Advances in Water Resources 25 (2002) 1313–1334

Advances in Water Resources

www.elsevier.com/locate/advwatres

Advances in the use of observed spatial patterns of catchment hydrological response

Rodger B. Grayson [a,*], Günter Blöschl [b], Andrew W. Western [a], Thomas A. McMahon [a]

[a] *Department of Civil and Environmental Engineering, Cooperative Research Centre for Catchment Hydrology and Centre for Environmental Applied Hydrology, University of Melbourne, Melbourne 3010, Australia*

[b] *Institut für Hydraulik, Gewässerkunde und Wasserwirtschaft, Technische Universität Wien, Wien 1040, Austria*

Received 26 September 2001; received in revised form 22 April 2002; accepted 22 April 2002

Abstract

Over the past two decades there have been repeated calls for the collection of new data for use in developing hydrological science. The last few years have begun to bear fruit from the seeds sown by these calls, through increases in the availability and utility of remote sensing data, as well as the execution of campaigns in research catchments aimed at providing new data for advancing hydrological understanding and predictive capability. In this paper we discuss some philosophical considerations related to model complexity, data availability and predictive performance, highlighting the potential of observed patterns in moving the science and practice of catchment hydrology forward. We then review advances that have arisen from recent work on spatial patterns, including in the characterisation of spatial structure and heterogeneity, and the use of patterns for developing, calibrating and testing distributed hydrological models. We illustrate progress via examples using observed patterns of snow cover, runoff occurrence and soil moisture. Methods for the comparison of patterns are presented, illustrating how they can be used to assess hydrologically important characteristics of model performance. These methods include point-to-point comparisons, spatial relationships between errors and landscape parameters, transects, and optimal local alignment. It is argued that the progress made to date augers well for future developments, but there is scope for improvements in several areas. These include better quantitative methods for pattern comparisons, better use of pattern information in data assimilation and modelling, and a call for improved archiving of data from field studies to assist in comparative studies for generalising results and developing fundamental understanding.

1. Introduction

In 1995, Hornberger and Boyer finished their report to the IUGG on recent advances in watershed modelling, 1991–1994, by saying:

> "It may seem strange to end a review of modelling with an observation that future progress is very strongly linked to the acquisition of new data and to new experimental work, but that, in our opinion is the state of the science."

What Hornberger and Boyer [50] were responding to was the fact that, while the technology of modelling had advanced over the previous few years, the fundamental development of modelling ideas had "run out of steam". There had been some technical developments in relation to the use of terrain data and development of distributed models (as well as improvements in the availability of remote sensing and chemical data) but most of these had been in the application of ideas that had existed for some time, their implementation being made possible by the ever decreasing cost of computing power. Beven and Moore [14] and Rosso [87] detailed some of these advances and explored the issues associated with what was then a relatively new technology of distributed hydrological models. But in essence, the problems that Hillel [49] and Klemeš [63] had noted almost a decade earlier—that progress was being stymied by a lack of appropriate data—were still very real.

What brought us to this point in the early 1990s where there was an explosion in the development and use of spatially explicit hydrological models, which has continued to today? There are at least two answers. The first is technological—the ever increasing availability of digital elevation data, GISs to manipulate spatial data

* Corresponding author.
E-mail address: rodger@civag.unimelb.edu.au (R.B. Grayson).

0309-1708/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0309-1708(02)00060-X

of all sorts, and the decreasing cost of computing power. The second is the rise in environmental awareness of the broader community and its subsequent impact on research into, and the management of, natural resources. We now want to know not only what is the quantity and quality of water in a stream, but also from *where* any contaminants came and *where* best to invest scarce financial resources to help rectify the problem. We now need predictions of the hydrological (and ecological) impacts of land use and climate change—predictions that must at least account for the *spatial variability* we see in nature, and more often provide *spatial estimates*, if they are to be of any practical use. Natural resource agencies have been amassing large amounts of spatial data to complement the temporal data traditionally measured, and are eagerly looking to use this for predictive spatial modelling of environmental response. In principle, we have the tools available to undertake this work and already, the spatial models and impressive colour graphics that our GISs generate are attractive to politicians and administrators [45]. But, as Hornberger and Boyer pointed out, there are significant problems with the representation of scale and heterogeneity, with the extent to which some of these data sets and models are appropriate for hydrological applications, and consequently with the scientific credibility of predictions.

In recent times there have been a number of general discussions on distributed catchment modelling, including applications to practical problems [1] and a comprehensive assessment of the current state of the art [8]. But these have served more to consolidate the work of the 1990s and propose new methodological advances, rather than focus on new data sources. Nevertheless, the calls of the 1980s and early 1990s for more research into representing spatial heterogeneity, the collection of data sets for the testing and development of distributed models, and methods to how best deal with issues of scale, have to some extent been heeded and it is these on which we will focus in this paper. Specifically, we will explore the use of spatial patterns of catchment response that have arisen from detailed field studies for (i) describing the structure of heterogeneity and (ii) advancing the testing and development of distributed hydrological models.

