A stochastic space-time rainfall forecasting system for real time flow forecasting I: Development of MTB conditional rainfall scenario generator

D. Mellor, J. Sheffield, P. E. O'Connell, A. V. Metcalfe

Hydrology and Earth System Sciences, 4(4), *603–615* (2000) © EGS

Hydrology & Earth System Sciences

A stochastic space-time rainfall forecasting system for real time flow forecasting I: Development of MTB conditional rainfall scenario generator

D. Mellor[1], J. Sheffield[1], P.E. O'Connell[1] and A.V. Metcalfe[2]

[1]Water Resource Systems Research Laboratory, Department of Civil Engineering, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, UK
[2]Department of Engineering Mathematics, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, UK
e-mail for corresponding author: P.E.O'Connell@newcastle.ac.uk

Abstract

The need for the development of a method for generating an ensemble of rainfall scenarios, which are conditioned on the observed rainfall, and its place in the HYREX programme is discussed. A review of stochastic models for rainfall, and rainfall forecasting techniques, is followed by a justification for the choice of the Modified Turning Bands (MTB) model in this context. This is a stochastic model of rainfall which is continuous over space and time, and which reproduces features of real rainfall fields at four distinct scales: raincells, cluster potential regions, rainbands and the overall outline of a storm at the synoptic scale. The model can be used to produce synthetic data sets, in the same format as data from a radar. An inversion procedure for inferring a construction of the MTB model which generates a given sequence of radar images is described. This procedure is used to generate an ensemble of future rainfall scenarios which are consistent with a currently observed storm. The combination of deterministic modelling at the large scales and stochastic modelling at smaller scales, within the MTB model, makes the system particularly suitable for short-term forecasts. As the lead time increases, so too does the variability across the set of generated scenarios.

Keywords: MTB model, space-time rainfall field model, rainfall radar, HYREX, real-time flow forecasting

Introduction

BACKGROUND

The need for real-time flow forecasting systems which can provide forecasts of discharge and river level with sufficient accuracy and lead time has long been recognised, both by the research community and agencies responsible for flood warning and flood prediction. To achieve a lead time which can enable timely flood warnings to be issued and acted upon, quantitative rainfall forecasts with a spatial resolution which is compatible with that of the flow forecasting model are frequently required. Numerical weather forecasting models cannot yet provide forecasts with the required spatial resolution and accuracy (Todini, 1997), so alternative approaches must be explored. Spatial forecasts of rainfall are often obtained by making simplifying assumptions about the way rainfall fields evolve in time, e.g. Diskin (1987), Einfalt *et al.* (1990) and Bremaud and Pointin (1993). An operational system in the UK, known as FRONTIERS (Forecasting Rain Optimised using New Techniques of Interactively Enhanced Radar and Satellite data), has implemented a rainfield centroid interactive tracking technique for short-term rainfall forecasting (Interagency Research Committee on the Hydrological Use of Weather Radar, 1993). However, the forecasts produced by such systems have a large degree of uncertainty, which cannot be quantified easily but which cannot be ignored when used as the basis of flood forecasts and warnings.

A new approach to rainfall forecasting has been developed in which the Modified Turning Bands (MTB) rainfall field model (Mellor, 1996) is fitted to and conditioned upon the latest observed radar images, and then used to generate, using Monte-Carlo simulation, an ensemble of space-time rainfall forecasts which encapsulate the uncertainty about the future evolution of rainfall, given the available data and the model. This ensemble of rainfall forecasts is then converted into ensembles of flow forecasts, using the physically based SHETRAN (Parkin *et al.*, 1996) and simpler ARNO modelling systems, and the ensembles compared to assess the effects of lumping in the simpler model. This aspect of the work includes an investigation of the effects of spatial averaging of rainfall inputs to rainfall-runoff models.

All rainfall forecasting schemes provide forecasts which are uncertain, and this will contribute to uncertainty in peak flow predictions. The aims were to keep this uncertainty as low as possible and to quantify it. The method for

generating the rainfall forecasts is described in this paper. The data requirements are radar echoes as the storm approaches. Radar images have been processed to remove spurious features such as bright bands and anomalous propagation, and calibrated to give the rainfall intensity observed at points on the ground. It is assumed that the storm has been tracked for long enough to determine its overall velocity, size and statistical characteristics, and that these features do not change as the storm develops. The development, and use, of the catchment modelling systems is described in a companion paper (Mellor *et al.*, 2000).

MOTIVATION FOR A NEW FORECASTING APPROACH

The past decade has seen a major upsurge in research on rainfall estimation, modelling and forecasting on space-time scales of interest to hydrologists engaged in real-time flood forecasting. In the UK, a spatial resolution of typically 2 km by 2 km is required for application to mid-size and small catchments. Forecasts of rainfall will be made hourly during storms that present a potential flood risk, and a resolution down to 1 minute may be required for input to rainfall-runoff models. In the UK, currently, six weather radar stations provide data at 2 km spatial resolution every 5 minutes. The spectrum of research activity has been broad, and ranges from developments in the dynamic modelling of rainfall using mesoscale atmospheric models to advances in the stochastic modelling of rainfall in space and time. No attempt is made here to review the extensive literature describing this research; for this, the reader is referred to excellent reviews by Foufoula-Georgiou and Georgakakos (1991), Georgakakos and Foufoula-Georgiou (1991) and Foufoula-Georgiou and Krajewski (1995). However, some comments on contemporary research activity relevant to HYREX research in this field are included to place the latter work in context.

The classical deterministic approach to weather/rainfall forecasting is through a dynamic numerical model based on a set of partial differential equations describing the conservation of mass, momentum and energy in the atmosphere. Applied on a global scale, such models do not have anything like the required spatial resolution for hydrological (and other) applications, and so much recent effort has been devoted to the development, testing and comparison of limited area models (LAMs). Typically, these models are applied to mesoscale areas and derive their time-dependent boundary conditions from a previously integrated global atmospheric model. Examples of such models are the Colorado State University – Regional Atmospheric Modelling System (CSU-RAMS), the National Centre for Atmospheric Research – Mesoscale Model Version 5 (NCR-NM5) model, and the University of Bologna limited area model (LAMBO) which is based on the University of Belgrade/NOAA-National Meteorological Centre of Washington model. Typically such models need convective parameterisation schemes which are the subject of continuous research efforts (see Molinari and Dudek, 1992 for a review), since they exercise a major influence on the predicted timing, location and amount of precipitation.

Although the ability of such models to produce accurate simulations of rainfall on the scales of interest is limited, several applications of such models to a variety of hydrometeorological problems have been reported in the literature (Foufoula-Georgiou and Krajewski, 1995). Moreover, major advances in workstation technology and communication systems enhance the prospects for real-time application to quantitative precipitation forecasting (Cotton *et al.*, 1984). However, as observed by Foufoula-Georgiou and Krajewski (1995), the performance of such models in this role remains largely unknown.

In the absence of a well proven, operationally feasible, LAM modelling approach at the present time, alternatives are needed to meet the current call to extend the lead time of flood and flash-flood forecasts. One interesting approach in this regard is to couple dynamic hydrological and meteorological model components at the catchment scale within a state estimation framework (Georgakakos, 1986a, b, 1987). Such models may be classified as stochastic dynamic models where the conservation of mass couples the two model components through both the dynamic model equations and state estimator feedback. The importance of coupling and the worth of various types of hydrometeorological data in flood forecasting have been demonstrated as a function of the ratio of forecast lead time to basin response time (Georgakakos and Foufoula-Georgiou, 1991). Improvements in short term quantitative precipitation forecasting using an enhanced stochastic dynamic model are discussed by Lee and Georgakakos (1992).

The stochastic modelling of rainfall in space and time has also seen important developments within the past decade. Early work on the use of point process models to describe the temporal structure of rainfall at a point (e.g. Rodriguez-Iturbe *et al.*, 1987) has given way to scaling approaches which seek to describe the spatial statistical structure of rainfall over a wide variety of scales with relatively few parameters. Scale invariance implies that small and large scale statistical properties are related to each other by a scale changing operator involving only the scale ratio. Developments in this field are reviewed by Foufoula-Georgiou and Krajewski (1995) who note that the current state-of-the-art scale invariant rainfall models revolve around multiplicative cascades which have their origin in the statistical theory of turbulence e.g. Tessier *et al.* (1994). Based on this approach various methods of parameter estimation and multifractal field analysis and simulation have been developed (e.g. Gupta and Waymire, 1993; Over and Gupta, 1994; Kumar and Foufoula-Georgiou, 1993a, b). However, such models are still in the realm of theoretical development and their practical application would appear to be some way off.

