the state space modelling framework is extended to the probabilistic quantitative precipitation forecast model. In section 3 the prior models are discussed and it is shown how these can be used generatively, to simulate from the model, even when no data has been seen. Section 4 discusses the results of the application of the model to both simulated (where we know the true parameters) and real data and conclusions are drawn in section 5.

2 Model framework

The state space for the precipitation model must consist of the key variables necessary to forecast precipitation over the 0–3 hour range. At a minimum there must be a representation of the instantaneous precipitation rate field, which we will denote $\boldsymbol{R}$. To account for advection, the state space must be extended to include a vector field to represent the continuous spatial behaviour of the advection of the precipitation, which we denote $\boldsymbol{u}$. The model dynamics are very simple, the evolution of $\boldsymbol{R}$ being given by:

$$\frac{\partial \boldsymbol{R}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{R} \approx 0 \ ,$$

while $\boldsymbol{u}$ is a purely stochastic process. This reflects the belief that the advection field changes more slowly (in most situations) than the precipitation field. The formulation of the model as approximately preserving precipitation in the Lagrangian sense limits the applicability to short forecast lead times, and the addition of external forcing terms will be the subject of future work.

The model chosen to represent $\boldsymbol{R}$ is a radial basis function model given by:

$$\boldsymbol{R}(\boldsymbol{x}) = \sum_{k=1}^{N} h_k \exp\left[-\frac{1}{2}\frac{(\boldsymbol{x}-\boldsymbol{c}_k)^2}{w_k^2}\right] \ ,$$

where h_k represents the height, w_k the width and $\boldsymbol{c}_k$ the centre of the k'th basis function and $\boldsymbol{x}$ represents a 2 dimensional position vector. By choosing a suitable number of basis functions, N, it is possible to approximate any continuous mapping (Bishop, 1995), thus this model is flexible enough to represent any precipitation field over a spatial domain. The Gaussian form of the basis functions, chosen for computational reasons to produce a continuous model, implies that $\boldsymbol{R}$ is never truly zero, which is not very physically realistic. When we use this model to produce predictions of precipitation rates then we post-process the output, and assume that any values below some threshold are assigned to zero.

The model for $\boldsymbol{u}$ is based on a vector Gaussian process, $N(\boldsymbol{0}, \Sigma_u)$, where the covariance matrix Σ_u is carefully defined so that $\nabla \times \boldsymbol{u} \gg \nabla \cdot \boldsymbol{u}$. The parameters that define the covariance matrix are set to give sensible space and energy scales (Nabney *et al.*, 2000).

[Figure 2 about here.]

In the state space modelling framework the aim is to propagate the probability distribution function of the state variables through time, updating

10

them once observations become available. The state vector for the model is $\{\boldsymbol{c}, w, h\}$ and $\boldsymbol{u}$, and it is necessary to propagate through time the means and (co)variances of these variables. The complete model framework is shown in Figure 2.

The model is initialised carefully using the first two radar images in a sequence, but since this need only be done once, the mechanism is not described here; essentially it is based on data assimilation steps, where the forecast is assumed to be very bad. This provides initial estimates at $t = 0$ of the state variables and their (co)variances. The model uses vector Gaussian process priors over $\boldsymbol{c}$ of the form $N(\boldsymbol{0}, \Sigma_c)$, Σ_c is a full covariance matrix, which ensures that centres can be correlated and thus take into account the structure in $\boldsymbol{u}$. The priors over h and w log-Gaussian priors, which ensures $\boldsymbol{R} \geq 0$ and constrain the values to be physically realistic.

These prior distributions are only used at the first step, however the decomposition of the posterior these imply is maintained. Thus the state vector probability distribution function, which in the complete form is $\mathrm{p}(\boldsymbol{c}, w, h, \boldsymbol{u}|\cdot)$ is represented as a product (factorising) distribution $\mathrm{p}(\boldsymbol{c}|\cdot)\mathrm{p}(w|\cdot)\mathrm{p}(h|\cdot)\mathrm{p}(\boldsymbol{u}|\cdot)$, which makes inference more simple. Both $\mathrm{p}(\boldsymbol{c}|\cdot)$ and $\mathrm{p}(\boldsymbol{u}|\cdot)$ are represented by Gaussian processes, and thus have full covariance matrices which allow for complex correlation structures. $\mathrm{p}(w\,|\,\cdot)$ and $\mathrm{p}(h\,|\,\cdot)$ are assumed to have log-normal distributions with diagonal covariance matrices: that is interactions between basis functions are restricted to the locations of the basis functions. This is reasonable since the aim is to capture the uncertainty in the model estimates of advection and propagate these in a consistent and computationally efficient manner.

