Hydrology and Earth System Sciences, 3(1), *95–108* (1999) © EGS

Hydrology & Earth System Sciences

Analysis of aggregation and disaggregation effects for grid-based hydrological models and the development of improved precipitation disaggregation procedures for GCMs

H.S. Wheater*, T.J. Jolley**, C. Onof*, N. Mackay* and R.E. Chandler†

*Department of Civil Engineering, Imperial College of Science, Technology & Medicine, London SW7 2BU, U.K.
**CEPA West, Rivers House, Murray Road, East Kilbride G75 OLA, U.K.
†Department of Statistics, University College, Gower Street, London WC1E 6BY, U.K.

Abstract

Appropriate representation of hydrological processes within atmospheric General Circulation Models (CGMs) is important with respect to internal model dynamics (e.g. surface feedback effects on atmospheric fluxes, continental runoff production) and to simulation of terrestrial impacts of climate change. However, at the scale of a GCM grid-square, several methodological problems arise. Spatial disaggregation of grid-square average climatological parameters is required in particular to produce appropriate point intensities from average precipitation. Conversely, aggregation of land surface heterogeneity is necessary for grid-scale or catchment scale application.

The performance of grid-based hydrological models is evaluated for two large (10^4km^2) UK catchments. Simple schemes, using sub-grid average of individual land use at 40 km scale and with no calibration, perform well at the annual time-scale and, with the addition of a (calibrated) routing component, at the daily and monthly time-scale. Decoupling of hillslope and channel routing does not necessarily improve performance or identifiability. Scale dependence is investigated through application of distribution functions for rainfall and soil moisture at 100 km scale. The results depend on climate, but show interdependence of the representation of sub-grid rainfall and soil moisture distribution.

Rainfall distribution is analysed directly using radar rainfall data from the UK and the Arkansas Red River, USA. Among other properties, the scale dependence of spatial coverage upon radar pixel resolution and GCM grid-scale, as well as the serial correlation of coverages are investigated. This leads to a revised methodology for GCM application, as a simple extension of current procedures.

A new location-based approach using an image processing technique is then presented, to allow for the preservation of the spatial memory of the process.

Introduction

Atmospheric General Circulation Models (GCMs) are important tools in the investigation of global climate and climate change (Department of the Environment, 1996), and the role of hydrological processes in determining the exchange of energy and water vapour at the earth's surface has been increasingly recognised. In the early 1960s, a basic representation of the surface hydrological cycle was incorporated in GCMs (e.g. Smagorinsky, 1963; Manabe *et al.*, 1965).This was followed by a number of relatively simple experiments which nevertheless demonstrated clearly that simulated climates were sensitive to soil moisture (Manabe, 1975; Walker and Rowntree, 1977; Rowntree and Bolton, 1978), evaporation efficiency (Charney *et al.*, 1977) and albedo (Charney *et al.*, 1977; Carson and Sangster, 1981). There has, thus, been progressive development and testing of more complex, physically based 'SVATs' for incorporation in the current generation of GCMs (e.g. Sellers *et al.*, 1986, Dickinson *et al.*, 1986, Henderson-Sellers, 1991, 1992).

However, the representation of surface hydrology in GCMs in particular, and in meteorological models in general, raises a number of methodological issues. The grid-scales of GCMs (typically of side 200–300 km) are large in comparison with hydrological systems, even at river basin scale. There is, thus, a major question of the

representation of surface heterogeneity in GCMs, for example in soil and vegetation characteristics, and of lateral transfers of water within and between grid-squares. These are the problems of *aggregation*.

The grid-scale of GCMs is also large in comparison with the spatial variability of meteorological variables, of which precipitation is particularly significant for hydrological response. Spatially averaged precipitation at these scales results in low intensity drizzle, which has no physical relevance to hydrological processes such as interception, infiltration and runoff generation. For example, in a simple hydrological model application of spatially-averaged meteorological variables at 3.75 × 2.5 degrees to the Nile basin, no flow was generated (Abourgila, 1992). Hence, there is the need for *disaggregation*, to produce hydrologically-meaningful precipitation and other meteorological fields. These issues of aggregation and disaggregation are compounded by the difficulty of making relevant process observations at appropriate scales.

Associated with these methodological problems is the issue of complexity of surface hydrological representation. Given that it is the effective properties at GCM grid-scale which are important in terms of atmospheric feedbacks, concern should not be focussed necessarily on point-processes, but on aggregated response. This can have important implications for model structure, since the extent to which point-process hydrological response can be integrated to larger scales is still an open question (Wheater *et al.*, 1993). Related issues include robustness of performance to uncertainty in physical properties, and requirements for computational efficiency within the overall GCM solution scheme.

This paper is divided into two major sections. In the first, performance of relatively simple grid-based hydrological models is explored, together with evaluation of sensitivity to sub-grid representation of surface heterogeneity and meteorological variables, through analysis of catchment-scale response. The objective is to use measurements of water balance at catchment scale to evaluate large-scale model performance and, further, to identify the controls of sub-grid scale heterogeneity on model performance. Interdependence between representation of sub-grid scale heterogeneity and rainfall disaggregation is demonstrated, which points to the need for independent determination of these effects. Hence, in the second section, data analysis is used to define procedures for disaggregation of precipitation and their parameterisation. Analysis of large-scale rainfall radar fields provides a basis for empirical improvement of current simple, distribution-based disaggregation procedures and shows scale-dependence of coverage properties. In addition, a new methodology is defined which preserves the spatial location and spatial memory of disaggregated rainfall.

Performance of a grid-based hydrological model

The availability of a routine service by the UK Meteorological Office provided the basis for evaluation of a relatively simple grid-based SVAT scheme. The Meteorological Office Rainfall and Evaporation Calculation System (MORECS) (Thompson *et al.*, 1981) uses the Penman-Monteith equation to estimate daily potential and actual evapotranspiration rates on a 40 km × 40 km grid over the UK, published as weekly values. Daily meteorological variables are interpolated and averaged over each square. The minimum surface resistance and albedo, which vary seasonally, are defined for 14 land-cover types. The model has a single interception store and a two-layer soil store. Water can evaporate freely from the upper soil store, which represents 40% of the available soil water capacity, but evaporation is restricted by the soil moisture deficit in the lower soil layer. Available water capacity varies seasonally to allow for crop development. If evaporation exceeds rainfall, a soil moisture deficit develops. If rainfall exceeds potential evaporation, the soil stores rewet until 'hydrologically effective rainfall' is generated, which can, in principle, generate runoff or groundwater recharge. The water balance calculation is carried out separately for each land-surface type present in a given grid-square, and the fluxes and soil moisture deficits are areally-averaged (i.e. a tile-based aggregation).

Results were presented by Jolley and Wheater (1996) for application of the uncalibrated water balance model to the Severn and Thames UK river basins, each of area approximately 10,000 km^2, but with contrasting rainfall and geology (Fig. 1). In the Severn basin, annual precipitation varies from over 1500mm in the west of the catchment (the Cambrian mountains) to less than 700 mm in the east. Catchment average annual precipitation is 824 mm.There is relatively little variation in potential evapotranspiration across the basin (average 547 mm). Soils vary from thin upland peat and podzolic soils in the west and north-west to poorly-drained clays in the east. Groundwater resources are relatively limited. In the Thames, there is little variability in relief and precipitation. Average annual precipitation is 686 mm and average potential evapotranspiration 577 mm. Soils range from well-drained sandy-loams to poorly-drained clays, and the catchment contains a major chalk aquifer.

Over a common ten-year period, catchment water balances were reproduced with an error in (naturalised) runoff of +1% and –6%, respectively (positive error indicating overestimation of runoff); over 20 years (Severn) and 30 years (Thames), the errors reduced to +2.3% and –0.5%, respectively. Additional routing was required to represent subsurface and network routing effects to reproduce monthly hydrograph dynamics. A simple (calibrated) 2-parameter routing model produced R^2 values of 90% and 88% respectively.

