P2.3 On quality indicators for radar-based river flow forecasts

C.G.Collier*
University of Salford Salford, Greater Manchester
United Kingdom

1. INTRODUCTION

In many river basins worldwide the time from rain reaching the ground to the arrival of the peak flow at a flood-sensitive point can be short, in some cases only a matter of minutes. The result, if the flood peak is large enough, is known as a flash flood which can bring death and destruction rapidly without warning.

Floods such as these may be induced by an isolated extreme rainfall event, two moderate rainfall events occurring with a short separation time, rainfall with snowmelt or the sudden release of water from a glacier or a dam either manmade or created through debris accumulation. During these events rainfall measurements are usually required every 15 minutes or less, and forecasts of rainfall at ½ hour, preferably more often, to allow hydrological models to update steamflow predictions constantly (for review see Collier, 2007). Rivers with longer response time are, in general, not as demanding, and for large river basins flow measurements upstream may provide adequate forecasts of downstream flows. The directions in which storms move, and the coincidence of peak flows from tributaries, also impact floods, and require knowledge of the actual and forecast distribution of precipitation.

A wide range of hydrological models have been developed over many years. Simple Input-Storage-Output (ISO) or transfer function models may work well over restricted areas and time periods if appropriate model parameter optimisation procedures are adopted, but the rainfall-runoff process is non-linear and time variant restricting model applicability in some situations. However, as model structures become more complicated the impact of physical processes and spatial heterogeneity become more evident, and so it is not necessarily true that complexity improves flood prediction reliability (for a review see Beven, 2001).

Model parameter updating procedures have been developed in which one or more model parameters are varied in real-time in the light of recent model performance. In this way, real-time correction factors provide a constant check on flow predictions. Unfortunately the range of parameter updating possibilities is large, and procedures may become complex if a high level of reliability is to be achieved. There remains an urgent need to provide operational hydrologists with the means of assessing the validity or otherwise of flow forecasts if reliable and timely flood warnings are to be made. In his paper we describe one approach to assessing the quality of flow forecasts prepared using radar or raingauge information.

2. THE NEED FOR CONFIDENCE INDICATORS

Although there has been considerable progress in improving the quality of radar estimates of rainfall problems remain. The unpredictable nature of radar errors continues to discourage many operational hydrologists from using radar data quantitatively as input to models. The tendency has been to wait for all errors to be removed. Unfortunately it is unlikely that all errors will be removed for a very long time if at all. One should recall that data from raingauge networks remain prone to errors even after very considerable work over very many years. Perhaps a more profitable approach is to investigate the real-time derivation of data quality flags in order to advise users of the likely reliability of the radar estimates or forecasts derived from them.

Complementary approaches may involve the application of stochastic state-space

* Corresponding author address: Chris. G. Collier, University of Salford, School of Environ. & Life Sciences, Salford, Gt. Manchester, M5 4WT, UK; email: c.g.collier@salford.ac.uk

models (see for example Collier and Robbins, 2007) to handle the error characteristics of rainfall input to hydrological models. Whilst work in this area continues, data quality flags could also be used in such research.

In what follows we investigate one approach to the derivation of data quality flags based upon the use of a simple rainfall-runoff modelling approach. An index is tested on data from the United Kingdom.

3. HYDROLOGICAL EVALUATION OF THE IMPACT OF ERRORS IN RADAR ESTIMATES OF RAINFALL

3.1 *Hydrograhs derived using radar rainfall data*

Collier and Knowles (1986) discussed a number of cases in which data from the Hameldon Hill C-band radar in North West England were used as input to an isolated event rainfall-runoff model. Hourly radar and raingauge totals were used from a number of different synoptic situations. River hydrographs derived using both the radar and raingauges were in near perfect agreement in some cases, and in other cases there were very wide differences due mainly to deficiencies in the input radar data. Similar results have been found by many authors over the last 20 years or so (see for example Schulz, 1987, Pereira and Crawford, 1995).