2.1 Forecast step

To produce forecasts from the model, the methods used for state evolution for $\boldsymbol{R}$ and $\boldsymbol{u}$ must be defined. $\boldsymbol{R}$ is predicted first (denoted the circled 1 in Figure 2). It is assumed that:

$$\frac{\partial \boldsymbol{R}}{\partial t} \approx -\boldsymbol{u} \cdot \nabla \boldsymbol{R} \,.$$

With the radial basis function model for $\boldsymbol{R}$, if $\boldsymbol{u}$ is assumed to be locally constant, this yields a forecast step for $\boldsymbol{R}$ given by:

$$\boldsymbol{c}_{t+1} = \boldsymbol{c}_t + \delta t \, \boldsymbol{u}_t + \boldsymbol{\epsilon}_c \,,$$

where δt is length of the forecast step and $\boldsymbol{\epsilon}_c$ is the error in the forecast due to the simplifications of the model and that not all apparent cell motion is due to advection. Since $\boldsymbol{c}$ and $\boldsymbol{u}$ are both Gaussian and $\boldsymbol{\epsilon}_c$ is assumed Gaussian, this forecast distribution is also Gaussian:

$$\hat{\boldsymbol{c}}_{t+1} \sim N(\bar{\boldsymbol{c}}_t + \delta t * \bar{\boldsymbol{u}}_t, \Sigma_{\boldsymbol{c}_t} + \delta t^2 \Sigma_{\boldsymbol{u}_t} + \Sigma_{\epsilon\boldsymbol{c}}) \,, \tag{1}$$

where the hat is used to denote forecast quantities and the over-bar denotes the expectation (or mean). $\Sigma_{\epsilon\boldsymbol{c}}$ is the covariance of $\boldsymbol{\epsilon}_c$ and it should be noted that at the prediction steps the covariance matrices are added together, implying that our uncertainty increase. The forecasts for w and h are more simple since there is currently no explicitly resolved growth and decay. Thus they can be written:

$$\hat{w}_{t+1} \sim N(\bar{w}_t, \Sigma_{w_t} + \Sigma_{\epsilon w}) \,, \tag{2}$$

and

$$\hat{h}_{t+1} \sim N(\bar{h}_t, \Sigma_{h_t} + \Sigma_{\epsilon h}) \,, \tag{3}$$

where $\Sigma_{\epsilon w}$ and $\Sigma_{\epsilon h}$ are the system noise covariances for w and h respectively and are assumed diagonal.

The assumptions made about the evolution of c (particularly the locally constant assumption on u) are not overly restrictive. The model is designed to run on a 5 minute time step (determined by the frequency of radar observations available) and typical advection speeds are of the order of 10 ms^{-1}, this implies that the locally constant assumption need only apply over a distance of ~ 3 km. It is not difficult to produce a better approximation using information on the gradient, even curvature of the u, but this was not felt to be a significant source of error within the model.

The forecast step for u (denoted the circled 2 in Figure 2) is straight forward since the assumption is made that this does not have temporal dynamics. Thus:

$$\hat{u}_{t+1} \sim N(\bar{u}_t, \Sigma_{u_t} + \Sigma_{\epsilon u}) \, , \tag{4}$$

where $\Sigma_{\epsilon u}$ is the covariance matrix of the system noise (model uncertainty) concerning u, and is a full covariance matrix, with a spatial structure similar to that imposed by the prior, but with much smaller variances. The estimation of these system noise variances is very non-trivial and will be addressed later.

2.2 Assimilation step

Having forecast the state variables, then next step is to update the state given the observations, which in the case of radar estimates of precipitation intensity, have quite a large degree of associated uncertainty. The hierarchical nature of the model means that R must be update before u is updated, and that the uncertainty in the estimates of R must be taken into account when updating u.