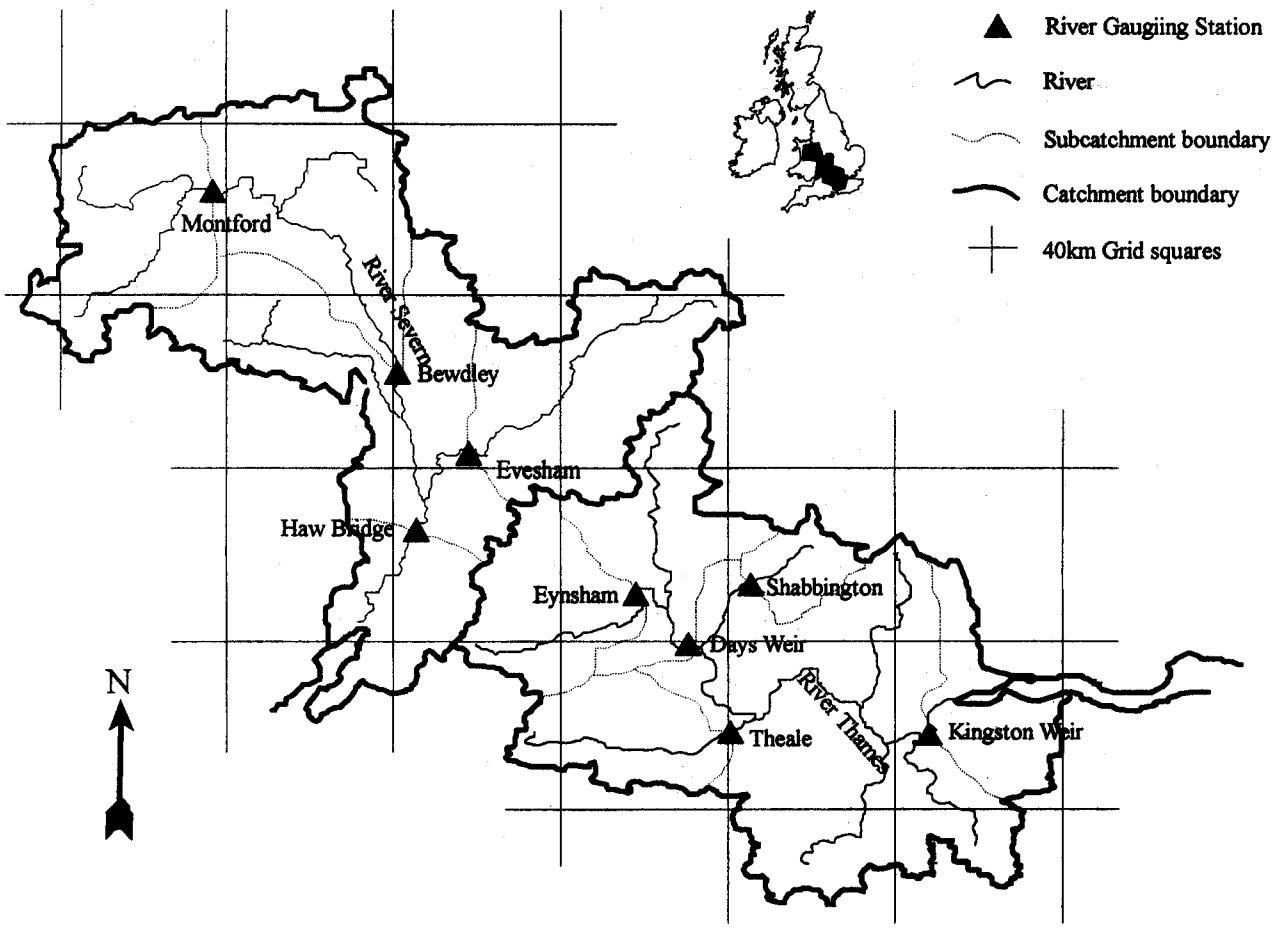


Fig. 1. *Severn and Thames catchments.*

The simple bucket model thus provided extremely good performance at whole-catchment scale. However, sub-catchment performance was poorer, with a systematic trend. Considering the climate ratio of annual precipitation to potential evaporation, the wetter sub-catchments showed positive errors in annual runoff, and as climate ratio decreased, progressively larger negative errors (flow underestimate) were observed (Fig. 2). It was apparent that performance degraded for drier climatic conditions.

A more extensive analysis of routing algorithms for the grid-based modelling was undertaken by Jolley and Wheater (1997b); this showed that the incorporation of physics-based channel network routing algorithms did not prevent the need for calibration to represent subsurface (hillslope and aquifer) routing effects.

Analysis of aggregation and disaggregation effects

A further analysis was made, based on a close approximation to the MORECS model, to investigate the issues of spatial aggregation and disaggregation with respect to catchment-scale response. In current GCMs, e.g. Warrilow *et al.*, 1986, an attempt is made to represent the fact that rainfall is expected only over a proportion of a grid square by arbitrarily assigning a proportion wet (dependent on rainfall type). A simple exponential distribution function is applied to represent the distribution of rainfall intensity. There is no spatial memory of rainfall location between time-steps, nor of soil moisture location. A distributed soil moisture calculation is averaged at the end of each time-step.

Performance based on the 40 km × 40 km grid was compared with 100 km × 100 km simulations under various treatments of spatial aggregation (Jolley and Wheater, 1997a). For example, when meteorological inputs were averaged at 100 km scale, but land use/soil moisture response was retained at 40 km resolution, a negative bias was introduced into the runoff simulation. Annual runoff decreased by 5.2% for the Severn and 2.3% for the Thames. Conversely, when meteorological inputs were retained at 40 km resolution, but soil moisture deficits were averaged at 100 km scale, a positive bias occurred. This was small for the wetter Severn catchment (+0.8

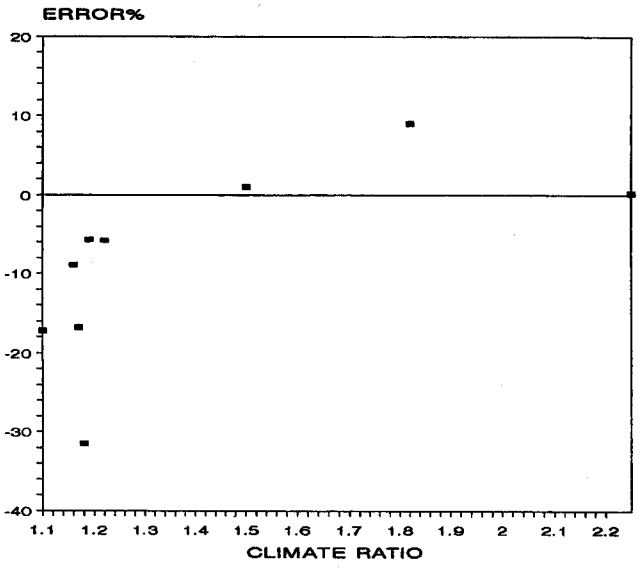


Fig. 2. *Error in simulated annual runoff as a function of climate ratio.*

mm/year, +0.0%), but larger for the Thames (+26.3 mm/year, +11.9%). The combined effect of averaging meteorological inputs, soil moisture deficits and soil moisture capacity at 100 km scale was relatively small, a reduction of 2.6% and 2.5% for the Severn and Thames respectively. It was evident that averaging meteorological inputs and soil moisture response had opposing effects, the magnitude of which varied according to the catchment climatic characteristics. For the drier Thames, loss of resolution in soil moisture led to a much larger increase in runoff than for the wetter Severn. For the Severn catchment, which has a marked gradient of annual rainfall, as noted above, averaging the meteorology produced a marked reduction in runoff, much more so than for the Thames, which has much less spatial variability in annual rainfall. For both catchments, the opposing influences tended to cancel out when a fully aggregated simulation was carried out.

In a further extension, a distribution function approach was applied to represent the sub-grid variability of rainfall and of the soil moisture upper store capacity, Smax1, at 40 km scale. In accordance with previous studies (Warrilow *et al.*, 1986; Eagleson *et al.*, 1987), it was assumed that the rainfall depth R is distributed exponentially over a fraction ε of a grid square, so that the distribution of rainfall is given by

$$f(R) = \frac{\varepsilon}{\bar{R}} \exp\left(-\frac{\varepsilon}{\bar{R}} R\right)$$

where $\bar{R}$ is the mean grid-square rainfall.

The capacity of the upper soil moisture store Smax1 was assumed to be distributed according to the two-parameter gamma distribution

$$f(x) = \frac{1}{a^{\beta}\Gamma(\beta)} x^{\beta-1} \exp\left(-\frac{x}{\alpha}\right)$$

where x is the random variable (Smax1 in this case), Γ is the gamma function and a and b are the scale and shape parameters, which can be related to the mean and coefficient of variation (CV) of Smax1.