To investigate these variations further three cases for catchments in North West England have been selected for analysis, a local thunderstorm situation over the Ewood Catchment, River Darwen, a case of warm sector orographic rain over the River Ribble catchment and a case of heavy showery rainfall over the River Croal. A fourth case of a radar-based rainfall forecast is also analysed as a test of the procedure. In the local thunderstorm case the radar greatly overestimates the rainfall during one particular hour, the resulting radar-derived hydrograph peak flow being very much an overestimate. This overestimate is probably due to the presence of hail in the radar beam. For the orographic rainfall case the radar underestimates the rainfall as much of it occurs at very low altitudes, and the peak flow is correspondingly underestimated. In the shower case the radar underestimates the peak rainfall rates, but estimates the storm total rainfall quite well.

3.2 *Spectral analysis*

In order to estimate the characteristic frequencies of a rainfall time series, the time series may be spectrally analysed to identify what frequencies are present within it. Figs 1-3 are plots of the normalized spectral density (NSD) (see for example Wilks, 1995) as a function of time period (frequency).

$$\text{NSD} = \frac{n/2 \ \ C_k^2}{(n-1) \ \ S_y^2} \tag{1}$$

Where n is the number of values in the time series; k is an integer referring to particular harmonics i.e. the kth harmonic; S_y is the variance of the time series.

$$C_k = [A_k^2 + B_k^2]^{1/2} \tag{2}$$

$$A_k = \frac{2}{n}\sum_{t=1}^{n} y_t \cos\left(\frac{2\pi kt}{n}\right) \tag{3}$$

$$B_k = \frac{2}{n}\sum_{t=1}^{n} y_t \sin\left(\frac{2\pi}{n} kt\right) \tag{4}$$

The smallest frequency is the fundamental frequency, 2 π/n, where n is the number of measurements in the series, and the highest frequency is the Nyquist frequency given by,

$$W_{n/2} = \pi \tag{5}$$

If the spectrum of the data series includes important physical processes that vary faster than the Nyquist frequency, the data series is said to be undersampled. Variations that occur at frequencies higher than the Nyquist frequency are spuriously attributed to some lower, but representative, frequency. These high frequency variations are said to be *aliased*. Unfortunately it is not possible to tell from the data values alone whether appreciable contributions to the spectrum have been made at frequencies higher than the Nyquist frequency, or how large these contributions might be.

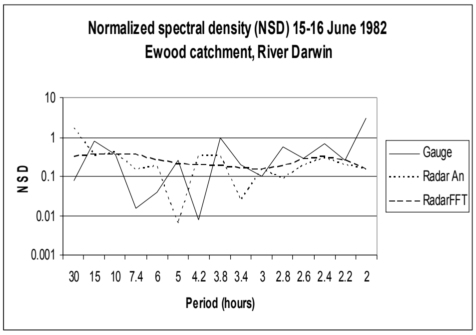

Figure 1: Spectral analyses of flow predictions for raingauge and radar time series inputs to an event-based hydrological model observed for the Ewood catchment, River Darwen, NW England 15-16 June 1982 (see Collier and Knowles, 1986).

The spectrum of a data series of values equally spaced in time may be computed analytically from the usual formulae for harmonic coefficients (see for example Wilks, 1995). This approach has been used to produce Figs 1-3. Whilst this is a reasonable approach for small datasets, large data sets require a more computationally efficient technique such as the Fast Fourier Transforms (FFTs). Unfortunately the way in which spectral estimates are calculated may cause them to be erratic depending upon the impact of aliasing. Even increasing the sample size does not give more precise information about individual frequencies. This problem is illustrated in Fig 1 which also shows the spectra for this case calculated using the

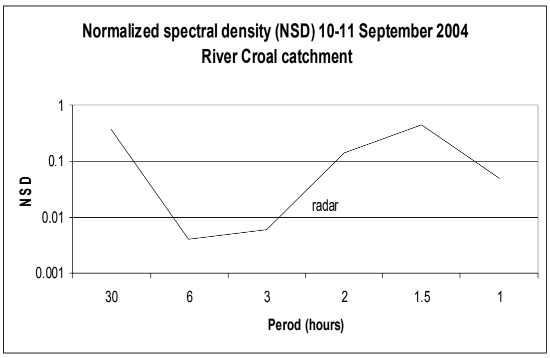

Figure 2: Spectral analysis for raingauge and radar time series for the inputs to stochastic hydrological model of the Croal catchment NW England, 10-11 September 2004.