Since the relation between the observations I and the parameterised representation of R (that is $\{c, w, h\}$) is not linear, this update (denoted

the circled 3 in Figure 2) is non-trivial . The aim is to determine

$$p(\boldsymbol{c}_{t+1}, w_{t+1}, h_{t+1} \mid \boldsymbol{I}_{t+1}) = \frac{p(\boldsymbol{I}_{t+1} \mid \boldsymbol{c}_{t+1}, w_{t+1}, h_{t+1})p(\boldsymbol{c}_{t+1}, w_{t+1}, h_{t+1})}{p(\boldsymbol{I}_{t+1})} \; ,$$

(5)

where the 'prior' $p(\boldsymbol{c}_{t+1}, w_{t+1}, h_{t+1}) = p(\hat{\boldsymbol{c}}_{t+1})p(\hat{w}_{t+1})p(\hat{h}_{t+1})$ is obtained from Equations 1, 2 and 3. The likelihood $p(\boldsymbol{I}_{t+1}|\boldsymbol{c}_{t+1}, w_{t+1}, h_{t+1})$ is defined by the errors on the radar observations, but is a function of the $\boldsymbol{R}$ model parameters. The normalising constant $p(\boldsymbol{I}_{t+1})$ is unknown.

In general, due to the non-linearity of the radial basis function model, although the prior has been assumed Gaussian and the noise model for the radar is also assumed Gaussian (see Collier (1999), Figure 1 for a justification), the posterior distribution is not Gaussian. In this work a choice was made to put the non-linearity into the observation process rather than the system evolution process, for computational reasons. There are many methods which could be used to resolve the estimation of the posterior distribution. The method we adopt is pragmatic; the model is based on propagating the first two moments of the pdf (and justifying the applicability of this under assumptions of Gaussianity, driven by the Central Limit Theorem), thus at the update step for the $\boldsymbol{R}$ model it is only necessary to determine the first two moments of the posterior pdf.

The first moment, the mean, can be approximated by optimising (the negative log of) Equation 5 with respect to the state vector for $\boldsymbol{R}$, as is frequently done in data assimilation (Ide *et al.*, 1997). In this work the optimisation was carried out using a scaled conjugate gradients algorithm (Nabney, 2001). Having determined the most probable value, the second moment can be approximated by analytically computing the Hessian of (the negative log of) Equation 5, again with respect to the state vector for $\boldsymbol{R}$. This Hessian gives an approximation to the inverse covariance matrix for

the Gaussian centred at the mean value. This approach is justified by the belief that although the posterior distribution for $\boldsymbol{R}$ is not Gaussian, the main features of the distribution can be described by the first two moments. In future work it would be sensible to sample from this posterior distribution using Markov Chain Monte Carlo methods and assess the effect of the Gaussian assumption.

The noise on the radar observations was assumed to be given by a Gaussian on the precipitation intensity with a standard deviation of 4 $mm\ hr^{-1}$. In theory it makes more sense to have a noise model over the logarithm of the precipitation intensity, with variance given by the assumption that the root mean square factor error on the radar observations

$$\exp\left[\sqrt{\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{R_i}{I_i}\right)^2}\right]\ ,$$

is approximately 2. This is likely to be a rather poor characterisation of the magnitude and spatial structure of the errors on radar data, however at present this was felt to be the optimal estimate. The UK Meteorological Office is currently running a project to address issues of radar errors, but the results of this are not currently available.

The update (or assimilation) step for $\boldsymbol{u}$ (denoted the circled 1 in Figure 2) is more simple. However, the model is hierarchical and it is necessary to integrate over the (conditional) uncertainty in $\boldsymbol{R}$ at both times:

$$\mathrm{p}(\boldsymbol{u}_{t+1}|\boldsymbol{I}_{t+1},\boldsymbol{I}_t) = \int \mathrm{p}(\boldsymbol{c}_{t+1}|\boldsymbol{c}_t,\boldsymbol{u}_{t+1})\mathrm{p}(\boldsymbol{c}_t|\boldsymbol{I}_t)\mathrm{d}\boldsymbol{c}_t\mathrm{p}(\hat{\boldsymbol{u}}_{t+1})\mathrm{p}(\boldsymbol{c}_{t+1}|\boldsymbol{I}_{t+1})\mathrm{d}\boldsymbol{c}_{t+1}\ . \tag{6}$$

This update for $\boldsymbol{u}$ is a double Gaussian integral and can be computed analytically, again invoking the approximation that $\boldsymbol{u}$ is locally constant, to give $\boldsymbol{u}_{t+1} \sim N(\bar{\boldsymbol{u}}_{t+1}, \Sigma_{\boldsymbol{u}_{t+1}})$ where

$$\Sigma_{\boldsymbol{u}_{t+1}}^{-1} = \Sigma_{\hat{\boldsymbol{u}}_{t+1}}^{-1} + \delta t^2 \left[\Sigma_{\boldsymbol{c}_t} + \Sigma_{\boldsymbol{c}_{t+1}} + \Sigma_{\epsilon\boldsymbol{c}}\right]^{-1}\ ,$$

and

$$\bar{\boldsymbol{u}}_{t+1} = \Sigma_{\boldsymbol{u}_{t+1}} \left[\Sigma_{\hat{\boldsymbol{u}}_{t+1}}^{-1} \bar{\boldsymbol{u}}_t + \delta t \left[\Sigma_{\boldsymbol{c}_t} + \Sigma_{\boldsymbol{c}_{t+1}} + \Sigma_{\epsilon \boldsymbol{c}} \right]^{-1} (\bar{\boldsymbol{c}}_{t+1} - \bar{\boldsymbol{c}}_t) \right] .$$