The model response is then calculated by integration to account for zones of hydrological response, as shown by Jolley and Wheater (1997a). It is, thus, dependent on the assumed proportion dry ε and on the CV of the soil moisture upper store. Results for the Severn (Fig. 3) show

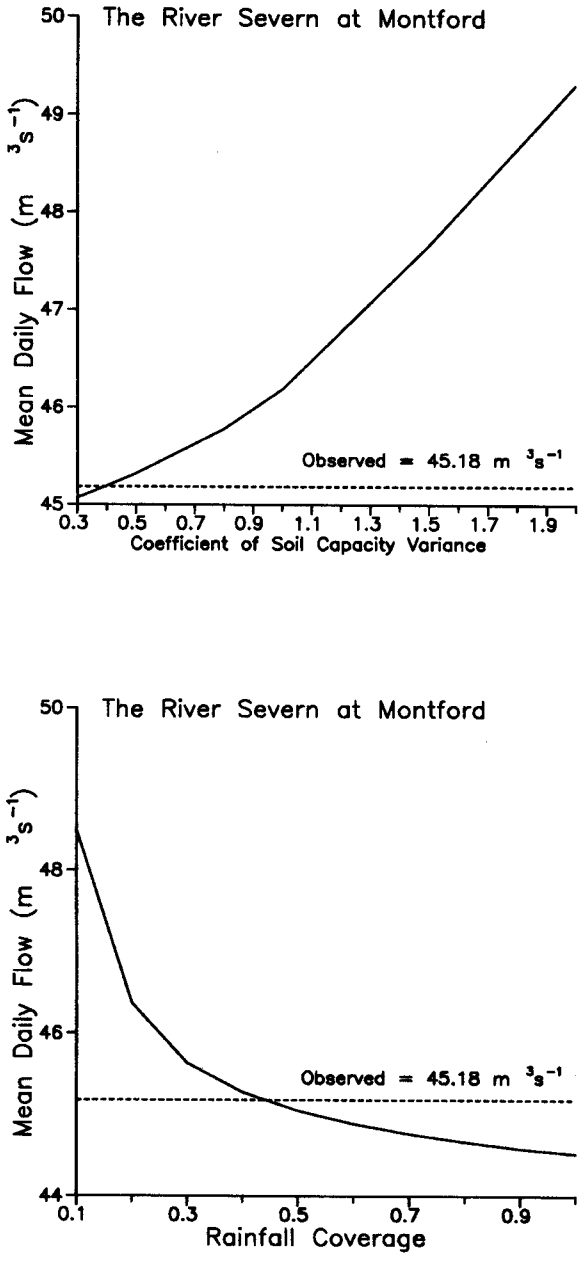


Fig. 3. *Sensitivity of mean daily flow to sub-grid variability of soil capacity and rainfall.*

runoff decreasing with increasing rainfall coverage and increasing with increasing variability of soil moisture capacity. Jolley and Wheater (1997a) showed that these effects are magnified progressively as the climate factor (rainfall/potential evaporation) decreases.

These results, therefore, confirm the increase in sensitivity of runoff to aggregation and disaggregation procedures, and show that there is an interdependence between the effects of rainfall disaggregation and soil moisture aggregation. Hence, parameters of the two distributions cannot be identified uniquely from model calibration alone.

Rainfall disaggregation

As noted above, in meteorological models in general and General Circulation Models (GCM) in particular, a single value of the rainfall depth is generated at each time-step for each grid-square, for grid-square sizes of the order of 200–300 km when the models are run on a continental scale. Clearly, the knowledge of an average over a large area is not sufficient for hydrological modelling. A disaggregation scheme, therefore, is required to downscale large scale rainfall so as to preserve the observed spatial distribution.

GCM MODEL

As noted above, GCMs commonly assume that precipitation occurs over a proportion ε of the grid-square and that, where it rains, point rainfall depth is distributed exponentially (Warrilow *et al.*, 1986) with parameter $\varepsilon/\bar{R}(w,t)$, where (w,t) is the unconditional mean rainfall intensity over the grid-square of size w^2 at time t, which is the value provided by the GCM atmospheric component. The values of ε are chosen equal to values of 0.1, 0.3 or 0.5 depending on the model and the rainfall type (Gregory and Smith, 1990).

Analyses of UK data have shown that the choice of a fixed value of the coverage ε, dependent solely upon the rainfall type, does not reflect the spatial structure of real radar fields (Onof and Wheater, 1996a) and that the time-series of consecutive coverages is highly correlated at short time-intervals and has a fairly long memory (Onof and Wheater, 1996b), a feature which is not reproduced by the current scheme. Improvements such as a choice of the coverage ε as the ratio of $\bar{R}(w,t)$ and the climatological rainfall intensity ρ (Eltahir and Bras, 1993) have been proposed and are discussed in Onof and Wheater (1996a).

We shall analyse here ways of improving upon this scheme by focusing upon the rainfall coverage as it is measurable from radar data and extending the data analyses presented in the two previous papers.

Analysis of coverage

DATA SETS AND METHODOLOGY

The data used for the analyses and for model calibration are the fully calibrated Arkansas Red River data (area: 850,000 km^2) and the uncalibrated Wardon Hill radar in the South-West of England (radius 210 km). For the first, hourly data at a resolution of 4×4 km^2 and for the second, 5 minute data at resolutions of 2×2 and 5×5 km^2 are available. The Arkansas data is the product of the integration of radar data from 17 WSR-88D radars calibrated using approximately 500 raingauges. Wardon Hill is a C-band radar with scanner height 255 m above sea-level, beamwidth 1° and National Grid coordinates (360900, 102300); a comparison with 50 raingauges showed that there was no systematic bias in this data set.

These data sets are used to estimate $\mathcal{E}(p,w,t)$ which is the coverage measured with a pixel size of p^2, of a window of size w^2 at time t. This estimate is obtained by taking averages over many non-overlapping and not completely dry windows of size w^2 of the proportion of wet pixels on the radar picture.

DEPENDENCE UPON WINDOW SIZE

Figure 4, for the Arkansas data, shows that the coverage is a scaling quantity for window sizes larger than a given threshold, i.e. that there is a gradient $\alpha(p,t)$ on the log-log plot such that:

$$\mathcal{E}(p,\lambda w,t) = \lambda^{\alpha(p,t)}\, \mathcal{E}(p,w,t) \quad \text{for } w \geq w_0 \qquad (1)$$

where the threshold may be the consequence of the uncertainty introduced in the estimation of the coverage when the window size approaches the pixel size. This limit w_0 appears to be about 20 km for Arkansas, and at the most, 4 km for Wardon Hill (Onof *et al.*, 1977).

DEPENDENCE UPON PIXEL SIZE

The variation of $\mathcal{E}(p,w,t)$ with p is shown for the Arkansas data in Fig. 5 (with w maximum, i.e. corresponding to the whole data range): scaling holds over a range of scales between p_0 and p_1 but there is a smaller dependence upon scale for smaller and larger pixels sizes:

$$\mathcal{E}(\lambda p,w,t) = \lambda^{\beta(w,t)}\, \mathcal{E}(p,w,t) \quad \text{for } p_0 \leq p \leq p_1 \qquad (2)$$

The double scale dependence exhibited above is characteristic of multi-fractals for which the existence of limits upon the range of scales over which such a multi-scaling behaviour holds has already been observed by many authors (Kumar and Foufoula-Georgiou, 1993).

TEMPORAL DEPENDENCE

There is obviously some long memory dependence in the time-series of radar rainfall coverages, as shown by the