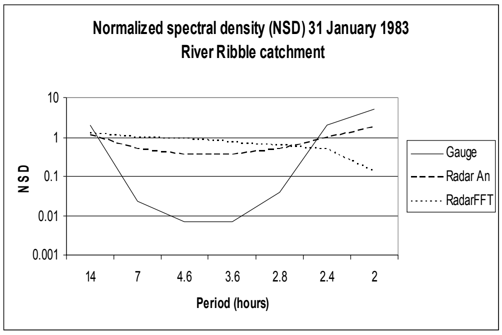

Figure 3: Spectral analyses of flow predictions for raingauge and radar time series inputs to an event-based hydrological model used for the Ribble cathment NW England, 31 January-1 February 1982 (see Collier and Knowles, 1986). The analyses for the radar data are carried out using both analytical and FFT techniques.

FFT method. Both spectra are similar in magnitude and trend, although the FFT method smoothes the more extreme fluctuations of those values derived analytically. Since we are not concerned here with absolute spectral values, the FFT approach is quite acceptable.
In order to extract a characteristic frequency from the spectral analyses, the derived spectral density values are used to weight the frequencies. For Fig 1, a thunderstorm case, using an event based model employing raingauge and radar inputs separately, this gives values of the characteristic time period of 2.1 hours for the raingauge time series, and 6 hour for the radar time series. In this case high frequencies are likely to be aliased into lower frequencies given the highly variable nature of thunderstorm rainfall. Clearly the hourly time step is adequate. The radar very much overestimates the rainfall for one particular hour making the aliasing even worse and increasing the characteristic time period.

We conclude that a spectral analysis of a river flow time series using a radar rainfall time series as input may provide useful additional information on the likely quality of the radar data. This possibility is examined further later.

3.3 *A data quality index*

If both raingauge and radar time series are available in real-time, the ratio of peak river flows calculated using both is a measure of hydrological impact. However, it is desirable to develop a quality index which may be derived from radar and other data which change more slowly (e.g. soil moisture) or change not at all (river catchment characteristics).

One possibility is to use a measure of maximum-likelihood that identifies whether values are part of a specific parameter distribution. Rather than use rainfall values, the use of peak river flow as a fraction of a threshold flow offers a parameter which reveals the severity of an event. However, we propose that the threshold peak flow is taken as the flow having a recurrence interval of once per year. This is a methodology reported in the Flood Estimation Handbook (FEH 1999). Observed peak flow divided by the threshold flow is designated the POT, Peaks Over Threshold. In what follows here we calculate the model peak flow using radar estimates of rainfall to calculate POT for comparison with values of the observed POT listed for specific catchments in the FEH. The FEH contains a comprehensive observed POT dataset for stations throughout the United Kingdom containing information on the maximum, median and arithmetic mean river flows and coefficients of variation of the series as a fraction (the standard deviation divided by the mean flow). We define a radar data quality index (QI) in terms of the difference between the calculated (or observed) POT ($POT_{c/o}$) and the POT for the median observed flow for each station within the area of radar coverage (POT_{med}) divided by POT_{med} that is,

$$\mathrm{QI} = \frac{POT_{c/o} - POT_{med}}{POT_{med}} \qquad (6)$$

Table 1 shows the values of QI for the observations and model using radar input observed in a number of case studies of peak flow. The range of QI is from – 1 to QI_{max} and is calculating taking $POT_{c/o}$ as the maximum value of POT so far observed. It might be that the $POT_{c/o}$ value exceeds the maximum value previously observed or is less than the median value of POT, and therefore the QI could take a value outside this range. However, in this case the likelihood that the values of rainfall and river flow are reliable is estimated by comparing the minimum and maximum values of POT with values which would be derived from a Poisson distribution over this range, where it is assumed that flood events occur randomly in time. In fact non-Poissoness is a feature of UK POT data and the impact of this is being investigated further.