Note that in this linearised update it is the inverse covariances which are added together, implying a reduction in the estimate of our uncertainty. It is also important that the uncertainties in the estimates of $\boldsymbol{c}$ at the two times is accounted for in the update, so that the model does not become over confident about the advection field without sufficient evidence in the data. As the model is run through time, at each update (assimilation) step the model probability distribution function for the state variables becomes more strongly peaked about the mean values, with the limit behaviour determined by the system and observation noises and the data distribution.

3   Generative model

The model has been completely specified, although there remain some (hyper)parameters (the system noise (co)variances) which must be defined. Ideally these parameters would be estimated from data, using a method such as Kalman smoothing. Time constraints meant this was not possible for this prototype model, thus the values for these priors were set on the basis of expert judgement. This is consistent with the Bayesian framework adopted.

The system noise covariances were estimated using arguments about what the model was, and was not, resolving. The key omission from the model is precipitation growth or decay dynamics. This could have a significant impact over a forecast time frame of 3 hours, so the model uncertainties in w and h, which would be most strongly affected by growth / decay should be quite large. It would be quite possible for an active convective system to completely decay within 3 hours, thus the system noise on h is quite large.

16

The system noise covariances can be understood in terms of the standard deviation of the added uncertainty, which in the case of h is $\log(16)$ mm over one hour. The widths, w are less likely to change as dramatically, until the system gets very weak, so these system noise on this is smaller, equivalent to adding an uncertainty having a standard deviation of $\log(2)$ km over one hour.

The uncertainty in the centres, $\boldsymbol{c}$, should be rather less since the model is capturing most of the uncertainty due to errors in the estimate of the advection field. The remaining errors come from two main sources. The first is the growth and decay of new cells, or so called daughter cell development (Pierce *et al.*, 2000) (which the model may partially capture as advection in any case) but is not explicitly accounted for. The second error concerns whether there is a single advection field for a precipitation field. Embedded precipitation, or scale dependent advection could be important, and it might be necessary to consider this in future work. At present the system noise on $\boldsymbol{c}$ is set equivalent to adding an uncertainty having a standard deviation of 1 km over one hour.

[Figure 3 about here.]

In order to assess whether the $\boldsymbol{R}$, $\boldsymbol{u}$ model is a reasonable approximation to reality, it can be very beneficial to regard the model in it's generative sense. This means that rather than trying to use data assimilation to condition the model on reality, a random realisation from the model (strictly the realisation is a sample from the prior distribution defined over the model parameters) is made and integrated forward in time. This is very simple to do when working with properly specified probabilistic models, and can produce visualisations which allow a subjective assessment of how well the model characterises the processes being modelled. Figure 3 shows a sequence

17

of synthetic radar images (with advection vectors) generated from the prior model alone (that is without using any data).

4 Results and discussion

[Figure 4 about here.]

The model was first tested on synthetic data generated from the model itself. This has the advantage that the true (generating) parameters are known and thus we can check the inference within the model is operating correctly. An example of a time series generated from the model, together with the fit to that data (the assimilated fields at the same time) is shown in Figure 4. While this sort of testing cannot be used to infer the ability of the model to forecast real precipitation events, it does provide confidence that modelling approach adopted is plausible, and consistent.

Testing the model on sequences of real radar images is more challenging, since in this case processes will be operating which are not explicitly resolved in the model, particularly processes associated with the growth and decay of precipitation. To test the model, one sequence of radar images was used, so this cannot be considered a complete verification of the model and the outcomes must be judged with care. Clearly further work on model verification is required before this can be considered an operational model.

[Figure 5 about here.]

In Figure 5 the results of running the model on a sequence of NIMROD processed (Golding, 1998) radar derived precipitation estimates is shown. The model is able to learn the advection velocities (here drawn at the centre of each basis function) as well as an approximation to the true underlying precipitation field. The representation of $\boldsymbol{R}$ is rather smooth, this being

18

largely due to the magnitude of the noise assumed on the radar images. Note that in producing these images a threshold of 0.5 mm has been selected below which the rain rate is plotted as zero.