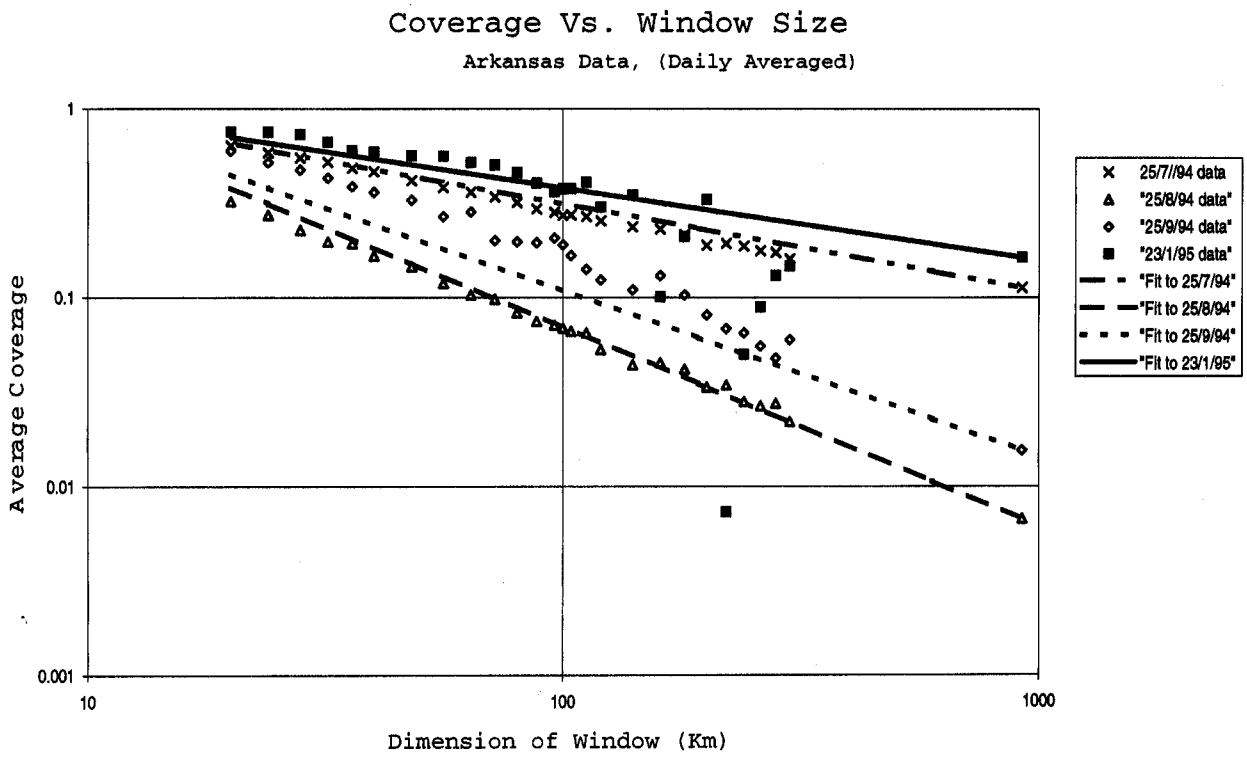


Fig. 4. *Rainfall coverage as a function of window size.*

autocorrelation function for the Arkansas data in Fig. 6a, which confirms the findings of Onof and Wheater (1996b). However, the current GCM scheme allows only for variability of ε if the rainfall type varies; constant ε is assumed in the case of a persistent rainfall type.

The mean areal depths exhibit an analogous autocorrelation pattern although, on the whole, the autocorrelations are smaller.

Using this analogy, but bearing in mind the lack of any linear relationship between the two series (since both large scale drizzles—large ε and small $\bar{R}$—or intense local storms—small ε and large $\bar{R}$ are observed), the relation between the logarithms of the $\bar{R}(p,w,t)$ and $\mathcal{E}(p,w,t)$ is examined. These show a clear linear dependence which stretches over the whole range of coverages and mean intensities (Fig. 6b) and which explains more than 80% of

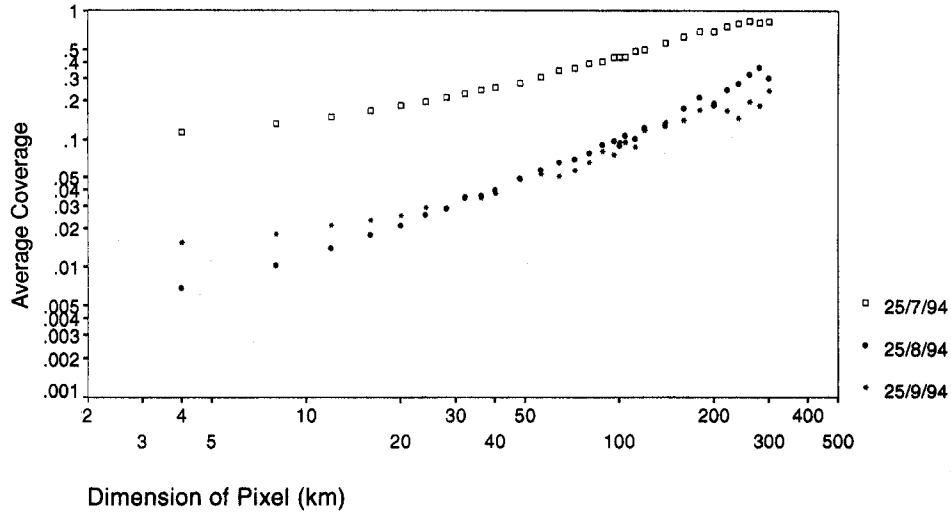
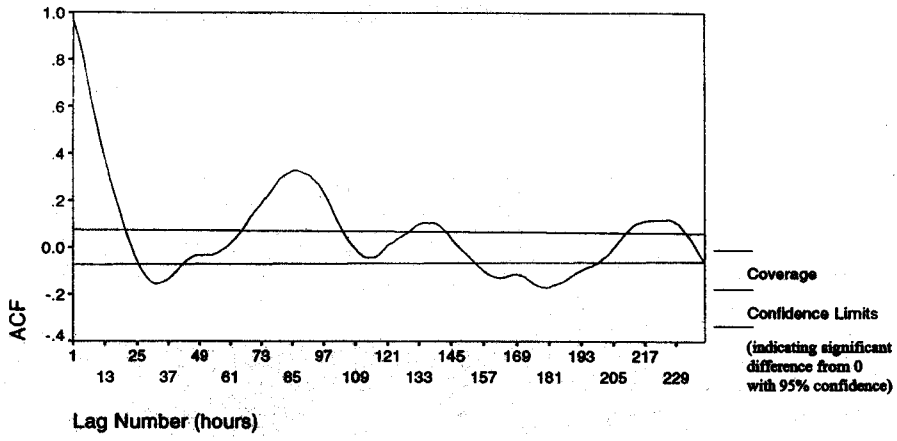
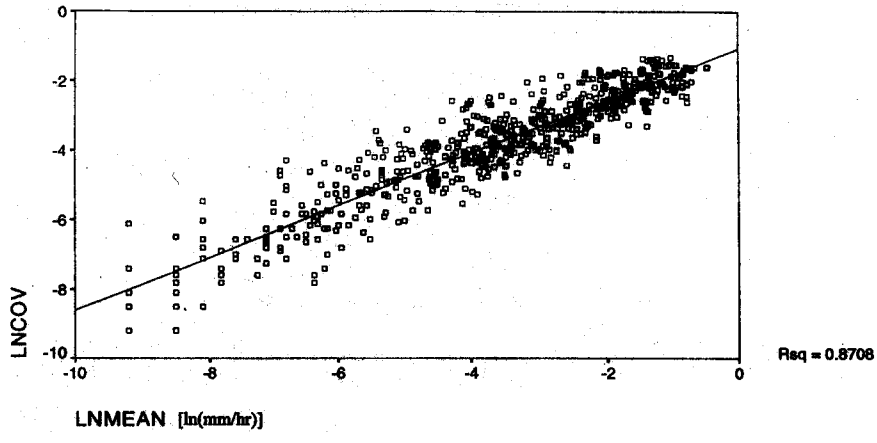


Fig. 5. *Rainfall coverage as a function of pixel size.*

(a)



(b)



(c)

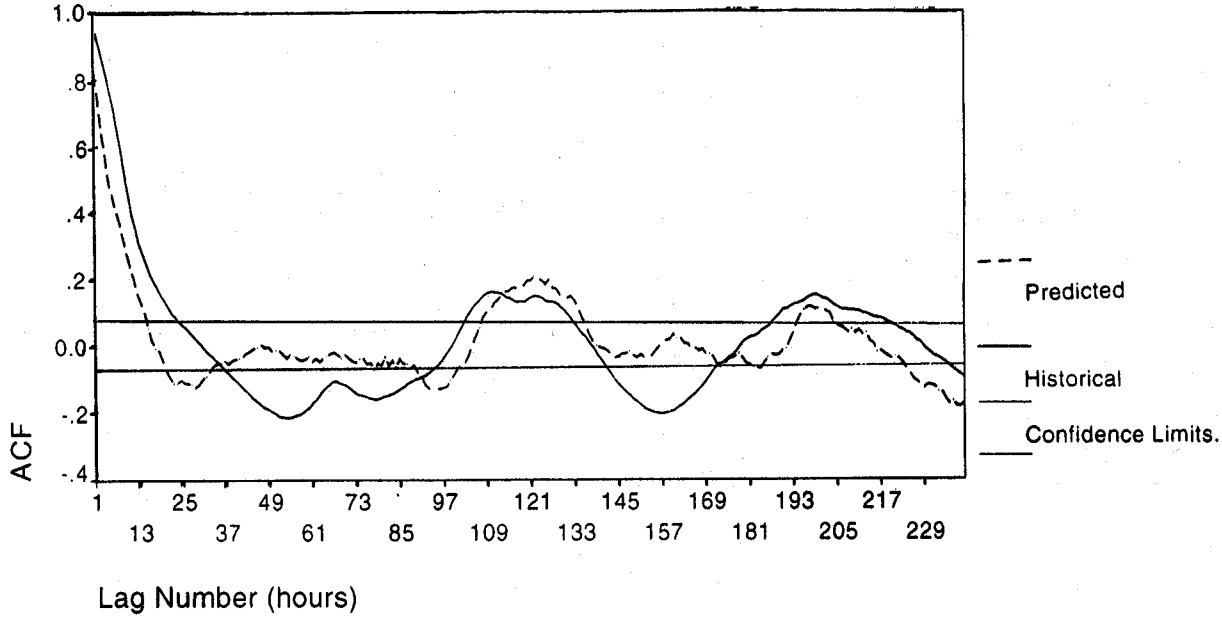


Fig. 6. *(a) Autocorrelation function of historical hourly rainfall coverages*
(b) Rainfall coverage as a function of mean intensity
(c) Autocorrelation function of historical And modelled hourly rainfall coverages.

the variance of the logarithms of $\mathcal{E}(p,w,t)$ over a whole month.