We begin by briefly discussing some advances in philosophical issues related to modelling, and a framework for how to deal with the scaling issues that arise in using data and models to understand processes, as well as for predictive purposes. We then describe, through example, the types of pattern information that are available to the modeller and some general considerations in their use. This is followed by a summary of recent results related to characterising the structure of heterogeneity that are based on recently completed field studies. We then briefly present three case studies from very different hydrological environments where pattern information was used to test distributed models and summarise what was learned from the use of observed spatial patterns. Finally we discuss pattern comparison techniques and propose some new approaches and avenues for further research that we believe will underpin further advances in both our understanding of spatial hydrological response and our predictive ability.

2. Advances in philosophy

The debate that had raged in the 1980s and early 1990s about the relative merits and capabilities of physically based and conceptual models has largely subsided. Problems with the use and interpretation of distributed models are better understood [1,10]) and the formalised testing procedures proposed for such models (e.g. [84,85]) have illustrated that the earlier tendency to believe the distributed output from models, having tested only an integrated value such as runoff, was unsound. This has placed greater importance on the collection and analysis of spatial response data to improve model performance.

Fig. 1 (after Grayson and Blöschl [46]) illustrates the conceptual relationship between model complexity, the availability of data for model testing, and predictive performance of the model. We use the term "data availability" to imply both the amount and quality of the data in relation to its use for model testing. Having data on spatial response is equivalent to "large" availability while just runoff data would imply "small" availability. We use the term "model complexity" to mean detail of process representation. Complex models include more processes and so are likely to have more parameters.

If we have a certain availability of data (e.g. solid line in Fig. 1), there is an "optimum model complexity" beyond which the problems of non-uniqueness become important and reduce the predictive performance. There

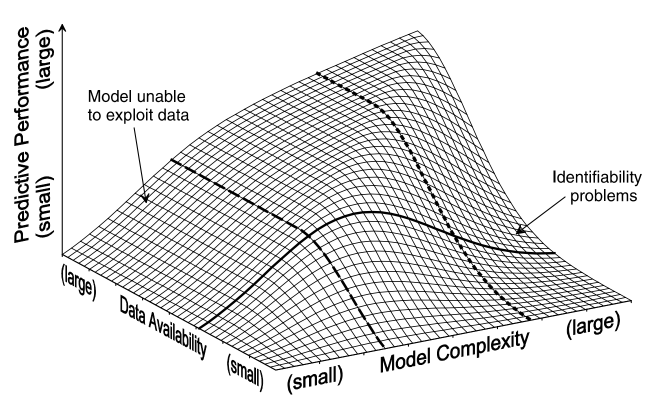


Fig. 1. Schematic diagram of the relationship between model complexity, data availability and predictive performance (after Grayson and Blöschl [46]).

are too many model parameters and not enough data to test whether the model is working, or is working for the right reasons, which means that both the model structure and the model parameters cannot be identified properly. We can use a simpler model than the optimum, but then we will not fully exploit the information in the data (e.g. intersection of solid and dashed lines). For given model complexities (e.g. dashed and dotted lines), increasing data availability leads to better predictive performance up to a point, after which the data contains no more "information" to improve predictions i.e. we have reached the best a particular model can do and more data does not help improve performance (the dashed and dotted lines flatten out as data availability increases). In this case, we could consider a more complex model to better exploit the information in the data. The more common situation for practical applications of distributed modelling is represented by the intersection of the dotted and solid lines, where we are using too complex a model with limited data and so have identifiability problems. Increased data availability is needed to significantly improve the predictive performance.

In practice, the "optimum model complexity" implied in Fig. 1 does not mean an optimum model or parameter set. Beven [13] introduced the term "equifinality" to describe the situation where there are a large number of models or parameter sets that equally well describe some given data. He argues that equifinality is endemic in environmental modelling and that the only way to sensibly deal with it is to explicitly account for the uncertainty that it introduces, and to focus attention on seeking data that enables rejection or falsification of models on the grounds that they fail to represent observed behaviour of the system [7,9,10]. Beven uses the terms "behavioural" and "non-behavioural" (following the terminology of Hornberger and Spear [51,52]) to describe models that "fit" or "do not fit" with observations. If there are few observations, many models will be "behavioural" (since they will not be falsified by the data), but there will be a great deal of uncertainty associated with their predictions. We believe that observed spatial patterns of hydrological response can be a powerful discriminator of "behavioural" and "non-behavioural" models. But to compare observed and simulated patterns, we must be sure that they are representing the same thing—not only the feature of interest, but also that the scales of observation and simulation match, or if they do not match, the effects of this mismatch need to be defined. These are not trivial problems and some significant conceptual advances have been made to assist in dealing with these scale issues.