[Figure 6 about here.]

Figure 6 shows the results of using the model to forecast the evolution of the precipitation field, with a lead time of one hour. This is a difficult case to forecast since the 30th of October 2000 was a very active situation with a deep depression (central pressure of 950 hPa over Yorkshire) and active cold front (with embedded convection) crossing the UK. Rainfall totals were between 40 and 60 mm over a wide area. Even in this rapidly developing situation the model is able to provide a plausible forecast, and the realisations from the forecast probability distribution function show that the uncertainty in the joint basis function advection process is being captured, maintaining the structure of the forecast precipitation field. While this falls a very long way short of model validation, it shows the method has potential and should be developed further.

5   Conclusions

This paper has described a fully probabilistic model for precipitation nowcasting. Wherever possible the assumptions which underly the model have been explicitly stated, so that it is possible to appreciate the drawbacks and advantages of the model. A methodology is proposed for placing ad-hoc extrapolation based nowcasting models into a consistent probabilistic framework. There is much that could be criticised about the model, and the model deficiencies might be ranked in the following order:

- The model assumes that the precipitation field is passively advected

19

over time, with no explicit growth or decay of precipitation. Thus in situations where there is daughter cell development, or significant decay of precipitation the model is likely to ascribe some of this to apparent motion. However, due to the affect of the probabilistic update for the advection field, $\boldsymbol{u}$, it is likely that most of these effects will be treated (correctly) as model errors on the positions of the centres and thus not affect $\boldsymbol{u}$.

- The assumption that the precipitation field and advection field probability distribution function can be fully characterised by the first two moments (the Gaussian assumption) is quite strong. This means that the model is not capable of resolving, for example a multi-modal probability distribution function for $\boldsymbol{R}$, however, it must be understood that the posterior distribution for $\boldsymbol{c}$ is a full covariance Gaussian, which can generate quite complex ensembles when sampled from. Further work is required to assess how restrictive this Gaussian assumption is, and could use Markov Chain Monte Carlo methods, such as particle filters (Doucet *et al.*, 2001), to undertake distribution assumption free modelling.

- In the present framework there is assumed to be a single advection field, $\boldsymbol{u}$, for all precipitation features. In the case of embedded convection, it might be expected that the advection would be scale dependent. It would be possible to extend this model to have multiple advection fields which applied (probabilistically) to different basis function sizes w, using a hierarchical approach, however assimilation in such a model would be complex to implement and might require more data than is currently available from the UK radar network.

- The hyper-parameters of the model (the system noise (co)variances) are at present specified using expert judgement. This is reasonable, since their magnitudes are probably well estimated, however the actual values might more sensibly be learnt from long time series of observations, in a manner similar to Kalman smoothing. Since it is possible that the hyper-parameters are state dependent, it could be argued that these should be estimated as part of the data assimilation process, in a hierarchical manner.

- The model only uses radar derived estimates of instantaneous precipitation rates. There are a much wider range of observations which might also be usefully assimilated in the model (Collier, 2002), including satellite observations, surface synoptic observations, real time raingauge data, even information from web cams or other visual sources. The probabilistic nature of the model means, that so long as it is possible to establish a link between the observations and the state variables, and the errors on the observations can be characterised it will be possible to assimilate the new data. For longer range forecasts it might be beneficial to consider using data from numerical weather prediction models, as applied in the NIMROD forecasting system (Golding, 1998).

- There are several numerical aspects of the model which could be improved, such as the initialisation and update of $\boldsymbol{R}$, which could be parallelised. This is important because this update takes 99 percent of the overall time it takes to run the model. Additionally the model would be more flexible were the radial basis functions to be given elliptical shape, rather than being forced to be circular. This additional flexibility in the $\boldsymbol{R}$ model would allow a precipitation field to be well

represented by a smaller number of basis functions. The use of Gaussian functions in the basis means that precipitation, as modelled by $\boldsymbol{R}$ is never zero, since the Gaussian function has infinite support. In practice a threshold is selected when the output of the model is visualised below which all values are treated as zero. The benefit of having infinite support comes during the data assimilation step for $\boldsymbol{R}$, which then always has well defined derivatives. Future models might consider representations with basis functions having finite support.