Given the current state-of-the-art of stochastic model-

ling, one of the questions which this HYREX project sought to explore is whether or not an ensemble of scenario forecasts from such a model might prove useful in extending the lead time of flood forecasts from a rainfall runoff model. Existing empirical rainfall forecasting procedures are often very crude (e.g. assuming that future rainfall will replicate that observed in recent time periods). They may also involve the selection of "analogue" rainfall sequences from observed past events, but it is never clear which part of the past event should be used as an acceptable 'forecast' of the current event (Todini, 1996). They are sometimes based on statistics extracted from past data (e.g. Schultz, 1994) but, in this case, the quantile curves do not display anything like the temporal variability of the actual rainfall.

The most straightforward and well developed area of stochastic rainfall modelling is for temporal rainfall at a point (e.g. Cowpertwait, 1994): this univariate modelling approach should also be applicable to rainfall averaged over an area. Rainfall runoff models with spatially averaged rainfall inputs perform well in many real-time flood forecasting systems, providing some justification for a univariate approach. As catchment area increases, the assumption of spatial averaging of rainfall becomes more questionable, and so it may be necessary to employ a multivariate approach in which rainfall is modelled at several points (or for several sub-catchments), which would involve reproducing the cross-correlation as well as the temporal structure of point rainfall (e.g. Sansó and Guenni, 1999). Alternatively, a multi-dimensional rainfall field model could be employed to characterize rainfall at all points in the catchment domain, not just those where measurements exist. However, such models are not easy to parameterise, and the necessary data are frequently not available. The availability of radar data within HYREX has enabled a rainfall field model to be employed for ensemble forecasting.

OPERATIONAL RADAR RAINFALL FORECASTING

The ability of radar to portray the spatial and temporal variation in rainfall estimates is significant for real-time flood forecasting. Radar measurements are relevant to both the qualitative and quantitative aspects of operational rainfall forecasting. In qualitative terms, the advantages of radar measurements are particularly apparent in detecting the convective storms which may be undetected by a conventional raingauge network and for frontal storms whose movement can be readily depicted by replaying radar pictures at successive time frames (Haggett *et al.*, 1991). Both are useful in an operational flood warning system. In quantitative terms, the radar measurements are relevant mainly for two reasons: (a) for estimating rainfall inputs to a catchment within an operational forecasting system; and (b) for extending the lead time of the forecasts through rainfall forecasting. Flood forecasting with a catchment model forms an integral feature of an operational flood-warning system. This is particularly important for a flash flood generated from an urban or mountainous catchment.

Spatial forecasts of rainfall are often obtained by making simplifying assumptions about the way rainfall fields evolve in time, for example, by assuming that the field appears stationary to a moving observer e.g. Diskin (1987), Einfalt *et al.* (1990), Bremaud and Pointin (1993) and Abdullah (1996). A forecaster may attempt to enhance these assumptions by identifying features of the rainfall field which evolve more predictably, and allowing for these changes e.g. Seo and Smith (1992). One of the principal uses of the FRONTIERS system (Browning, 1979, 1981) is for hydrological forecasting. The FRONTIERS forecast is produced using a three stage process, where the forecaster checks the radar data for quality control, uses the Meteosat satellite data to derive likely areas of precipitation beyond the radar coverage, and produces a forecast for up to six hours ahead (although forecasts for six hours ahead are not necessarily reliable, Einfalt and Semke, 1997). These stages are performed during a half-hourly cycle. During the first stage of the process, the forecaster interactively assembles a composite picture of the radar images from a radar network. This permits the possibility of filling missing data with data from an adjacent radar. The satellite image, where the rainfall areas are defined as "wet" or "dry", is merged with the composite radar image for the rainfall extrapolation in the forecasting step. The velocities of radar features (i.e. clusters of echoes, representing raincells) are determined semi-automatically, on two radar images separated by one or two hours, and the forecast is performed by linearly extrapolating the features using their estimated velocities. However, it is difficult for the forecaster to track several clusters simultaneously because of human limitations. Another development is the Complex Method (e.g. Duda and Blackmer, 1972; Elvander, 1976; Einfalt *et al.*, 1990). This utilizes an advanced pattern recognition technique to process the radar echo images and track the motions of individual echoes. For instance, the description of echo geometry using a Fourier analysis technique provides the locations of the edges, which may be matched at separate times to give the displacement of the echo. Similar analyses using statistical echo distribution (e.g. intensity distribution of an echo) may also be used for echo matching (Collier, 1991).

Research and Development to automate FRONTIERS forecasts and to develop improved procedures using artificial intelligence techniques continues. This work includes investigation of improvements to radar measurements of precipitation, including the use of estimates of the vertical reflectivity profile, bright-band correction procedures and rain gauge representativity, and improvements to the estimation of rainfall from visible and infrared satellite data (Interagency Research Committee on the Hydrological Use of Weather Radar, 1993). NIMROD (Nowcasting and Initialization for Modeling using Regional Observation

Data) is a fully automated system which has now replaced FRONTIERS at the Met Office. Cranston (2000), for example, has compared NIMROD rainfall radar observations with recordings at raingauges for an upland site in Southwest Perthshire, Scotland. NIMROD works better with frontal systems than during convective events. GANDOLF (Generating Automated Nowcasts for Deployment in Operational Land-based Flood forecasts), based around an object-oriented model of the life cycle of a convective cell (Hand, 1996), has been developed at the Met Office as a means of improving forecasts of heavy convective precipitation.

Operational radar rainfall forecasting within flood-warning systems is not widespread and is confined mainly to urban drainage flow forecasting (Huff *et al.*, 1981; Bellon and Austin, 1978; Einfalt *et al.*, 1990). The main problem is that the lead time for reliable forecasting is typically very short (of the order of one hour). The World Meterological Organization (WMO) has instigated forecast demonstration projects (FDP) under its World Weather Research Programme. The first of these is the Sydney 2000 Olympics FDP (e.g. WMO, 2000).

In this paper, a different approach to forecasting rainfall is described. It involves the fitting of a sophisticated stochastic model of the rainfall process, in such a way that the model will produce scenarios of rainfall fields which are consistent with current observations, i.e. it will produce possible future scenarios conditioned on the observations. The model chosen for this purpose, the Modified Turning Bands (MTB) model (Mellor, 1993, 1996), has both deterministic and stochastic features, making it ideal for this application as the range of future scenarios generated can be constrained by an experienced operator through the deterministic aspects of the model. Natural variability in the rainfall process, and uncertainty in the future evolution, is represented by the stochastic details. This model operates at four distinct scales of space and time, permitting different advection velocities and rates of evolution at each scale. Hence, forecasts can be produced with lead times from a few minutes to several hours. The range of variability across the scenarios is expected to increase with the lead time.

The MTB model is conditioned on the most recent radar image available, which is reproduced at the beginning of each generated scenario. However, the preceding image is also taken into consideration to determine how the fitted raincells should be distributed in time in order that the evolution of the rainfall field between the two images can be carried forward into future scenarios. This allows for an accurate reproduction of the current rainfall field which respects the ageing process of raincells, i.e. which places raincells in the growing or decaying parts of their life cycles to account for the most recent changes in the radar images. As the scenarios develop, the structure of the MTB model is used to generate new raincells as existing raincells die away. Thus, the raincells fitted to the observed rainfall field are replaced, gradually, with stochastically generated ones.