Table 1: POT and QI value for cases of thunderstorm, orographic and shower rainfall in three catchments in North West England, and a case of extreme forecast rainfall over a small urban river catchment in north London.

Catchment	Rainfall Type	Threshold Cumecs	POT_{med}	POT_{max}	$POT_{c/o}$		QI*		Range of QI
					Observed	Radar model	Observed	Radar model	
Ribble	Orographic	350	1.70	3.49	0.72	0.26	-0.58	-0.84	-1 to + 1.05
Darwen	Thunderstorm	11.6	2.64	7.88	1.09	5.47	-0.59	+1.07	-1 to + 0.85
Croal	Shower	5.0	11.38	23.89	3.0	4.8*	-0.73	-0.31	-1 to + 1.10
Silk Stream London	Embedded convection forecast	4.1	2.85	10.24	-	119.5	-	+40.9	-1 to + 2.59

**Stochastic model used; for other cases an event-based deterministic model is used.*
+May be converted to a reliability factor using the range of QI and a Poisson distribution.

Unfortunately equation (6), whilst indicating whether the peak flow of the hydrograph lies outside the range of values which have occurred in the past, does not give absolute confirmation that the peak flow is to be believed. It may be that the calculated flow has not ever been observed in the past record, but is indeed accurate.

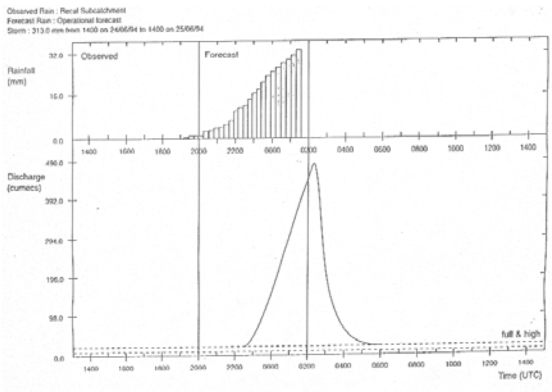

Figure 4: Radar-based rainfall and hydrograph forecasts for the Silk Stream, North London on 24-25 June 1994 (courtesy Environment agency Thames Region)

Fig 4 shows an example of forecast radar rainfall and the hydrograph derived using these data for a small urban river catchment, the Silk Stream in north London, on 24-25 June 1994. The QI for this case is also given in Table 1. This hydrograph was generated in real-time. Fortunately the operational hydrologists recognised that it was very unrealistic and a major flood warning was not issued. However, had the radar rainfall been somewhat less a difficult decision would have been required. It is such extreme cases which we <u>must</u> have confidence in the veracity of the hydrograph if weather radar is to become a totally reliable operational tool. The validity of the QI must be tested using additional information.

We must put aside the rainfall rate volume since there is no way yet of objectively determining whether its value is correct, unless the errors which are present in the data are known. This may be possible in the future as algorithms have been, and are being, developed to test for the presence of bright-band (melting snow), hail, ground clutter etc. using single polarisation, multi-parameter and Doppler radar data (see for a review Collier, 1996). However, at present, the reliability of these algorithms, whilst in some cases being individually acceptable, remains collectively unreliable for continuous objective operational use. Alternatively, the frequency of the rainfall time series may offer some deterministic method of identifying whether an extreme event is actually occurring. We assess this approach in what follows.

3.4 *Spectra of time-domain autoregressive models*

Rainfall is a continuous variable, and therefore the correlation structure of a time-series of rainfall measurements may be represented by a class of time series models known as Box-Jenkins models (Box and Jenkins, 1976). The types of time dependence produced by different autoregressive models generate characteristic spectral signatures that can be related to the autoregressive parameter ϕ (see for example Wilks, 1995), which in its simplest form is the sample lag one autocorrelation coefficient.