- The model assumes that $\boldsymbol{u}$ has no dynamics. This is clearly not the case, since the large scale movement of synoptic systems means that in some situations $\boldsymbol{u}$ might change quite rapidly. The large scale evolution of the atmosphere is generally well resolved and forecast in numerical weather prediction models, thus it would make sense to take advantage of this information, although a lack of information about the error structure on these numerical weather prediction forecasts would require further assumptions to be made.

- The Gaussian assumption on the radar derived precipitation estimates should be improved when better estimates of the errors become available.

Of course a key issue which remains to be resolved is whether this model adds any benefit to the operational forecaster. This issue is planned to be addressed through a rigorous verification and comparison with existing methods such as GANDOLF and NIMROD. It seems like that in many situations skill will be lost in quite a short time, but the advantage of this model is that this will be apparent, since the probability distribution function will be very broad. To assess this will require the use of methods for

probabilistic forecast verification, such as Relative Operator Characteristics (ROC) curves and reliability diagrams.

Overall, the model produces a forecast probability distribution function of $\boldsymbol{R}$ (and $\boldsymbol{u}$). From the forecast probability distribution function it is possible to tailor the product delivered to the user. If the user wants realisations or ensembles, then as many as desired can be constructed, or if the user just wants one forecast the mean can be produced. It should be noted that the realisations produced from this model will provide a sample from $\boldsymbol{R}$, which has the 'right' spatial structure: that is all the basis functions are simulated together from their joint distribution. These realisations can then processed to provide radar like images, or catchment averages, or indeed whatever the user desires since the model for $\boldsymbol{R}$ is continuous in space. Further work is required to produce an operational model.

Acknowledgements

References

Antolik, M. S. 2000. An overview of the National Weather Service's centralised statistical quantitative precipitation forecasts. *Journal of Hydrology* **239**, 306–337.

Bishop, C. M. 1995. *Neural Networks for Pattern Recognition.* Oxford: Oxford University Press.

Chatfield, C. 1996. *The Analysis of Time Series: An Introduction.* London: Chapman and Hall.

Collier, C. G. 1999. The impact of wind drift on the utility of very high spatial resolution radar data over unrban areas. *Physics and Chemistry of the Earth* **24**, 889–893.

Collier, C. G. 2000. Quantitative precipitation forecasting. *Journal of Hydrology* **239**, 1–2.

Collier, C. G. 2002. Developments in radar and remote-sensing methods for measuring and forecasting rain. *Philosophical Transactions of the Royal Society of London A* **360**, 1345–1361.

Doucet, A., N. deFreitas, and N. E. Gordon 2001. *Sequential Monte Carlo Methods in Practice.* London: Springer-Verlag.

Evensen, G. 2001. *Ocean Forecasting: Conceptual Basis and Applications (N. Pinardi and J.D. Woods eds.)*, Chapter Sequential Data Assimilation for Nonlinear Dynamics: The Ensemble Kalman Filter, pp. 101–120. Berlin: Springer-Verlag.

Fox, N. and C. G. Collier 2000. Physical dissagregation of numerical model rainfall. *Hydrology and Earth System Science* **4**, 419–424.

Georgakakos, K. P. 2000. Covariance propagation and updating in the

context of real-time radar data assimilation by quantitative precipitation forecast models. *Journal of Hydrology* **239**, 115–129.

Golding, B. 1998. Nimrod: A system for generating automatic very short range forecasts. *Meteorological Applications* **5**, 1–16.

Golding, B. W. 2000. Quantitative precipitation forecasting in the UK. *Journal of Hydrology* **239**, 286–305.

Grecu, M. and W. F. Krajewski 2000. A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting. *Journal of Hydrology* **239**, 69–84.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc 1997. Unified notation for data assimilation: Operational, sequential and variational. *Journal of the Meteorological Society of Japan* **75**, 181–189.

Krzysztofowicz, R. 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research* **35**, 2793–2750.

Kuligowski, R. J. and A. P. Barros 1999. High-resolution short-term quantitative precipitation forecasting in mountainous regions using a nested model. *Journal of Geophysical Research - Atmospheres* **104**, 31553–31564.

Mellor, D., J. Sheffield, P. E. O'Connell, and A. V. Metcalfe 2000. A stochastic space-time rainfall forecasting system for real time flow forecasting I: Development of MTB conditional rainfall scenario generator. *Hydrology and Earth System Science* **4**, 603–615.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis 1996. The new ecmwf ensemble prediction system: methodology and validation. *Quarterly Journal of the Royal Meteorological Society* **122**, 73–119.