The MTB model

DEFINITION OF THE MTB RAINFALL FIELD MODEL

The Modified Turning Bands (MTB) stochastic rainfall field model was developed originally to simulate the spatial and temporal distribution of rainfall in frontal storms (Mellor, 1996). The main structural features of frontal rainfall (raincells, cluster potential regions and rainbands) are represented by the MTB system. The construction of the rainfall field is begun by placing two sets of parallel parabolic prisms, in the $x - y$ domain (Fig. 1, level 1). The two sets are at different angles, e.g. $\pm 30°$ to the storm velocity vector, and the areas of intersection will correspond to higher cluster potential regions. The prisms are distributed randomly and independently according to a Poisson process. Once this initial distribution, corresponding to time zero, is established, the prisms are moved along the lines. The velocities are approximately opposite, so that when the prisms are added the high zones tend to move in a direction which is parallel to the y-axis. The speeds at which the prisms move along the two lines are independent.

A modulating function, which travels in the direction of the storm, is then applied to the aggregated parabolic prisms to give a potential field function for raincells. The modulating function is the sum of two sinusoids of slightly different frequencies, which produce a 'beating' effect, and a

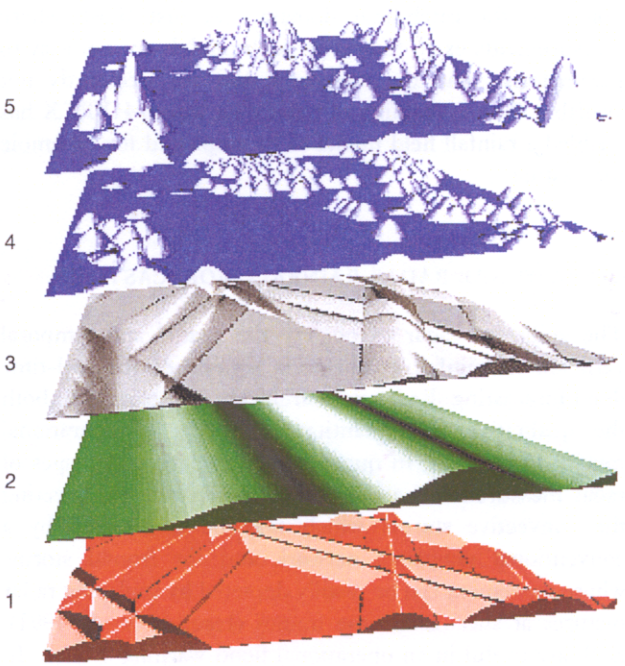

Fig. 1. *Construction of the MTB model, showing from bottom to top, the placing of two sets of parallel parabolic prisms, the storm/rainband modulating function, the resulting potential function of raincells, a distribution of raincells, and the final rainfall field obtained on summing the contributions of all the raincells.*

constant to make it non-negative. The modulating function is set at zero outside a fixed number of beats, which represent the width of the storm, and slides along a line in the direction of the storm with some constant velocity (Fig. 1, level 2). The sinusoids that make up the beat have a different velocity to that of the modulating function and this produces a rainband effect inside the storm.

The field obtained by adding the parabolic prisms and multiplying the sum by the modulating function represents the potential function for raincells in the plane, as shown in the third level of Fig. 1. That is, an inhomogeneous Poisson process is realised, over time, in the two spatial dimensions of the catchment, with the non-stationary rate function given by the raincell potential field. Raincells are born at the point occurrences of this process.

The raincells themselves are described as inverted parabolas of revolution, with a peak that grows and decays quadratically in time. This function is truncated to zero where it would otherwise be negative, giving raincells with a finite lifetime and spatial extension, as depicted in Fig. 2. If g is the rainfall contributed by a raincell born at the origin to a point (x, y) on the ground at time t, then g takes the form

$$g(x,y,t) = h(x - c_x t, y - c_y t, t) \qquad (1)$$

where

$$h(x,y,t) = \frac{4M}{d^2 w^2}(w^2 - 4x^2 - 4y^2)\ t\ (d - t)$$
$$\text{when } x^2 + y^2 < \frac{w^2}{4} \text{ and } 0 < t < d, \qquad (2)$$
$$\text{and 0 otherwise.}$$

M is the maximum rainfall intensity that can occur inside a raincell (at the centre, half-way through the lifetime), d is the lifetime (duration), w is the width (diameter), and (c_x, c_y) is the velocity. The MTB model assumes that all raincells in any storm have the same set of parameters.

Once the raincells are distributed on the field of points of the inhomogeneous Poisson process as indicated in the fourth layer of Fig. 1, they are summed in the plane to produce the final rainfall intensity field as indicated in the top layer of the figure. The final model exhibits features in the synthetic rainfall fields resembling raincells, cluster potential regions and rainbands, all of which are observed to a greater or lesser extent in real radar data.

The final rainfall field displayed in the top layer of Fig. 1 can be sampled at a set of points on a square grid, and at regular time intervals. For example, if the overall dimensions of the area are 420 by 420 kilometres, and the rainfall field is sampled on a 5 kilometre grid at five minute time intervals, then data in the same format as MO radar can be synthesised. For greater realism, the field can be sampled at a higher resolution in space and then averaged over each radar grid square. Also, the data could be sampled in a radial pattern in time to simulate the action of the radar.

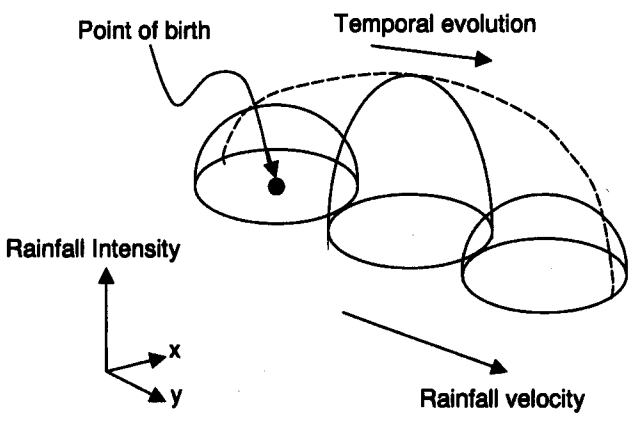

Fig. 2. *Schematic showing the position of a raincell at three stages of its lifetime.*

TECHNIQUES FOR PARAMETER ESTIMATION

The MTB model parameters are estimated by focusing on structural features of the rainfall field at different scales. Raincell parameters are estimated using the (non-stationary) covariance structure of the MTB model (Mellor, 1993, 1996), and the technique of Full Correlation Analysis. Extensive Monte-Carlo simulations have been used to verify these and to determine confidence intervals (Mellor and O'Connell, 1996). Methods have also been developed for the estimation of the parameters of the modulating function under the assumption that storms are without curvature (Mellor and Metcalfe, 1996).

Production of a forecast from radar data

OVERVIEW OF THE FORECASTING PROCEDURE

The primary data requirement for the forecasting process is enough radar echoes for the overall velocity and size of the incoming storm to be estimated. The model described earlier synthesizes storms which move strictly from west to east and are oriented in a straight north-south direction. Therefore, the next stage is to transform the real storm to such an idealized structure. This is achieved using a graphical user interface to the MTB system, developed for the HYREX project; this allows the user to fit frames by eye and thereby identify the curvature, angular rotation and translational velocity of the storm, as well as the number of rainbands inside and their position relative to the overall storm profile. The storm signal is passed through a non-linear coordinate transformation, so that the rainbands are constrained by parallel lines, perpendicular to the direction of the storm and the angular rotation is removed. It is then possible to estimate the parameters of the MTB model, including those of the distribution of raincells so that the small-scale statistics and structure of the rainfall field are

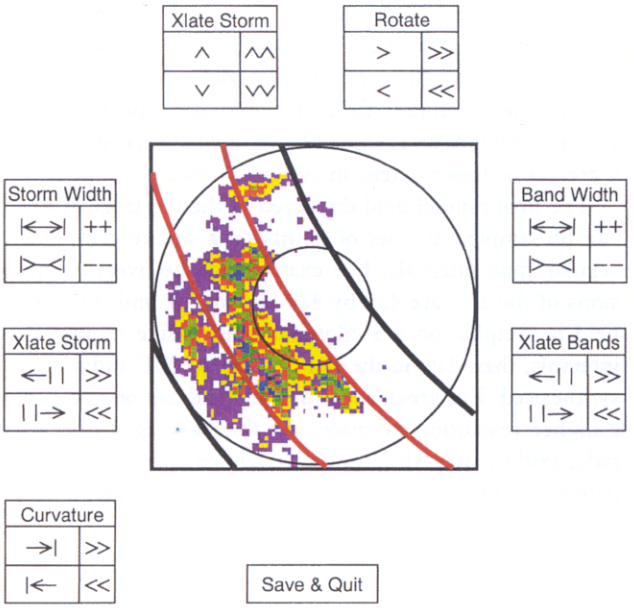

Fig. 3. *Computer screenshot showing the graphical software used by the operator to fit a set of frames to an observed radar image.*

reproduced in simulations. An outline of the forecasting strategy follows.