This enables an approximation of the value of $\mathcal{E}(p,w,t)$ by using the following regression:

$$\ln\{\mathcal{E}(p,w,t)\} \_ a(p,w) \ln\{\bar{R}(p,w,t)\} + b(p,w) \qquad (3)$$

This relationship provides a simple model which reproduces much of the observed autocorrelation in the radar data by estimating the coverage using the mean rainfall depths (Fig. 6c).

IMPROVED DISAGGREGATION TECHNIQUE: METHODOLOGY AND VALIDATION

These results can be brought together into a simple model to improve current procedures for the selection of the rainfall coverage ε for a meteorological model with grid-square size γ^2. For GCMs, a reasonable assumption consists in choosing $\bar{R}(w,t) \simeq \bar{R}(\gamma,t)$, which latter is given by the model at each time-step; more generally, for meteorological models, this assumption can be applied as a first approximation:

1. $\mathcal{E}(p,w,t)$ is obtained as a function of $\bar{R}(w,t)$, using the regression relationship (3) for a window size w which is that of the data analysis (e.g. radar range);
2. $\mathcal{E}(p,\gamma,t)$ is obtained as a function of $\mathcal{E}(p,w,t)$ using the scaling relationship (1);
3. ε is then chosen as considered appropriate in view of the observed weak dependence upon scale for small pixel sizes, for instance $\mathcal{E}(p_0,\gamma,t)$ for the smallest available scale p_0 of measurement.

As validation for this proposed methodology, apart from the data analyses above, it is interesting to consider the interaction of the different scaling effects with the log-log dependence of coverage upon rainfall. By aggregating the original 4 km Arkansas data to p = 8 km and looking at this dependence in windows of side w = 40 km, a regression with parameters a(8,40) and b(8,40) is obtained.

However, it is also possible to obtain these parameters by using the original a(4,930) and b(4,930) for the whole Arkansas range and scaling down using the relationships (1) and (2). The latter change for the pixel size corresponds to a multiplication by 1.2. The parameters shown in Fig. 7 indicate a good match between the two methods, which is an important result for the validation of consistency of the proposed disaggregation methodology.

This methodology, for which more detailed results are presented in Onof *et al.* (1997), allows for a clear improvement upon the original GCM disaggregation scheme by taking into account the evidence from extensive data analysis of rainfall coverages. However, because it is a distributional approach, it is not adapted to represent the spatial memory of the precipitation process, the importance of which was demonstrated in the Severn and Thames analyses. The question of how knowledge of the statistical properties of rainfall in space and time can be used to develop a location-based disaggregation method is addressed in the next section.

An algorithm for spatial location

The work described in the previous sections was primarily concerned with improving the representation of coverage within existing disaggregation methodology, and did not seek to address the problem of determining the spatial location of wet and dry regions within a GCM grid square. In this section a new approach to the problem is described; this incorporates spatial location, as the ability to reproduce fine-scale spatial pattern may have a major impact upon the usefulness of GCM output for hydrological purposes. Other approaches include the use of multi-scaling properties as the basis for a successful disaggregation approach (Perica and Foufoula-Georgiou, 1996) but this does not reproduce the spatial memory of the process. Research is currently ongoing into developing space-time random cascades which would form the basis for a multi-scaling disaggregation model (Over and Gupta, 1996, Marsan *et al.* 1996) but questions remain about the limits of the validity of the multi-scaling hypothesis.

The approach is motivated by the 'cascade algorithm', suggested by Jennison (1986) in the context of statistical image restoration. This is an area which has much in common with rainfall disaggregation; in both cases, the objective is to reconstruct an unknown 'image' from imperfect data. In the case of image restoration, the imperfection is due to noise added to the original image, whereas in the disaggregation problem, the imperfection arises because the data are only available as a very coarse-scale spatial average. The methodology, as applied to the disaggrega-

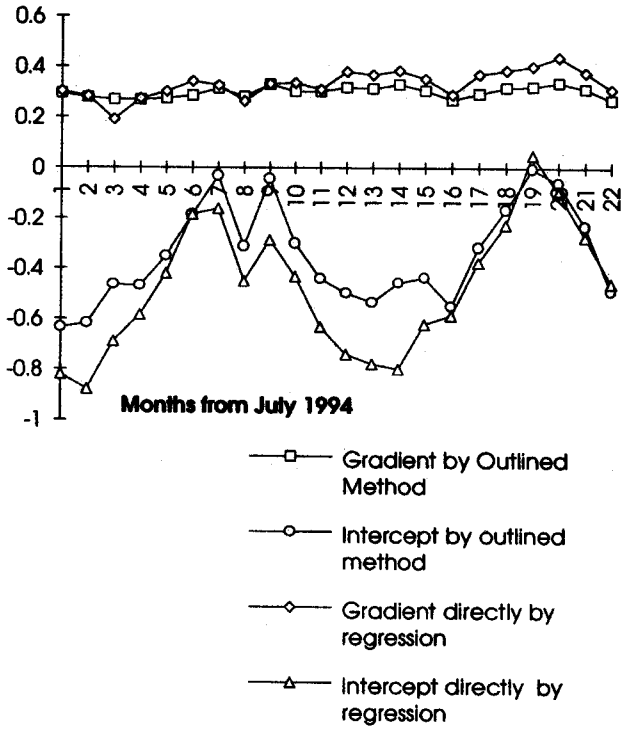


Fig. 7. *Model validation by comparison of regression parameters.*

tion problem, is described in detail in Chandler *et al.* (1997): here an overview of the main concepts is given.

OVERVIEW OF THEORY

For the development of a scheme incorporating fine-scale spatial pattern, it is convenient to subdivide each GCM square into smaller squares (*pixels*) at the scale of interest. The main assumption underlying the scheme is that, at any instant in time, the spatial pattern of the rainfall at the scale to which we are disaggregating can be modelled as a Markov Random Field (Besag, 1974; Isham, 1981; Ripley, 1988), which is essentially the spatial equivalent of an autoregressive time series model. This section concentrates solely on the wet/dry rainfall pattern, so deals with a binary random field.

The essence of a Markov Random Field (MRF) model is the following: for all pairs of pixels in a scene, a symmetric *neighbourhood relation* is defined which states whether or not the two pixels are neighbours; furthermore, each pixel is defined to be its own neighbour. A random field has the Markov property with respect to a particular neighbourhood structure if the conditional probability that a pixel is wet or dry, given values at all other pixels, depends only on the configuration in its neighbourhood. Following common image reconstruction practice, an autologistic model (Besag, 1974) has been adopted to model the wet/dry pattern. In this model, the neighbourhood of a pixel is defined to be the set of all pixels immediately horizontally, vertically and diagonally adjacent. Horizontal and vertical neighbours are said to be of *type 1*, while diagonal neighbours are of *type 2*. With this in mind, the MRF model used in this work may be specified as

$$\frac{Pr\{Pixel\ is\ wet \mid rest\ of\ scene\}}{Pr\{Pixel\ is\ dry \mid rest\ of\ scene\}} = \exp\left[\alpha = \sum_{k=1}^{2} (\beta_k \times no.\ of\ wet\ type\ k\ neighbours)\right]$$

where α, β_1 and β_2 are parameters to be estimated. β_1 and β_2 measure in some sense the dependence between a pixel and its type 1 and 2 neighbours; α is most intuitively thought of as the log odds ratio for a pixel being wet when it is completely surrounded by dry pixels. All three parameters can be estimated straightforwardly from radar data using methods described by, for example, Besag (1974).

In addition to the MRF model, which plays the role of prior distribution in a Bayesian framework and reflects the belief as to what a 'typical' rainfall field looks like, temporal dependence is incorporated into the disaggregation procedure at each time point by conditioning upon the disaggregated image at the previous timestep. In the implementation used to date, this requires two parameters: p_{11}, the probability that a pixel was wet at time $t-1$ if it is known to be wet at time t, and p_{10}, the probability that a pixel was wet at time $t-1$ if it is known to be dry at time t. These are *transition probabilities*, and they are easily estimated.