Mylne, K. R., C. Woolcock, J. C. W. Denholm-Price, and R. J. Darvell 2002. Operational calibrated probability forecasts from the ECMWF ensemble prediction system – implementation and verification. In *Symposium on Observations, Data Assimilation, and Probabilistic Prediction.* American Meteorological Society.

Nabney, I. T. 2001. *Netlab: Algorithms for Pattern Recognition.* London: Springer.

Nabney, I. T., D. Cornford, and C. K. I. Williams 2000. Bayesian inference for wind field retrieval. *Neurocomputing Letters* **30**, 3–11.

Pierce, C. E., P. J. Hardaker, C. G. Collier, and C. M. Haggett 2000. GANDOLF: a system for generating automated nowcasts of convective precipitation. *Meteorological Applications* **7**, 341–360.

Toth, E., A. Brath, and A. Montanari 2000. Comparision of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology* **239**, 132–147.

Wheater, H. S., V. S. Isham, D. R. Cox, R. E. Chandler, A. Kakou, P. J. Northrop, I. Oh, C. Onof, and I. Rodriguez-Iturbe 2000. Spatio-temporal rainfall fields: modelling and statistical aspects. *Hydrology and Earth System Science* **4**, 581–601.

Zou, X. and Y. H. Kuo 1996. Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Monthly Weather Review* **124**, 2859–2882.
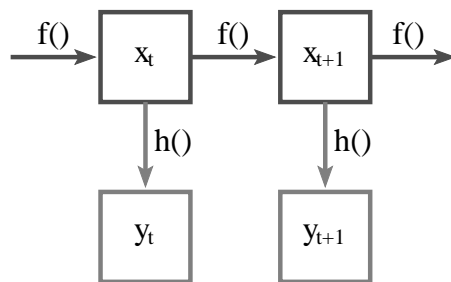
Figure 1: The state space modelling framework, showing the relation between the states over time and the observations, which are assumed conditionally independent over time.