The simplest Box-Jenkins model is the first-order autoregression {AR(1)} model, sometimes called the *Markov process*. In this case positive values of ϕ induce a memory into the time series that tends to smooth over high frequency (small period) variations, and emphasize the low frequency (long period) variations. This leads to more spectral density at low frequencies and less spectral density at high frequencies, in this case the AR (1) process being known as “red noise”. For negative values of ϕ the opposite is true, and an AR (1) process tends to generate erratic short-time variations in the series. In this case the AR (1) process is known as “blue noise”, which is rare in environmental time series other than those heavily aliased. Finally a zero value of ϕ indicates an AR (1) process consisting of a series of temporally uncorrelated data values i.e. a time series of truly independent data. The spectrum is flat and the process is referred to as a “white noise” process.

Higher order autoregressions allow data values progressively further back in time as predictors. The spectral characteristics of higher order processes exhibit a wide variety of behaviour including periodicities.

In Fig 1, the thunderstorm case, we see that the raingauge time series has more spectral density in the high frequencies, a “blue noise” process, than the radar time series. This could indicate a movement towards dominant “red noise” processes namely those in which the most erratic point-to-point variations in the uncorrelated series are smoothed out. However, the spectral contribution in the low and high frequencies is very similar for the radar time series, and this tends to suggest that this time series is better represented by a higher order autoregressive model.

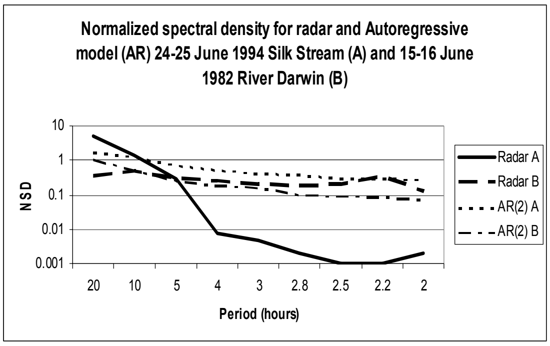

Figure 5: Sample theoretical spectral of a second order auto-regressive model superimposed upon the spectral density of the radar time series shown in Figs 1(B) and 4 (A).

This is further demonstrated in Fig 5 where sample theoretical spectra of a second order autoregressive model [AR (2)] are superimposed upon the spectral density of the radar time series. Two cases are shown namely the thunderstorm case calculated using the FFT method (Fig 1 and the case of extreme forecast rainfall 313mm in 24 hours, derived from the radar rainfall estimates on 24-25 June 1994 (Fig 4) calculated using an AR (2) model with ϕ_1 = 0.1 and ϕ_2 = 0.5 allowing for aliasing effects. However, the extreme rainfall case may only be modelled with very moderate success using an AR (2) model with ϕ_1 = 0.1 and ϕ_2 = 2.0. These parameter values fall outside the allowable parameter space for stationary AR (2) processes. Non stationarity can be seen as a drifting of the mean value. In this case this would indicate some problem with the radar calibration, and therefore the data set is suspect. Taken together with an extremely small value of the QI as shown in Table 1, the radar estimated peak flow is regarded as highly suspect.

4. A REAL-TIME PROCEDURE FOR THE GENERATION OF A HYDROLOGICAL RADAR QUALITY INDEX

4.1 *A real-time hydrological quality indicator*

Use of a hydrological model offers a simple procedure, based on the previous discussion, for quality controlling radar data from a hydrological point of view. The proposed procedure is as follows:

- At the time of each radar rainfall measurement estimate the likely peak flow for the particular catchment(s) of interest using the rainfall time series up to the current time
- Evaluate the POT_c from this peak flow
- Extract values of POT_{med} and POT_{max} from the catchment(s) history using the FEH.
- Calculate QI and compare against the range of QI previously experienced
- Carry out spectral analysis of the rainfall time series
- Fit autoregressive model(s) to the spectral analysis
- Assess the validity of fit and compare with QI analysis
- Set quality flag indicating likely reliability of hydrographs calculated using radar data

In the following section we go some way to testing this procedure.