Finally, it is advantageous to incorporate information from the GCM rainfall intensities into the disaggregation procedure, as there is clearly a relationship between the mean rainfall intensity over a grid square and the number of pixels within that square which are wet—this has been demonstrated empirically in a previous section. To incorporate this information into a coherent framework, it is necessary to specify, for each GCM square, the conditional distribution of coverage given the mean rainfall intensity. The linear relationship, noted above, between the logarithms of these quantities, is the most obvious choice given that mean grid square rainfall is simply the product of the proportion of wet grid squares and the mean rainfall within those grid squares. However, this is unsatisfactory for predicting coverages for a variety of reasons, not least of which is that it is quite feasible to obtain predicted coverages which are greater than 1; it is therefore necessary to seek some other form of dependence. For the purposes of implementation, it is convenient if regression residuals appear to have the same distribution across the entire range of observation: from this point of view, the most promising relationship found is between the empirical logit (Cox and Snell 1989) of the coverage and the log mean rainfall. This requires 3 parameters: the regression intercept and slope and the residual standard deviation—again, these are easily estimated from radar data. From exploratory data analysis, the residuals appear to have an approximately symmetrical distribution, which is approximated by a normal distribution with appropriate continuity corrections.

These various model components may be combined using Bayes' Theorem to obtain a joint probability distribution for the wet/dry configuration at time t, given both the reconstructed image at time $t-1$ and the GCM rainfall intensity information at time t: full details of the calculations are omitted here for reasons of brevity, but may be found in Chandler *et al.* (1997). Some assumptions are made in the process: these are felt to be physically reasonable, at least as a first approximation. For example, it is assumed that the fine-scale wet/dry pattern at time $t-1$ is conditionally independent of the coarse-scale rainfall *intensities* at time t, if the fine-scale wet/dry pattern at time t is known—again, see Chandler *et al.* (1997) for full details.

Sampling directly from this high-dimensional joint distribution is non-trivial: however it may be accomplished indirectly in a straightforward way using the Gibbs sampler (Geman and Geman 1984; Ripley 1988), and this enables an ensemble of scenarios each of which is consistent with the coarse-scale GCM output.

PARAMETER ESTIMATION

Based upon a 12-month data set (July 1994 – June 1995), a fixed set of parameters was calculated for each month of

the year for the Arkansas-Red River Basin assuming grid-square dimensions of 40 × 40 km and a pixel size of 8x8 km. Parameters were calculated for each image in this period and over each month a weighted average was calculated (parameters were weighted by the conditional mean intensity of the rainfall to place emphasis on hydrologically significant events).

The parameters thus calculated show obvious seasonal trends but contain some noise (not unexpected with only one year's data). Therefore, the parameter estimates are smoothed using a 3-point moving average window with weights (0.25, 0.5, 0.25), with June 1995 and July 1994 considered to be adjacent for this purpose.

The resulting monthly parameter estimates are given in table 1.

VALIDATION

A 15 hour event was chosen (1700, 14/7/95 to 0700 15/7/95) to test and validate the disaggregation scheme. This event was selected as it was outside the period over which the parameters were estimated and was also preceded by a completely dry hour for which the disaggregated field is trivially known. The data were first aggregated from 4 × 4 km pixels to 8 × 8 km pixels to give a 'true image', then the 8 × 8 km pixels were aggregated to 40 × 40 km to give an aggregated image.

The probability of misclassifying a pixel (wet or dry) using the proposed model and knowing the true state of all the pixel's neighbours was calculated for each pixel in the field and an average misclassification probability calculated for the whole field. These Perfect Information Misclassification Probabilities (PIMPs) provide the upper limit the model can achieve, as the true state of each pixel's neighbours is unknown while the dissagregation is being performed. This criterion also allows for comparison of the existing model with other possible model extensions. However, model extensions incorporating more comprehensive neighbourhood relationships showed limited potential for improvement and were not evaluated further at this stage.

The aggregated image was given as input to the scheme and the parameters listed in Table 1 were used in the disaggregation process. 100 realisations of the storm were generated and statistics of the results are based on these. A sample image is shown in Fig. 8 along with the real and aggregated images for visual comparison. This image appears to reproduce the general sparseness of the observed rainfall but the obvious deficiencies are the lack of diagonal banding in the disaggregated rainfall field and edge effects. The edge effects are probably due to having a border around each image which is fixed at the start of the disaggregation procedure and does not change subsequently as the field is disaggregated. This is one of the areas of the model which will be addressed in future work.

Looking at the images quantatively, several features of the wet/dry disaggregated field can be evaluated. Figs 9a and b show the proportion of pixels in the disaggregated rainfall field which are correctly classified, considering only wet grid-squares. The thick solid line represents the performance of the disaggregation scheme over 100 realisations. Figure 9a shows the theoretical best performance for this model, i.e. (1-PIMPs); also shown is a plot of the model performance using parameters calculated from each real image in the event instead of the seasonal parameters. It is evident that little information is lost by using seasonal parameters for this storm. In Fig. 9b, three representations of the expected effective performance of the original GCM scheme with fixed coverages of ε_{GCM} = 0.1, 0.3 and 0.5 are shown. This is calculated from:

$$1\text{-PIMPs} = \varepsilon_{GCM} \times E + (1 - \varepsilon_{GCM})(1 - E)$$

where E is the observed coverage. Additionally, a reference line is plotted of the proportion of correctly classified

Table 1.

	Markov Random Field Parameters			Transition Probabilities		Regression Parameters			Gamma scale parameter
Month	α	β_1	β_2	p_{11}	p_{10}	Intercept	slope	RSD	
January	−5.92	2.55	0.48	0.66	0.25	3.14	0.74	1.37	1.478
February	−5.79	2.57	0.44	0.60	0.23	2.98	0.73	1.38	1.495
March	−5.66	2.52	0.44	0.62	0.24	2.87	0.73	1.37	1.25
April	−5.50	2.46	0.41	0.63	0.25	2.63	0.73	1.33	0.772
May	−5.22	2.36	0.36	0.60	0.25	2.19	0.71	1.33	0.428
June	−4.96	2.25	0.35	0.56	0.22	1.80	0.67	1.37	0.31
July	−4.78	2.21	0.31	0.53	0.20	1.57	0.62	1.42	0.32
August	−4.71	2.19	0.29	0.51	0.19	1.68	0.61	1.47	0.475
September	−4.83	2.17	0.36	0.54	0.20	2.04	0.61	1.48	0.59
October	−5.17	2.22	0.46	0.60	0.24	2.36	0.65	1.40	0.612
November	−5.56	2.30	0.56	0.66	0.28	2.68	0.70	1.32	0.86
December	−5.86	2.43	0.56	0.69	0.28	3.05	0.73	1.33	1.26

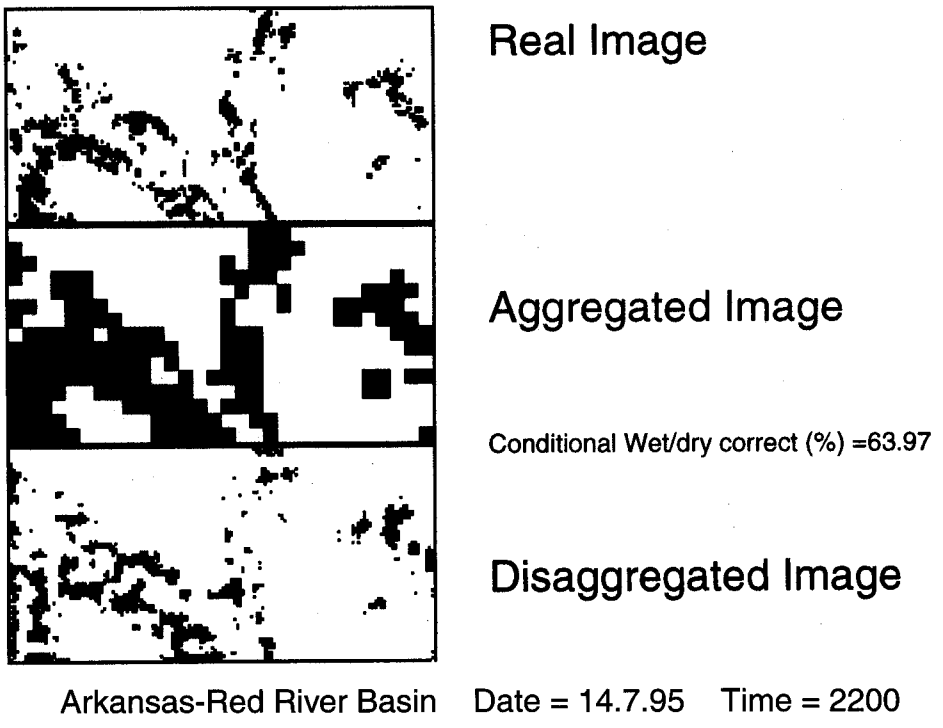
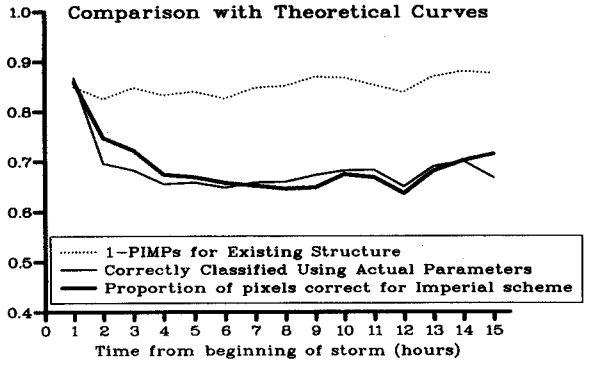