The data are first transformed to remove the identified curvature and angular rotation, and the parameters of the model are then estimated. The next step is to determine a distribution of raincells, with the property that, when their contributions are added, the currently observed rainfall field is reproduced. Once such an arrangement of raincells has been established, a distribution of cluster potential regions and rainbands is inferred such that, when the underlying raincell birth-rate function is reconstructed, the distribution of the fitted raincells is explained.

Since the underlying birth-rate function of raincells is treated as a deterministic process, it can be calculated for the future rainfall field based on the fitted function at time zero. Using this as a probability distribution function for future raincell births, a random number generator is used to produce a field of future raincells and, hence, generate a stochastic future rainfall field. Finally, the transformation which was applied to straighten the original data is applied in reverse to the generated data, so that the original curvature, angular and translational velocity of the storm are carried through the future scenarios.

DETERMINING THE CURVATIVE, ANGULAR ROTATION AND TRANSLATIONAL VELOCITY OF RAINSTORMS

Figure 3 is a screenshot of the workstation at which the user sits, showing a radar image superimposed by a frame of equi-distant curves which the operator has fitted to rainbands by eye. The frame may exhibit arbitrary orientation and curvature, and these attributes and the location of the centre of the storm can change in time, i.e. they can vary from one frame of the radar image to another. The system uses this information to determine the translational and rotational velocities of the storm, and the rate of change of curvature if any; this derived information is subsequently used to 'straighten' the storm out so that the MTB model can be fitted, and to compute a set of extrapolated frames on the basis that the shape of the storm will continue to evolve with the currently observed rates of change of the parameters (orientation, curvature, location). However, this assumption could be relaxed if there is reason to suppose some alternative trend in these parameters is more realistic. Subsequently generated future rainfall fields are inserted into these extrapolated frames to produce the final forecasted rainfall field. The effect of straightening the images in the frames, so that the storm moves in the west to east (positive x-direction) is shown in Fig. 4. The data are now in a form for estimation of the MTB model parameters.

ESTIMATION OF MODEL PARAMETERS

The parameters of the MTB model fall into three groups according to the space-time scale at which they apply. Each group has its own methods for parameter estimation.

At the smallest scales, the parameters of the raincells are estimated using the methods given in Mellor and O'Connell (1996). The raincell lifetime, width and velocity are estimated through the use of a technique known as full correlation analysis, which identifies features in the correlation structure related directly to these aspects of the raincells. Internal covariance analysis is used to determine the density of raincells and their maximum intensity, by equating the theoretical covariance structure with the empirical covariance of observed rainfall.

The medium scale variation is modelled by the cluster potential regions. The associated parameters are: the angles between the lines and the x-axis; the width of the parabolic prisms; the velocities of the prisms; and the rate of occurrence of prism sections along the lines. The line processes do not correspond to physical features; they are just a device to make cluster potential regions move in an approximately orthogonal direction to the leading edge of the storm. The associated parameters have little effect on the statistical features of storms, and it is not practical to estimate specific values for each storm from rainfall data (Mellor and Metcalfe, 1996). The following fixed values have been found to be consistent with meteorological observation and lead to visually realistic synthetic storms. The angles between the x-axis and the two lines along which the prisms slide are plus 30° (A-line) and minus 30° (B-line). The angle between the two lines governs the eccentricity of the elliptical cluster potential regions. The velocities of the prisms along the A and B lines are 20 km h^{-1} and -28 km h^{-1} respectively. The rate of occurrence of

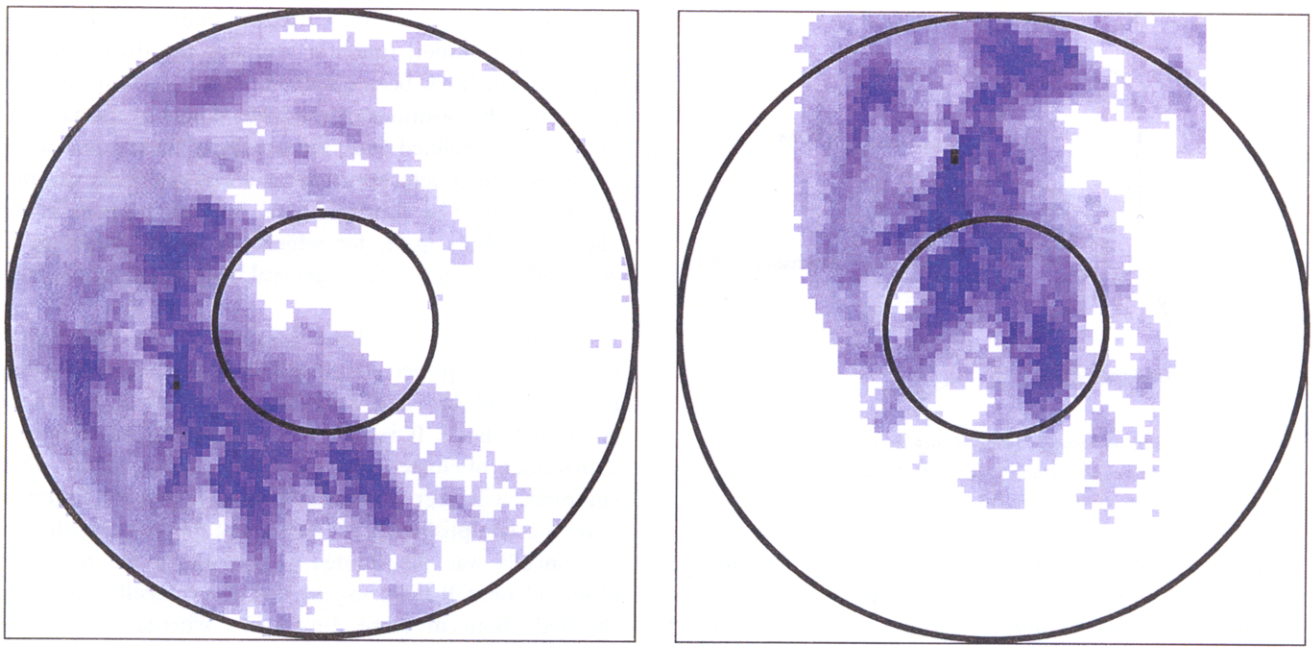

Fig. 4. *Real radar image, diameter 420 km, (left) at five km spatial resolution and the same data after being straightened out (right).*

prisms and their width are 0.03 km^{-1} and 20 km respectively.

The modulating function (second level of Fig. 1) governs the large scale aspects of the model. The parameters of these features are determined at the same time as the operator identifies the curvature and angular motion of the storm. The operator first places curves that delimit the edges of the storm in the original observed radar data. Then curves are placed to enclose the rainbands within the storm. The gap between any pair of curves in each frame is constant and the computer software adjusts for this automatically. However, the position of the lines within the frame relies on the operator's judgement. Once the data have been straightened, the sizes and velocities of the components of the modulating function can be estimated.

PRODUCTION OF A FIELD OF FITTED RAINCELLS

The fitting of stochastic models to data usually entails the determination of model parameters so that important statistical properties of the data (for example, the mean, variance, correlation structure, proportion of dry days) are matched by the model. However, in this case the conditioning is much stronger because the model matches not only the statistical aspects of the rainfall fields, and thus implicitly the geometrical and dynamic structure as well, but also duplicates the observed parts of the rainfall field. Hence, the term 'strong conditioning' is used here.