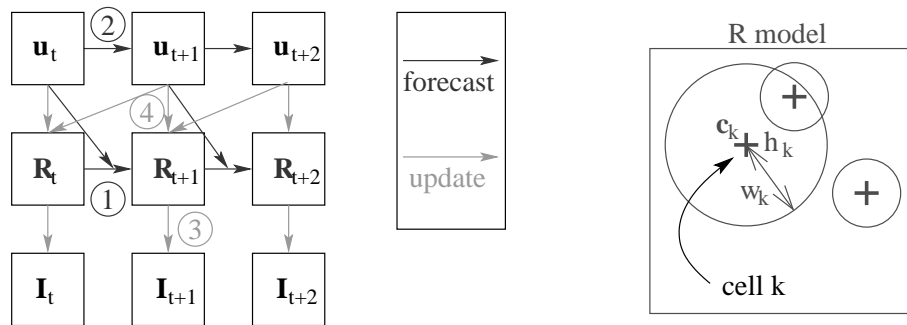
Figure 2: The model framework adopted for the probabilistic nowcasting model. The hierarchical model represents the precipitation field $\boldsymbol{R}$, which is used over two time steps to infer the advection $\boldsymbol{u}$, and generates the radar observations $\boldsymbol{I}$. The model for $\boldsymbol{R}$ is given by a radial basis function model, illustrated on the right side of the state space model.
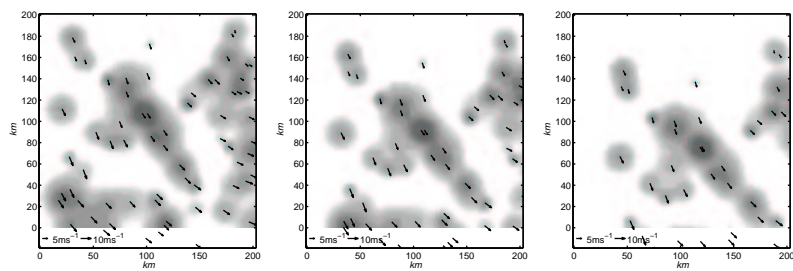
Figure 3: A series of radar images generated from one realisation of the proposed model.
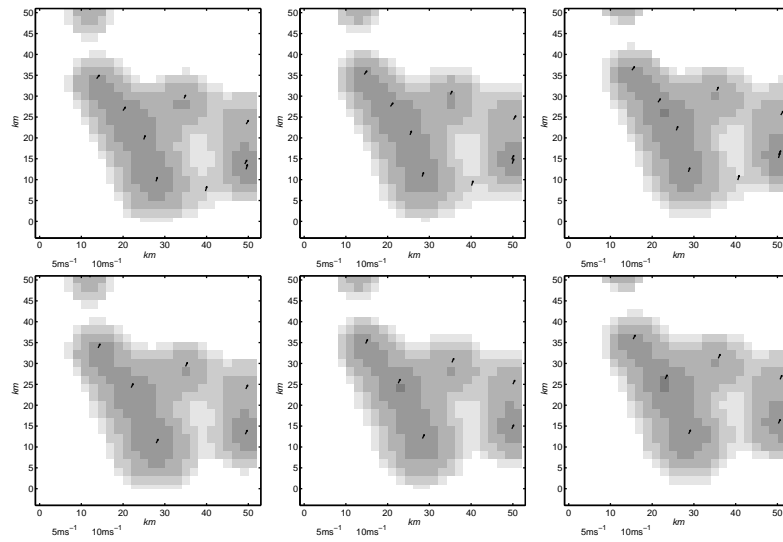
Figure 4: A series of radar images generated from one realisation of the proposed model (top), and the associated precipitation field $\boldsymbol{R}$ and advection vectors $\boldsymbol{u}$ learnt from these simulated images using the state space model framework (bottom).
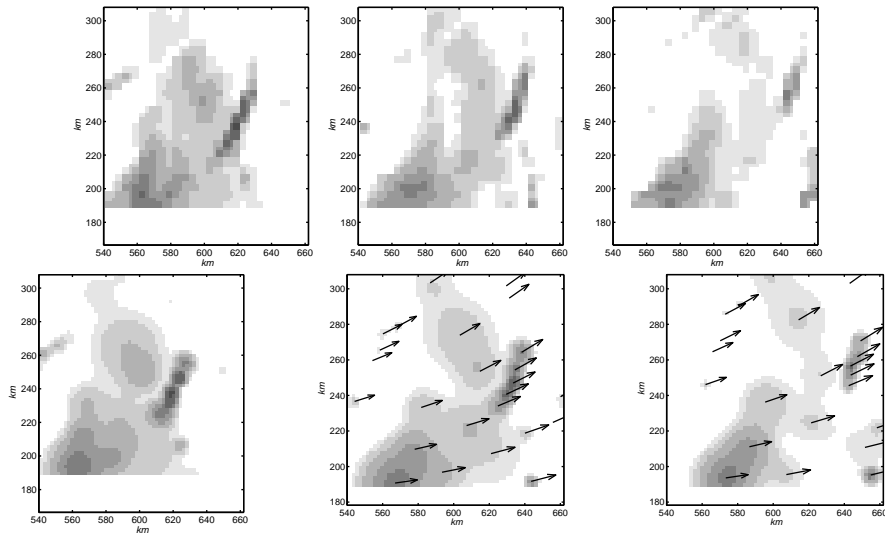
Figure 5: A series of radar images from 30th October 2000 (top), and the associated assimilated precipitation field $\boldsymbol{R}$ and advection vectors $\boldsymbol{u}$ (bottom).
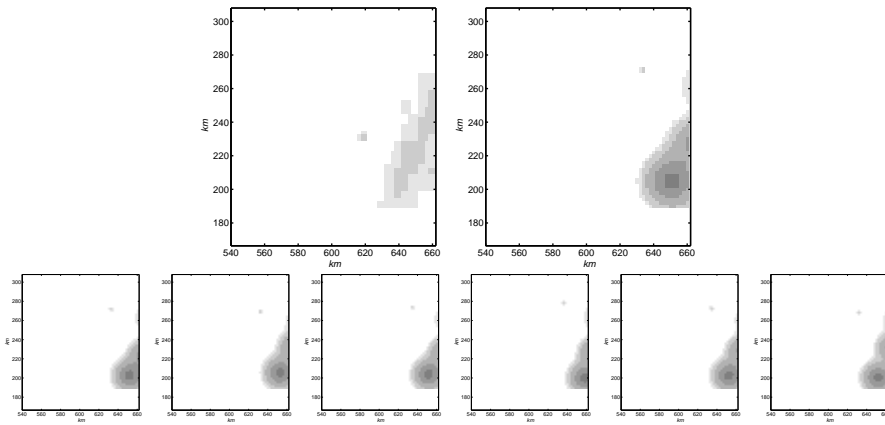
Figure 6: The verifying and one hour forecast radar images (left and right respectively) from 30th October 2000 (top), and six realisations (samples) from the forecast probability distribution function of the precipitation field R (bottom).