Fig. 8 *Imperial scheme disaggregated wet/dry field*



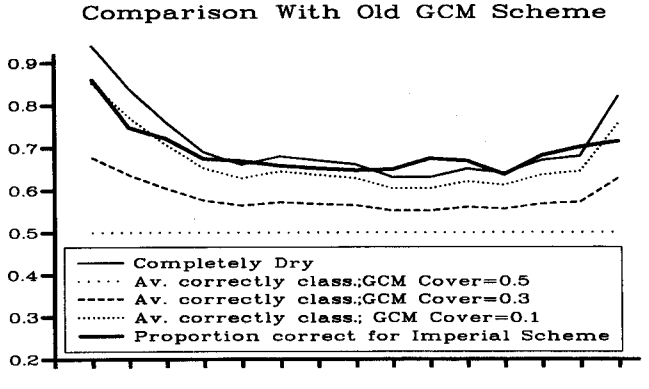


Fig. 9 *Proportion of correctly classified pixels*
(a) Comparison of Imperial scheme in/out of sample and theoretical limit
(b) Comparison of Imperial and other schemes

pixels given a completely dry disaggregated rainfall field. This is shown to highlight that such assessment must be made in the context of other features of the data. The proposed scheme (referred to as the Imperial disaggregation scheme) correctly classifies approximately 65–75% of the pixels, which is a similar performance to the completely dry scene, slightly better than the $\varepsilon_{GCM} = 0.1$ version and significantly better than the other fixed coverage estimates.

Other basic performance criteria include the coverage of the disaggregated field. Figure 10 shows graphs of the bias in coverage between the disaggregated rainfall field and the true rainfall field for the Imperial scheme and the GCM fixed coverage schemes with coverage $\varepsilon_{GCM} = 0.5$, 0.3 and 0.1 (plus the zero coverage case). It appears that the Imperial scheme reproduces the true coverages of the event very satisfactorily. The only version of the GCM scheme which comes close to the correct coverage for this event is that with a coverage of $\varepsilon_{GCM} = 0.3$, which did not perform well for the previous classification criterion.

EXTENSION TO INTENSITIES

Having established the wet/dry rainfall field, rainfall intensities must then be assigned, which sum to the grid-square total rainfall. Spatial rainfall is generally best modelled by a gamma distribution (Matsubayashi *et al.* 1984). Given a number of gamma distributed variables (n), a scale parameter (g) for the gamma distribution of the rainfall across the whole field and a total depth (T), we can

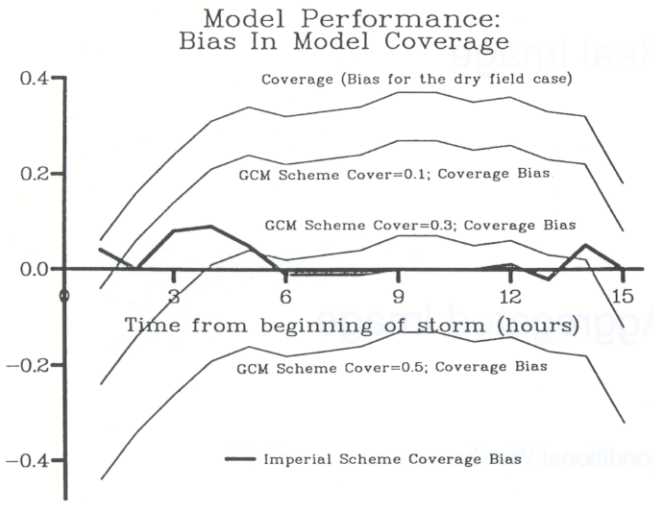


Fig. 10 *Bias in the disaggregated rainfall coverage*

generate n intensities from a gamma distribution with a scale parameter g which sum to T (the shape parameter of the gamma distribution is defined by g and T).

This leaves the problem of assigning the generated intensities to the wet pixels. If the values of these intensities and of the distances of each pixel to its nearest dry neighbour are ranked, over half the pixels analysed are such that higher ranked intensities have higher ranked distances. This result may be used to produce a rule of thumb for the allocation of intensities to wet pixels: these are randomly assigned in descending order by sampling pixels according to probabilities which are proportional to a power function of the distance to their nearest dry neighbour (power 5 is found to be optimal). Other methods for the assignment of intensities are the subject of ongoing research.

Figure 11 shows a realisation of the full disaggregation methodology for the same event as Fig. 9. The image appears to match the true field quite well but on close inspection, reveals that the intensities nearer the edges of each mass of rainfall are too low. When the gamma parameter is varied from that calculated for July, a higher rainfall level can be achieved round the edges but the distribution of intensities does not then reflect the true variation in intensities over the field. This is perhaps a reflection of the inadequacy of the Gamma distribution for reproducing rainfall intensities either in general or specifically at this time of year.

Conclusions

In this paper, the sensitivity of a grid-based hydrological model to sub-grid variability in soil moisture and precipitation is demonstrated for two large UK catchments. Performance of a non-calibrated tile-based SVAT scheme was excellent in terms of long-term water balance at catchment-scale. Calibrated routing was required to reproduce catchment flow dynamics at the shorter time-scales. The use of grid-scale averages for precipitation and soil moisture has climate-dependent opposing effects upon the resulting runoff. Sub-grid distribution of rainfall tends to increase runoff; sub-grid distribution of soil moisture tends to decrease runoff. The sensitivity to aggregation and disaggregation increases markedly as the ratio of

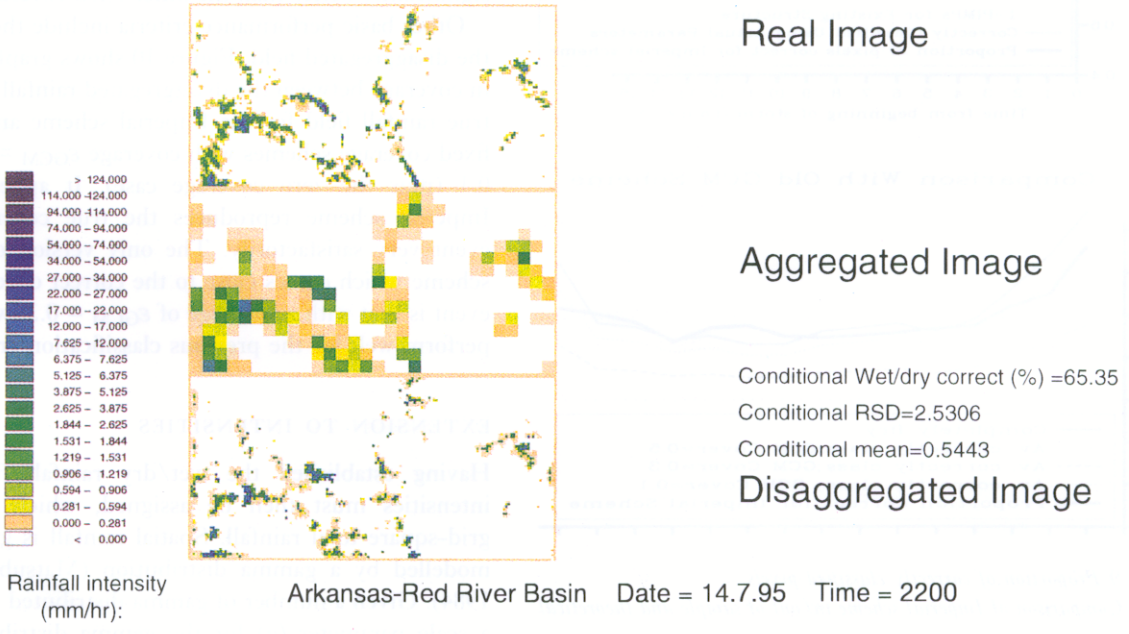


Fig. 11 *Imperial scheme disaggregated rainfall field*

annual precipitation to potential evaporation decreases. Thus, if a distributional approach is used to represent sub-grid scale variability, the required parameters cannot be estimated from model calibration alone.