4.2 *The Easter 1998 floods in the Midlands, England*

Rainfall during the period 8-10 April 1998 on already saturated land caused record flooding over extensive areas of Wales and Central and Eastern England. Convective cells were seen to be embedded within more stratiform rainfall. The hyetograph for a raingauge located near Banbury with the corresponding hydrograph is shown in Fig 6. The synoptic situation was a thundery, near stationery, low pressure system.

Real-time radar estimates of the rainfall were adversely affected during the first part of the event by an inability of an objective bright-band correction (Kitchen et al, 1994) procedure to handle correctly the convective cells embedded in stratiform rainfall consequently the radar estimates were reduced causing a significant underestimate. During the second half of the event the radar performed well in representing this rainfall. Hydrograph predictions using these estimates resulted in peak flows which were too low.

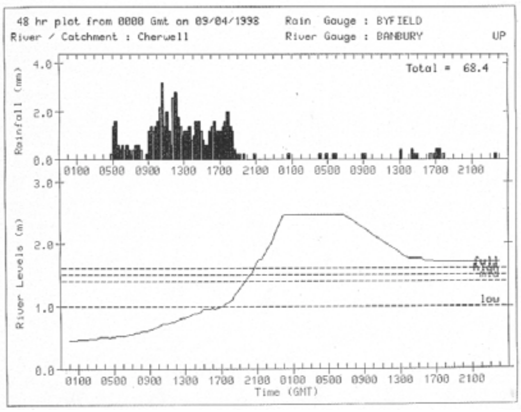

Figure 6: **Hyetograph for the Byfield raingauge located near Banbury and the Banbury hydrograph for the River Cherwell (courtesy Environment Agency, Thames Region).**

The deficiency (over estimate) of the radar estimate is revealed by the analysis of the raingauge and radar rainfall time series shown in Fig. 7 during the first half of the event (03-08 UTC).

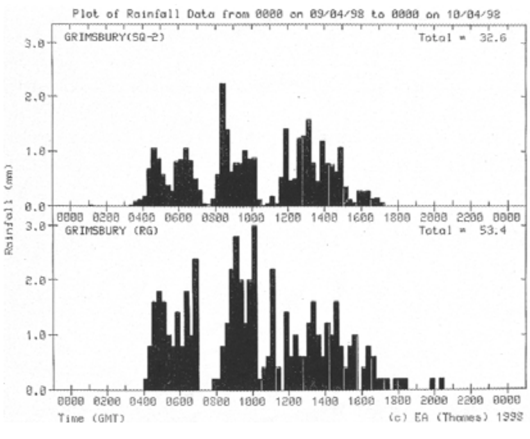

Figure 7: Grimsbury raingauge (upper panel) and corresponding 2km x 2km radar estimate time series from 0001 UTC 9 April 1998 to 0001 UTC 10 April 1998.

During the period 03-08 UTC the spectral analysis of the radar data is quite different from that of the raingauge data (Fig 8a). It is not possible to fit a physically realistic autoregressive model to the radar spectral analysis, and therefore the radar data are suspect. However, during the second half of the event, the spectral analyses of the raingauge and radar time series are very similar (Fig 8b). It is possible to fit a physically realistic autoregressive model to these data as shown. During the latter period it would appear that the radar is performing reliably.