The strong conditioning of the MTB model starts with the determination of a field of raincells which reproduces the observed data. The lifetime, widths and velocity of raincells have already been estimated, as described earlier. Each raincell has three degrees of freedom, its location and time when it reaches its peak intensity which occurs at mid-life. The strategy adopted is to fit a raincell so that the peak intensity, in both space and time, coincides with the largest observation in the data, as indicated schematically in Fig. 5. This may cause the raincell to deliver more rainfall than is actually observed at this point, or other observation points in its domain of influence. If this happens, the raincell is moved backwards in time, that is, the time of birth of the raincell is decreased, (see Fig. 6) until it no longer exceeds any of the observed data. In practice, the raincell is shuffled about with decreasing increments, until a fit is found for which the raincell contributes as much as possible to the observed peak without exceeding the rainfall at any other observation point. Actually, the line of the raincell's evolution in space-time makes an angle to the plane of the radar image because of the raincell's velocity, but the

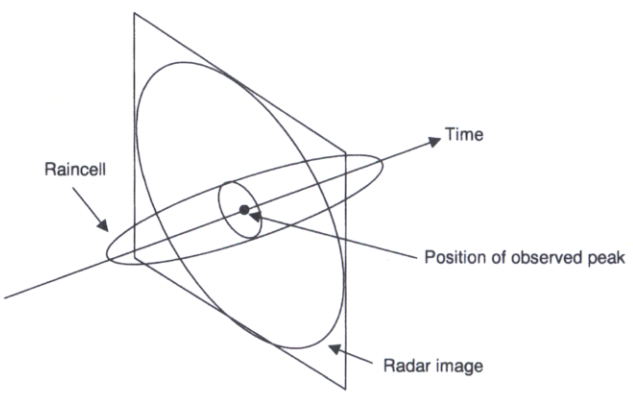

Fig. 5. *Schematic showing a raincell intersecting a radar snapshot, indicating the initial placement of the raincell.*

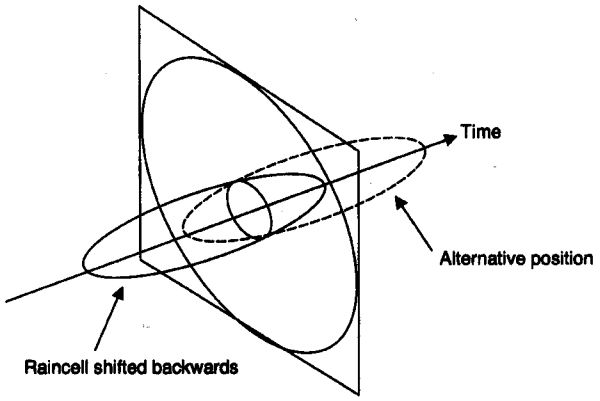

Fig. 6. *Schematic showing a raincell displaced backwards in time to make a diminished contribution to the radar image, and an alternative placing of the raincell which makes the same contribution.*

principle of the method can still be applied. Once a satisfactory fit of the raincell is made in this way, its contribution to the straightened radar image is subtracted from the data, and the process is repeated with the new peak, until all the rainfall data are explained by a collection of raincells. The algorithm has been coded in C++, as a subroutine for the main computer program which implements this work.

At this stage, the most recent straightened radar image will be explained completely by a set of raincells. However, since these have all been displaced backwards in time, they will all be in the decaying phase of their lifetime when they make their contribution to the radar data. A consequence of the structure of the MTB model, is that, at any time, half of the raincells are expected to be in the decaying phase of their life, and the other half in the growing phase. Furthermore, the raincells in the MTB model, described earlier, are symmetrical about the point in space and time where they attain their highest intensity. It follows that a raincell fitted to the most recent radar image can be 'flipped' about its intersection with the radar plane, so that it makes exactly the same contribution to the radar image but is now in the growing phase (Fig. 6).

To decide which of the fitted raincells should be flipped in this way, the radar image immediately before the current one is considered. The set of fitted raincells is scanned to find those that make the biggest unwanted contribution to the previous radar echo, and these are flipped forwards so that their contribution to the previous echo is decreased.

DETERMINATION OF THE UNDERLYING RAINCELL BIRTH-RATE FUNCTION AND PRODUCTION OF A SINGLE STOCHASTIC FORECAST (FUTURE SCENARIO)

Having determined a distribution of raincells which gives rise to a synthetic rainfall field identical to the one observed, a distribution of cluster functions is determined along the inclined lines which, when used to build the underlying raincell distribution function, explains the distribution of raincells which has been obtained. This is accomplished by projecting the positions of the birth points of the raincells onto the two inclined lines; the density of these points along the lines is then determined using a kernel density estimating function. The function used for this purpose is the cluster parabola itself, for which the width parameter is assumed to be known from general meteorological observations.

It follows that the best estimate of the distribution function of the projected raincell birth points along the line is just the observed density function of these points, as computed above. It is therefore necessary to determine a distribution of cluster parabolas along the line which, when summed, will reproduce this observed density function. This is accomplished in the same way that the distribution of raincells was determined to reproduce the originally observed rainfall field, except that the raincells had to be shuffled about in three dimensions whereas the cluster functions only need to be fitted in one dimension.

At this stage, all the details of the MTB model explain all the observed features of the rainfall field. However, these observed aspects of the model will not specify completely the model at points away from the observations, i.e. at future times, where the fitted features will be either too sparse or non-existent, and will not therefore reproduce the structure of the MTB model properly. Thus, it is conjectured that extra details of the model must be added to those fitted above, and this is achieved using the stochastic generation mode of the model. Specifically, cluster parabolas are distributed randomly on the parts of the inclined lines which currently project outside of the observed data field (since the cluster parabolas slide along the lines with some velocity, to produce a rainfall field at future times it is necessary to construct details along the inclined lines which will enter the field of observations only at later times).

Once a complete distribution of cluster parabolas has been obtained, the details on the inclined lines can be projected back into the field of fitted raincells and hence a complete raincell birth-rate function can be reconstituted in two-dimensional space and time. This is then used to scatter raincells randomly in the unobserved parts of the raincell field (i.e. at future times) to augment the raincells which have been fitted to the observed parts; this completes a realisation of the MTB model which is a combination of the deterministically fitted and the stochastically generated components. None of these extra raincells is allowed to contribute to parts of the space that have been observed, so the fitted raincells at these points continue to reproduce the observations exactly, but the extra details extrapolate the field into the future in a consistent way.

The final forecasted rainfall field is obtained by summing the complete set of raincells, which will reproduce all the original observations implicitly, as well as producing an extrapolated rainfall scenario.

PRODUCTION OF AN ENSEMBLE OF STOCHASTIC FORECAST

Once the model has been used to produce a future scenario of the straightened storm, the transformation is reversed so that the curvature and orientation, apparent in the original data, are preserved. Furthermore, the process of generating future raincells can be repeated as often as required without having to refit the raincells which explain the observed data. Thus, it is an easy matter to produce an ensemble of future rainfall scenarios, once the outline of the original storm has been identified and the MTB model has been conditioned.

In Fig. 7, the top left panel shows radar frames, taken at five minute intervals during the storm of 3rd February 1994, following the field of observations shown in Fig. 3, which also appears at the top left of the panel. The other three panels show radar frames taken from MTB forecast scenarios following the same initial field of observations. All three of the scenarios start out by reproducing the last observed radar frame (after low-intensity (<0.1 mm hr^{-1}) background rainfall has been removed), and the simulations become increasingly different as the lead time increases. This is the result of the original field of fitted raincells dying out gradually and being replaced by different fields of stochastically generated raincells in each scenario. It should be noted, though, that many properties of the original field are preserved in all the scenarios: the overall shape and movement of the storm, the number and relative positions of the rainbands, and the fitted attributes of the MTB model are maintained throughout each scenario (raincell velocity, width and intensity; mean, variance and lag-one internal spatial covariance structure; important peaks in the spatial correlation structure implied by explicitly modelled features, e.g. raincells, cluster potential regions). As scenarios can be generated quickly an ensemble of at least ten scenarios would usually be presented, as shown in Fig. 9.