Data analyses show how the parameter identification in a widely used rainfall disaggregation model can be improved by a proper estimation of the rainfall coverage which takes into account its dependence upon scale and rainfall intensity. This provides a way of reproducing the temporal dependence of the coverage.

The issue of the spatial location and memory requires a new disagregegation scheme. The Imperial model presented in this paper enables much of the spatial structure to be reproduced while avoiding excessive complexity which would impair its applicability as a potential interface between meteorological and hydrological models. Although the rainfall intensity representation in this model is being improved, the proposed scheme already provides a powerful tool for the generation of spatially and temporally consistent sub-grid scale rainfall fields.

Acknowledgements

The authors wish to acknowledge the support of the Natural Environment Research Council (NERC) who fund the TIGER programme within which this research was carried out.

References

Abourgila, A.E., 1992. *Large-scale hydrological modelling of the Nile basin.* Unpubl. MSc thesis, Imperial College, London.

Besag, J.E., 1974. Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc., Series B,* **36**, 192–236.

Carson, D.J. and Sangster, A.B., 1981. The influence of land-surface albedo and soil moisture on general circulation model simulations. In: (Rutherford, I.D.) *GARP/WCRP: Research Activities in Atmosphere and Oceanic Modelling. Numerical Experiment Programme*, Report no 2, pp 5.14–5.21.

Chandler, R.E., Mackay, N.G., Wheater, H.S. and Onof, C., 1997. Bayesian image analysis and the disaggregation of rainfall. *Research Report No. 184, Department of Statistical Science, University College London.*

Charney, J.G., Quirk, W.J., Chow, S.H. and Kornfield, J.K. 1977. A comparative study on the effects of albedo change on drought in semi-arid regions. *J. Atmos. Sci.*, **34**, 1366–1385.

Collier, C., 1992. The application of a continental-scale radar database to hydrological process parameterization within general circulation models, *J. Hydrol.*, **142**, 301–318.

Cox, D.R. and Snell, E.J., 1989. *Analysis of binary data (2nd edition).* Chapman and Hall, London.

Department of Environment, 1996. *Review of the Potential Effects of Climate Change in the United Kingdom.* UK Climate Change Impacts Review Group, March, HMSO, London, pp 247.

Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J. and Wilson, M.F., 1986. *Biosphere-atmosphere-transfer scheme (BATS) for the NCAR community climate model.* National Centre for Atmospheric Research, Boulder, Colorado, Tech. Note/TN-275+STR.

Eagleson, P.S., 1978. Climate, soil and vegetation, 2. The distribution of annual precipitation derived from observed storm sequences, Wat. Resour. Res., **14**, 713–721.

Eagleson, P.S., Fennessey, N.M., Wang, Q. and Rodriguez-Iturbe, I., 1987. Application of spatial poisson models to air mass thunderstorm rainfall. *J. Geophys. Res.*, **92**(D8), 9661–9678.

Eltahir, E.A.B. and Bras, R.L., 1993. Estimation of the fractional coverage of rainfall in climate models. *J. Climate*, **6**, 639–644.

Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intel.* **PAMI-6**, 721–41.

Gregory, D. and Smith, R.N.B., 1990. *Unified Model Documentation paper 25: Canopy, surface and soil hydrology, Version 1*, Meteorological Office, Bracknell, pp 19.

Henderson-Sellers, A., 1991. 'Incorporating' vegetation and soil schemes into Atmospheric General Circulation Climate Models. In: *Hydrological Interactions between Atmosphere, Soil and Vegetation* (Proc. of the Vienna Symposium, August 1991), IAHS Publ. No. 204, 11–29.

Henderson-Sellers, A., 1992. Assessing the sensitivity of a land-surface scheme to parameters used in tropical-deforestation experiments. *Quart. J. Roy. Meteorol. Soc.*, **118**, 1101–1116.

Isham, V., 1981. An introduction to spatial point processes and Markov Random Fields. *Int. Stats. Review,* **49**, 21–43.

Jennison, C., 1986. Contribution to the discussion of Professor Besag's paper. *J. Roy. Statist. Soc., Series B,* **148**, 288–9.

Jolley, T.J. and Wheater, H.S., 1996. A large-scale grid-based hydrological model of the Severn and Thames catchments. *J. Chart. Inst. Wat. Engrs. Managers*, **10**, 253–262.

Jolley, T.J. and Wheater, H.S., 1997a. An investigation into the effect of spatial scale on the performance of a one-dimensional water balance model. *Hydrol. Processes,* **11**, 1927–1944.

Jolley, T.J. and Wheater, H.S., 1997b. The introduction of runoff routing into large scale hydrological models. *Hydrol. Processes,* **11**, 1917–1926.

Kumar, P. and Foufoula-Georgiou, E., 1993. A new look at rainfall fluctuations and scaling properties of spatial rainfall using orthogonal wavelets, *J. Appl. Meteorol.*, **32**, 209–222.

Lovejoy, S. and Schertzer, D., 1990. Multifractals, Universality classes and satellite and radar measurements of cloud and rain fields, *J. Geophys. Res.*, **95**(D3), 2021–2034.

Manabe, S., Smagorinsky, J. and Strikler, R.F., 1965. Simulated climatology of a general circulation model with a hydrologic cycle. *Mon. Wea. Rev.*, **93**, 769–798.

Manabe, S., 1975. A study of the interaction between the hydrological cycle and climate using a mathematical model of the atmosphere. Proc. of the Conference on Weather and Food, Mass Inst Tech, Cambridge, Mass, pp 10.

Marsan, D., Schertzer, D. and Lovejoy, S., 1996. Causal space-time multifractal processes: predictability and forecasting of rain fields. *J. Geophys. Res.*, **101** (D21), 26333–26346.

Matsubayashi, U., Takagi, F. and Tonomwa, A., 1984. The probability density function of areal average rainfall, *J. Hydrosci. Hydraul. Eng.*, **2**, 63–71.

Oh, L., 1993. *Analysis of rainfall disaggregation in GCM*, MSc Thesis, Imperial College.

Onof, C. and Wheater, H.S., 1996a. Analysis of the spatial coverage of British rainfall fields, *J. Hydrol.*, **176**, 97-113.

Onof, C. and Wheater, H.S., 1996b. Modelling of the time-series of spatial coverages of British rainfall fields, *J. Hydrol.*, **176**, 115–131.

Onof, C., Mackay, N., Oh, L. and Wheater, H.S., 1997. An improved rainfall disaggregation technique for GCMs. *J. Geophys. Res.*, **103** (D16), 19577–19586.

Over, T. and Gupta, V., 1996. A space-time theory of mesoscale rainfall using random cascades. *J. Geophys. Res.*, **101** (D21), 26319–26331.

Perica, S and Foufoula-Georgiou, E., 1996. A model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *J. Geoph. Res.*, **101** (D21), 26347–26361.

Ripley, B.D., 1988. *Statistical Inference for Spatial Processes.* Cambridge University Press.

Rowntree, P.R. and Bolton, J.A., 1978. Experiments with soil moisture anomalies over Europe. In: (Asselin, R.) *The GARP Programme on Numerical Experiments: Research Activities in Atmospheric and Oceanic Modelling*, GARP, Report no. 178, August 1978, pp 63.

Sellers, P.J., Mintz, Y., Sud, Y.C. and Dalcher, A., 1986. A simple biosphere model (SiB) for use within general circulation models. *J. Atmosph. Sci.*, **43**, 505–531.

Smagorinsky, J., 1963. General circulation experiments with primitive equations, I. The basic experiment. *Mon. Wea. Rev.*, **93**, 99–164.

Thompson, N., Barrie, I.A. and Ayles, M., 1981. *The meteorological office rainfall and evaporation calculation system: MORECS (July 1981)*, Hydrological Memorandum No. 45, Meteorological Office, Bracknell, pp 72.

Walker, J. and Rowntree, P.R., 1977. The effect of soil moisture on circulation and rainfall in a tropical model. *Quart. J. Roy. Meteorol. Soc.*, **103**, 29–46.

Warrilow, D.A., Sangster, A.B. and Slingo, A., 1986. *Modelling of land surface processes and their influence on European climate.* Dynamical Climatology Technical Note No. 38, Meteorological Office, Bracknell, UK, pp 92.

Wheater, H.S., Jakeman, A.J. and Beven, K.J., 1993 Progress and directions in rainfall-runoff modelling. In: (A.J. Jakeman, M.B. Beck & M.J. McAleer, Eds.) *Modelling Change in Environmental Systems*, Wiley, Chichester, UK, pp 101–132.