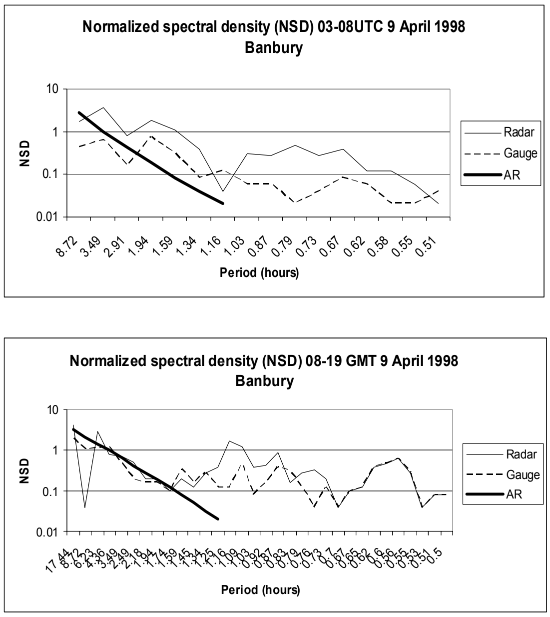

Figure 8: Spectral analyses for the raingauge and radar time series shown in Fig 7 for (a) 03-08 UTC (upper panel) and (b) 08-19 UTC (lower panel) on 9th April 1998. The radar analysis is shown as a dashed line and the raingauge analysis is shown as a solid line. Also shown (heavy-solid line) is an autoregressive model curve with the parameters given

We conclude that the use of the spectral analysis technique with both raingauge and radar data would have flagged the deficiency in the radar estimates of rainfall during the first half of the event. In practice the spectral analysis would have to be carried out continuously, and further consideration should be given to the minimum size of the time series necessary for reliable results.

5. CONCLUSIONS AND FUTURE WORK

A method of quality controlling radar estimates of rainfall has been described.

The technique is based upon a twofold approach. Firstly river peak flow is estimated using the radar data and a hydrological model. The calculated peak flow is compared against previously occurring peak flows using the FEH.

A further test involving the spectral analysis of raingauge, or river flow using the input radar time series data, is the carried out. Attempts to fit a physically realistic autoregressive model to these data are made. Failure to fit such a model to the radar time series spectral analysis indicates deficiencies in the radar estimates of rainfall.

Further case studies are recommended, as is the implementation of the system for real-time operation. The extent to which the spectral analysis can be associated with specific types of radar error should be investigated.

Acknowledgement

This work was partially carried out using data collected daily the EUFP4 Microwave link Project and the NERC Bolton II Project. Thanks are due to Dr Chris Haggett of the Environmental Agency for supplying the radar, raingauge and river flow data for the case study in section 4.2.

References

Beven, K.J., 2001: Rainfall-runoff modelling. The Primer, John Wiley & Sons Ltd, Chichester, 360pp

Box, G.E.P. and Jenkins, G.M., 1976: Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 375pp

Collier, C.G.. 1996: Applications of Weather Radar Systems. A guide to uses of radar data in meteorology and hydrology, Wiley-Praxis, Chichester, 390pp

Collier, C.G., 2007: Flash floods: What are the limits of predictability?, *Quart. J.R.Met.Soc.*, **133**, 3-23

Collier, C.G. and Knowles, J.M., 1986: Accuracy of rainfall estimates by radar Part III: Application of short-term flood forecasting, *J. Hydrology*, **83**, 237-249

Collier, C.G. and Robbins, G.L., 2007: Using stochastic state-space modelling to specify and reduce hydrograph error, *Hydrological Processes* (in press)

FEH, 1991: Flood Estimation Handbook (Five Volumes), Centre for Ecology & Hydrology, NERC, Wallingford, UK

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

Pereira, A.J. and Crawford, K.C., 1995: Integrating WSR-88D estimates and Oklahoma mesonet measurements of rainfall accumulations: hydrologic response, Proc. II Int. Symp. Hydro. App. Wea. Radars, 20-23 August 1995, ABRH/IAHR, Sao Paulo, 169-178

Schulz, G.A., 1987: Flood forecasting based on rainfall forecasting in the Federal Republic of Germany, Chapter 13 in Weather Radar and Flood Forecasting, ed. V.K.Collinge and C.Kirby, John Wiley & Sons Ltd, Chichester, 191-207

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences, Academic Press, San Diego / London, 467pp