Case study

STORM DESCRIPTION

The River Brue catchment, relative to the range of the Cobbacombe radar, is shown in Fig. 8. Detailed radar images are available throughout a frontal storm on 3rd February 1994, and rainfall in the Brue catchment was monitored by the network of gauges. This storm was used as a test case. Forecasts of total rainfall in the Brue catchment

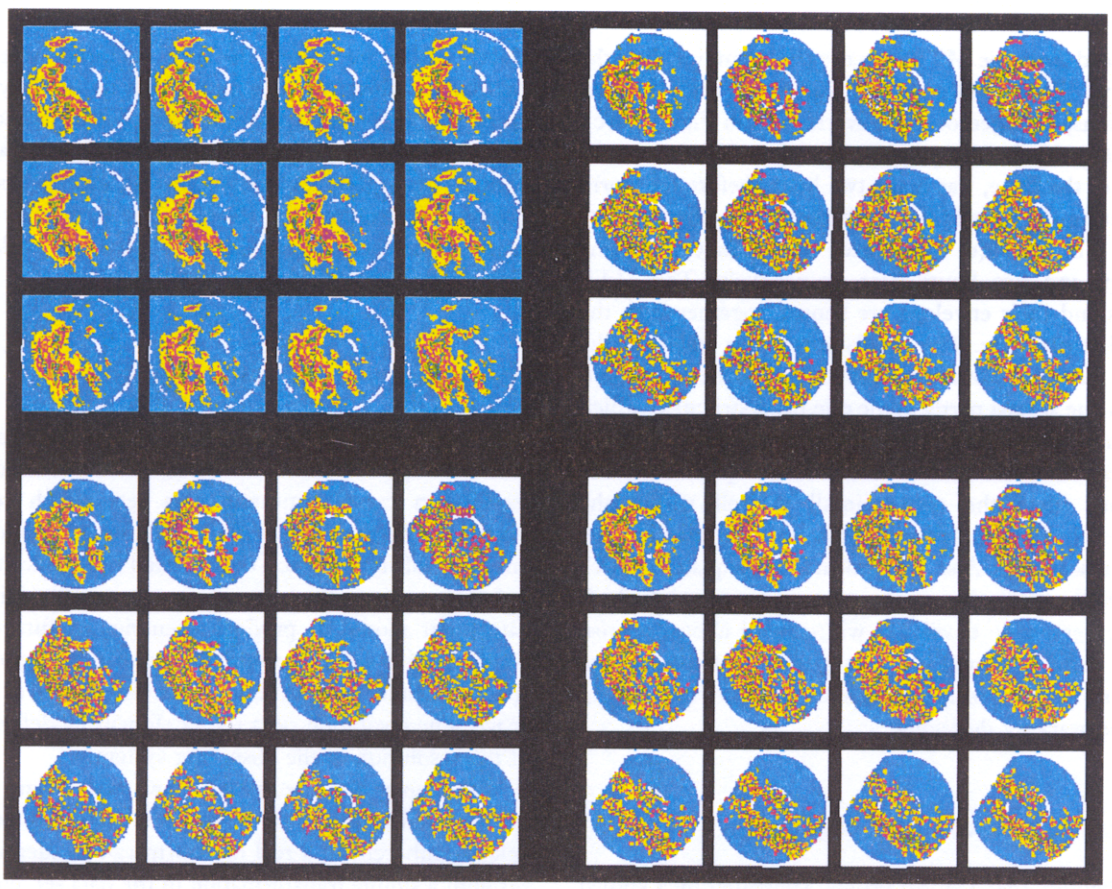

Fig. 7. *Original real radar sequence from storm of 3rd February 1994, frames at 5 minute intervals, (top left panel) and three scenarios generated by the MTB model from the same initial frame (top left frame in all four panels).*

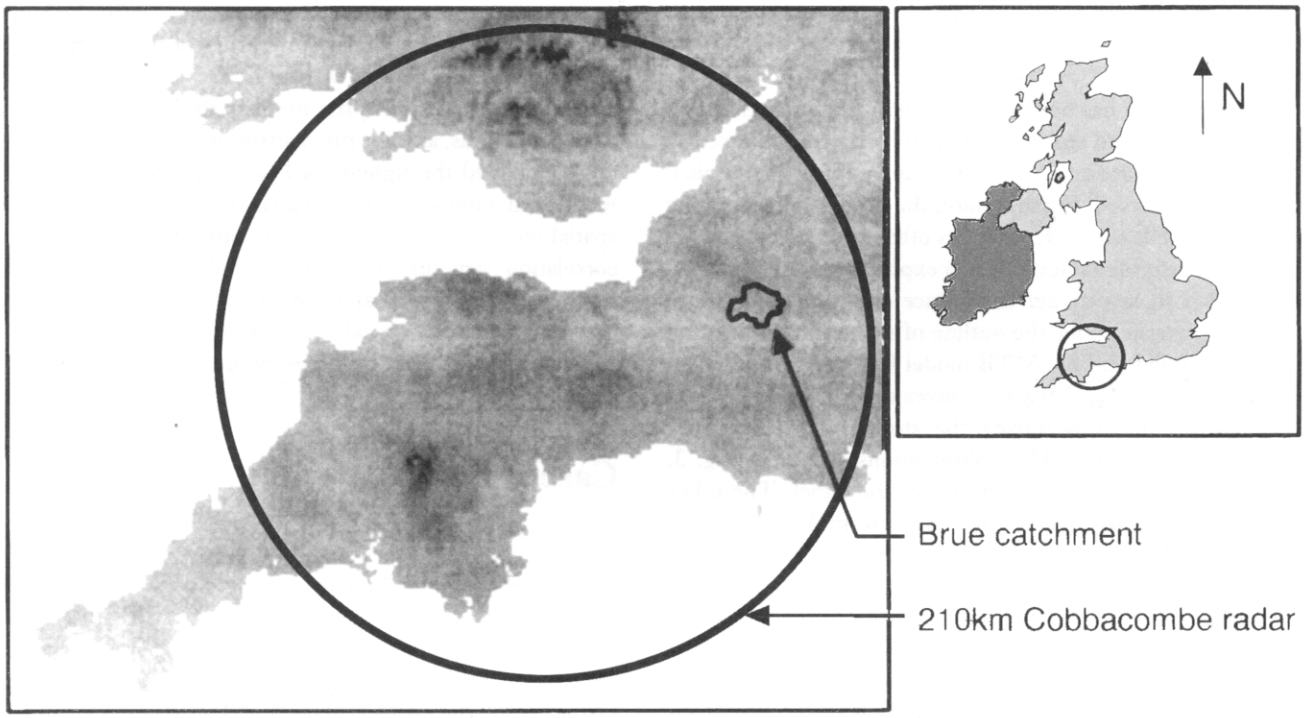

Fig. 8. *Location of the Brue test catchment shown relative to the extent of the Cobbacombe radar range.*

are deduced from the MTB scenarios by summing over the catchment. These forecasts are compared with ground observations, averaged over the catchment, in Fig. 9(a)–(c). In Fig. 9(a) the forecast, for the complete duration of the storm over the Brue, was made two hours before the storm front first hit the edge of the catchment, while the forecasts in Figs. 9(b) and 9(c) were made at the time it hit the catchment and two hours later, respectively. Ten MTB scenarios and their envelope are shown, together with the observed rainfall rate. The envelope is rather wider for the forecasts made two hours after the storm hits the catchment. This was due at least partly to the rain band structure, which had been defined clearly in the radar images at the earlier times, becoming less pronounced. There is generally good agreement, since the total rainfall is well within the ensemble envelope – although this is also true for the individual scenarios. Also, one storm is inadequate to check the reliability and value of the scenarios. A model which produces very wide envelopes will contain almost all real storms. Useful envelopes will usually turn out to contain the actual storm events, whilst being sufficiently narrow to help with flood warning decisions.

ASSESSMENT OF ENSEMBLE FORECASTS

The probability that a storm profile, generated by a known stochastic mechanism, will be outside an envelope of n storms generated by the same mechanism in at least one place, could be found to any required accuracy by simulation. However, this probability is unlikely to have much practical significance. A more useful criterion would be whether some crucial feature such as total rainfall, or perhaps peak rainfall, is consistent with the bands set by the scenarios. A method for assessing the performance of the overall flood forecasting system, in respect of peak run-off, is discussed in Mellor *et al.* (2000).

Summary and conclusions

A system has been developed which strongly conditions a stochastic space-time rainfall model on real radar data. The model can create synthetic data sets which reproduce exactly the original observations and, when extrapolated, an ensemble of future space-time rainfall scenarios. It is, thus, possible to use these to drive a spatially distributed catchment model to produce scenarios of future runoff. Furthermore, since the rainfall model is continuous over space and time, the future fields can be sampled at whatever space-time resolution suits the catchment model, regardless of the resolution of the observing system.

The paper has described how real radar data are manipulated interactively so that any curvature or angular velocity inherent in a real rainstorm can be identified and, through a simple transformation of the data set, eliminated. The parameters of the MTB model are estimated from these straightened data, and a distribution of model raincells is

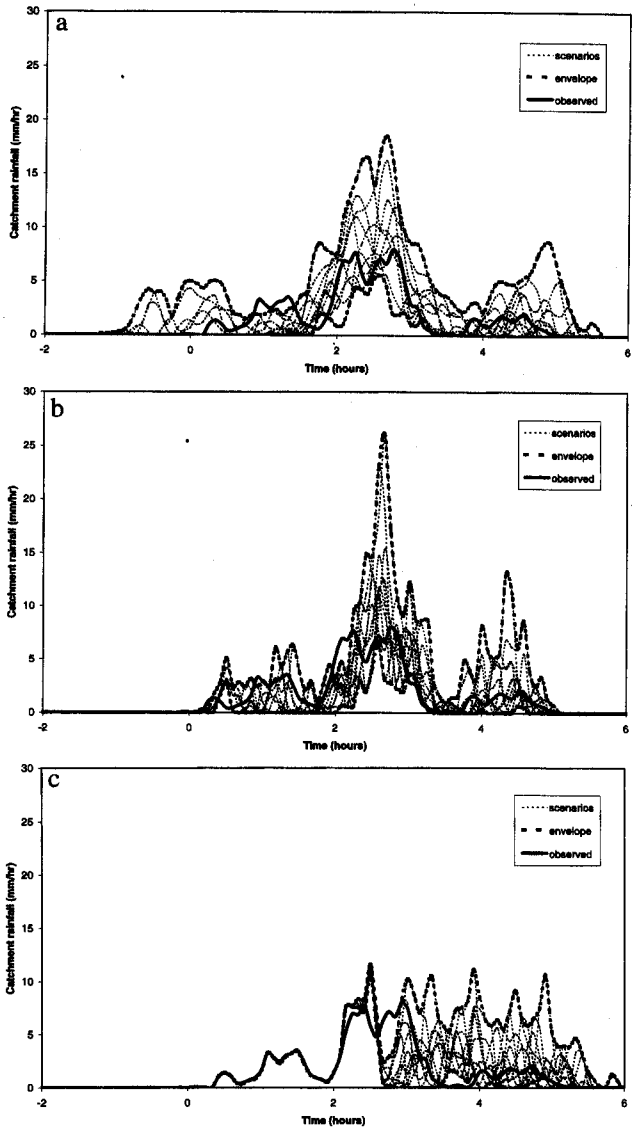

Fig. 9. *Storm of 3rd February 1994. Envelope of rainfall forecast scenarios generated by the MTB model conditioned on the observed rainfall: (a) 2 hours before the storm hits the catchment (−2 hours); (b) at the time the storm hits the catchment (0 hours); (c) 2 hours after the storm hits the catchment (+2 hours).*

determined which reproduces the most recent radar image. The underlying distribution function of these raincells is then determined, and this defines the probability distribution of raincells at future times. Based on this information, the model is used to generate randomly future raincells. When the curvature and angular rotation of the original storm is transformed back into the generated scenario, a possible future rainfall field is produced. The generation of the stochastic parts of the model may be repeated easily and quickly with a random number generator starting in a different state, so that ensembles of future scenarios can be generated. On a Hewlett-Packard 9000/715-50 desktop workstation, the production of the rainfall scenarios takes about two minutes for the examples displayed in this paper, plus an extra allowance of about two minutes for the manual fitting of the overall storm profile (the curvature and angular rotation). Thus, the overall production of future scenarios takes place sufficiently quickly for real-time forecasting of the rainfall field to be feasible.

The above paragraphs have summarised the developments of the MTB model, and a demonstration of the potential utility, that could be achieved in the time available for the HYREX project. The work has concentrated on a single storm event that took place in the Brue catchment. Despite the strong conditioning, the variability across the ensemble of forecasts appears to be high. From a forecasting standpoint this is inconvenient, but it may be an intrinsic characteristic of the rainfall process rather than a practical limitation of the proposed rainfall model.

The performance of the forecasting system in the Brue case study shows that it is capable of producing envelopes of rainfall intensity which take account of the observed rain band structure before, at the time when, and after a storm first hits the catchment. For this case at least, all sets of 10 envelopes include not only the peak intensity and the total rainfall, but also most of the storm profile. An ongoing appraisal of the system would include storms with a less well defined structure such as convective events. Further work is desirable to evaluate the varied nature of inputs from different operators of the system, and thus to determine how much scope the system has to utilize the operator's own skill and judgement. The programme of work described in this paper has been designed to investigate the ability of the MTB rainfall forecasting system to encapsulate the uncertainty in the future rainfall field, in isolation from variability induced by measurement errors and imprecise model parameters. The effect of errors, uncertainties and variability in these methods, after being compounded with the inherent variation of the MTB system, should be assessed. The techniques developed also have potential as a stochastic interpolator of rainfall fields (in both space and time), and this area also warrants further exploration.

Arising out of interaction with radar hydrologists and mesoscale atmospheric modellers within HYREX, a number of interesting avenues of research could be explored in the future, as outlined below.

Coupling of MTB forecasting methodology to radar

The work reported here has used data which have been processed and formatted by the Meteorological Office, but radar has potentially much greater flexibility than this, leading one to contemplate a situation in which the radar is controlled directly by a computer in real-time running the MTB system, thus allowing the MTB model to dictate dynamically its own inputs as the prevailing atmospheric conditions change. This would be a long-term research aim, but one which potentially could provide an avenue of ground-breaking advances in radar operation and processing and theoretical stochastic model development.

Extension to decision-making for flood warning

Ultimately, the results of this research may be used in a real-time decision-making environment, as it is necessary to devise methods which take advantage of the dynamic uncertainty in the forecasts to allow probabilistically optimal decisions to be made. Such decision-making would of necessity be based on ensembles of flow forecasts, as generated in Mellor *et al.* (2000).

Nesting of MTB model within larger scale dynamic forecasting models

The MTB model is an event-based model working on scales up to the synoptic level. Wider applicability could be gained by nesting the MTB model inside Mesoscale Forecasting Models, thus providing a means of downscaling numerical model forecasts to the spatial scales of interest in hydrology.

References

Abdullah, R., 1996. *Rainfall forecasting algorithms for real-time flood forecasting*, Ph.D. thesis, University of Newcastle upon Tyne, UK.

Bellon, A. and Austin, G.L., 1978. The Evaluation of two years of real-time operation of a short-term precipitation forecasting procedure (SHARP), *J. Appl. Meteorol.*, 17, 1778–1787.

Bremaud, P.J. and Pointin, Y.B., 1993. Forecasting heavy rainfall from rain cell motion using radar data. *J. Hydrol.*, 142, 373–389.

Browning, K.A., 1979. The FRONTIERS plan: a strategy for using radar and satellite imagery for short range precipitation forecasting. *Meterol. Mag.*, 108, 161–184.

Browning, K.A., 1981. Maximizing the Usefulness of Rainfall Data from Radars *Nowcasting Symp. at 3rd Assembly of IAMAP*, Hamburg, 47–64.

Collier, C.G., 1991. The combined use of weather radar and mesoscale numerical model data for short-period rainfall forecasting. In: *Hydrological applications of weather radar*, I.D. Cluckie and C.G. Collier (Eds.), Ellis Horwood, Chichester, UK, 331–348.

Cotton, W.R., Thompson, G. and Mielke, P.W., 1994. Real-time mesoscale prediction on workstations. *Bull. Amer. Meterol. Soc.*, 75, 349–362.

Cowpertwait, P.S.P., 1994. A generalised point process for rainfall. *Proc. Roy. Soc. London A*, 447, 23–27.

Cranston, M.D., 2000. A comparison of rainguage data and rainfall-radar observations in Scotland. *Hydrology in Scotland 2000* poster paper. http://www.wraffius.freeserve.co.uk/paper1.htm

Diskin, M.H., 1987. On the determination of the speed of moving rainfall patterns. *Hydrol. Sci. J.*, 32, 1–14.

Duda, R.O. and Blackmer Jr., R.H., 1972. *Application of pattern recognition techniques to digitised weather radar data*, Final Report Contract 1-36092, SRI Project 1287, Stanford Research Institute, Menlo Park, Cal., 135 pp.

Einfalt, T., Denoeux, T. and Jacquet, G., 1990. A rainfall radar forecasting method designed for hydrological purposes. *J. Hydrol.*, 114, 229–244.

Einfalt, T. and Semke, M., 1997. Reliable quantitative radar data–how long do hydrologists still have to wait? In: *Weather Radar Technology for Water Resources Management*, B.P.F. Braga and O. Massambani (Eds.), Unesco Press, Montevideo, 124–136.

Elvander, R.C., 1976. An evaluation of the relative performance of three weather radar echo forecasting techniques. *Proc. 17th Weather Radar Conf.*, Seattle, Amer. Meteorol. Soc., 526–532.

Foufoula-Georgiou, E. and Georgakakos, K.P., 1991. Hydrologic advances in space-time precipitation modelling and forecasting. In: *Recent Advances in the Modelling of Hydrologic Systems*, D.S. Bowles and P.E. O'Connell (Eds.), D. Reidel Publishing Company, Boston, MA, 47–65.

Foufoula-Georgiou, E. and Krajewski, W., 1995. Recent advances in rainfall modelling, estimation and forecasting, *Rev. Geophys.*, supplement, 1125–1137.

Georgakakos, K.P., 1986a. A generalised stochastic hydrometeorological model for flood and flash-flood forecasting. I. Formulation. *Water Resour. Res.*, 22, 2083–2095.

Georgakakos, K.P., 1986b. A generalised stochastic hydrometeorological model for flood and flash-flood forecasting. II. Case studies, *Water Resour. Res.*, 22, 2096–2106.

Georgakakos, K.P., 1987. Flash flood prediction, *J. Geophys. Res.*, 92, 9615–9629.

Georgakakos, K.P. and Foufoula-Georgiou, E., 1991. Real time coupling of hydrological and meteorological models for flood forecasting. In: *Recent advances in the modelling of hydrologic systems*, D.S. Bowles and P.E. O'Connell (Eds.), D. Reidel Publishing Company, Boston, MA.

Gupta, V.K. and Waymire, E., 1993. A statistical analysis of mesoscale rainfall as a random cascade. *J. Appl. Meteorol.*, 32, 251–267.

Haggett, C.M., Merrick, G.F. and Richards, C.I., 1991. Quantitative use of radar for operational flood warning in the Thames area. In: *Hydrological applications of weather radar*, I.D. Cluckie and C.G. Collier (Eds.), Ellis Horwood, Chichester, UK, 590–601.

Hand, W.H., 1996. A technique for nowcasting heavy showers and thunderstorms. *Meterol. Apps.*, 3, 31–41.

Huff, F.A., Vogel, J.L. and Changnon Jr., S.A., 1981. Real-time rainfall monitoring-prediction system and urban hydrologic operations, *Water Resources Planning and Management Div. Proc. A.S.C.E.*, 107, 419–435.

Interagency Research Committee, 1993. *First Report on the Hydrological Use of Weather Radar*. Institute of Hydrology, Wallingford, UK.

Kumar, P. and Foufoula-Georgiou, E., 1993a. A multicomponent decomposition of spatial rainfall fields: I. Segregation of large and small scale features using wavelet transforms, *Water Resour. Res.*, 29, 2515–2532.

Kumar, P. and Foufoula-Georgiou, E., 1993b. A multicomponent decomposition of spatial rainfall fields: II. Self-similarity in fluctuations. *Water Resour. Res.*, 29, 2533–2544.

Lee, T.H. and Georgakakos, K.P., 1992. A stochastic-dynamic model for rainfall forecasting. In: *Proc. of the CCNAA-AIT joint seminar on prediction and damage mitigation of meteorologically induced natural disasters*, C.H. Yen and W.F. Krajewski (Eds.), Nat. Taiwan University, 304–315.

Mellor, D., 1993. *The Modified Turning Bands (MTB) Model for space-time rainfall*, Ph.D. thesis. University of Newcastle upon Tyne, UK.

Mellor, D., 1996. The Modified Turning Bands (MTB) Model for space-time rainfall. I. Model definition and properties, *J. Hydrol.*, 175, 113–127.

Mellor, D. and O'Connell, P.E., 1996. The Modified Turning Bands (MTB) Model for space-time rainfall. II. Estimation of raincell parameters, *J. Hydrol.*, 175, 129–159.

Mellor, D. and Metcalfe, A.V., 1996. The Modified Turning Bands (MTB) Model for space-time rainfall. III. Estimation of the storm/rainband profile and discussion of the future model prospects, *J. Hydrol.*, 175, 161–180.

Mellor, D., Sheffield, J., O'Connell, P.E. and Metcalfe, A.V., 2000. A stochastic space-time rainfall forecasting system for real time flow forecasting II: Application of SHETRAN and ARNO rainfall runoff models to the Brue catchment. *Hydrol. Earth System Sci.*, 4, 617–626.

Molinari, J. and Dudek, K., 1992. Parameterization of convective precipitation in mesoscale numerical models: A critical review, *Monthly Weather Rev.*, 120, 236–344.

Over, T.M. and Gupta, V.K., 1994. Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on the large-scale forcing, *J. Appl. Meteorol.*, 33, 1526–1542.

Parkin, G., O'Donnell, G., Ewen, J., Bathurst, J., O'Connell, P.E. and Lavabre, J., 1996. Validation of catchment models for

predicting land-use and climate change impacts. 2. Case study for a Mediterranean catchment. *J. Hydrol.*, 175, 595–613.

Rodriguez-Iturbe, I., Cox, D.R. and Isham, V., 1987. Some models for rainfall based on stochastic point processes. *Proc. Roy. Soc. London*, A410, 269–288.

Sansó, B. and Guenni, L., 1999. Venezuelan rainfall data analyzed by using a Bayesian space-time model. *J. Roy. Statist. Soc.* C, 48, 3, 345–362.

Schultz, G.A., 1994. Remote sensing for forecasting of floods, in coping with floods. In: *Proc. NATO Advanced Study Institute, Erice (Italy), 3–15 November, 1992*, G. Rossi, N. Harmancoiglu and V. Yevjevich (Eds.), NATAO ASI SERIES E Applied Sciences, Vol. 257, Kluwer, Dordrecht, 459–472.

Seo, D.J. and Smith, J.A., 1992. Radar-based short-term rainfall prediction. *J. Hydrol.*, 131, 341–367.

Tessier, Y., Lovejoy, S. and Schertzer, D., 1994. Multifractal analysis and simulation of the global meteorological network. *J. Appl. Meteorol.*, 33, 1572–1586.

Todini, E., 1996. *AFORISM A comprehensive forecasting system for flood risk mitigation and control.* Commission of European Communities Final Report Research Contract EPOC-CT90–0023.

Todini, E., 1997. The role of rainfall measurements and forecasts in real-time flood forecasting and management. In: *Weather Radar Technology for Water Resources Management*, B.P.F. Braga and O. Massambani (Eds.), Unesco Press, Montevideo, 201–220.

WMO, 2000. *Forecast Demonstration Projects.* http://www.wmo.ch/web/arep/wwtp-fdps.html

Yoshino, F., Abe, E. and Tamamoto, S., 1991. Real-time calibration methods of radar rainguages. In: *Hydrological applications of weather radar*, I.D. Cluckie and C.G. Collier (Eds.), Ellis Horwood, Chichester, UK. 95–104.