A persistent problem in hydrological modelling in general has been how to deal with the different scales (in both space and time) on which processes operate, data are available and models are formulated. Blöschl and Sivapalan [16] proposed a framework for considering these issues. They introduced the notion of a "scale triplet" defined by spacing, extent and support. Spacing refers to the distance (or time) between samples, extent refers to the overall coverage of the data (in time or space), and support refers to the averaging volume or area (or time) of the samples. All three components of the scale triplet are needed to uniquely specify the space and time dimensions of measurements of a pattern. A similar scale triplet can be defined for a model. Blöschl and Sivapalan [16] and Blöschl [17] then examined how the measurement spacing, extent and support will change the true pattern to be reflected in the data; and how will the model spacing, extent and support change the data to be reflected in the predictions. The basic idea is that there is some similarity between these two steps. Generally, if the spacing of the data is too large, the small-scale variability will not be captured. If the extent of the data is too small, the large-scale variability will not be captured and will translate into a trend in the data. If the support is too large, most of the variability will be smoothed out. It is clear that some sort of filtering is involved, i.e. the true patterns are filtered by the properties of the measurement that are reflected in the data. These scale effects can be quantified (see [6,17,36]) and, maybe more importantly, a qualitative use of these ideas provides a guideline for how to best match the scales of observed and simulated patterns, and if this match can not be achieved what will be the effect of the scale mismatch on the pattern comparison.

3. Types of pattern information

In this section we provide an overview of the types of information that are available to modellers and some general considerations in their use. We can identify at least three distinct types of pattern information (i) "lots of points" (LOP) where there is a sufficiently dense array of point measurements to be interpreted as a pattern; (ii) binary data and (iii) surrogate data. Each will be described through examples from field studies.

3.1. Lots of points

Developments in measurement technology have enabled the rapid measurement of a number of variables of hydrological interest that in earlier days had been time consuming. Volumetric soil moisture as measured by time domain reflectrometry (TDR) is a classic example where it now takes only seconds to measure moisture over a depth of interest. These devices can be placed in situ to enable detailed temporal patterns of soil moisture to be obtained, or combined with GPS technology and an all terrain vehicle with hydraulic insertion of the TDR probes, to produce a tool for the rapid collection of spatial soil moisture patterns over small catchments (e.g. [96]). This approach has been used at a number of locations including the 10 ha Tarrawarra catchment in Australia [110], the 50 km^2 Mahurangi catchment in New Zealand [121], at Nerrigundah in Australia [108], at Weiherbach in Germany [4,24,69], at Dartmoor, England [76] and sampling associated with the Southern Great Plains experiments in the United States [77]. Also, the general interest in spatial patterns, and availability of rapid position fixing via GPS, has made it more common to collect patterns of even "traditional" data such as snow depth or snow density. For example, the 26 ha upper sheep creek in the Reynolds creek watershed has snow depth patterns measured over several seasons [71,72] and Yang and Woo [122] and Young et al. [123] collected patterns of snow-related variables in the Canadian Arctic. Along similar lines, McDonnell et al. [73] measured a detailed pattern of soil depth to bedrock in the 41 ha Panola catchment for use as input data to a terrain-based model of water movement. While there have been no major advances in the rapid measurement of soil hydraulic properties, it has still been possible to get LOP patterns in small catchments such as the 0.75 La Cuenca [31] or the 10 ha R5 at Chickasha [70,91].

Key considerations in the use of LOP patterns are whether there are really enough measurements to justify a pattern, and how representative the point measurement is of a larger area (i.e. how the support for the measurement relates to the support of the model, and how the measurement error compares to any underlying pattern in the field). There is no direct answer to the question of "how many points constitutes a pattern?" since it depends on the scale triplet of the measurement in relation to the process affecting the measurements and on the measurement error characteristics in relation to the variability of the field being measured, but one answer is "sufficient to infer the spatial structure", although this itself is open to interpretation (see later discussion on uncertainty in variograms). Grayson and Blöschl [47] discuss several LOP patterns where more than 100 points are generally available. The question of "representativeness" also relates to the scale triplet, and to the "signal-to-noise" ratio of the measurement field. It is common for some form of filtering of field data to be undertaken so that the scale of measurement better matches the model to which it is being compared and to deal with the measurement error. This filtering generally takes the form of interpolation methods of varying complexity [4,22,56]. Simple interpolation methods generally fail to distinguish between features of patterns that are hydrologically significant, and "noise". More sophisticated methods such as co-kriging and external drift kriging [4,22,42] are improvements, but require some understanding of the physical significance of underlying patterns to provide sound results. One approach to filtering that takes account of hydrologically significant patterns is to correlate the patterns with a deterministic measure such as terrain characteristics. This deterministic pattern is then removed from the data, the residuals smoothed (since these are more normally distributed and spatially random, choice of method is less critical) and the smoothed residuals added back to the deterministic pattern. The residuals represent a combination of the measurement error and the "sub-grid" variability resulting from the support of the measurement being smaller than the model grid size. This approach has been used with rainfall, where elevation is the most important terrain feature (e.g. [43]) and soil moisture, where a compound terrain index was used [111].

3.2. Binary patterns

The most widespread binary patterns in hydrology are probably snow cover, derived from aerial photographs (e.g. [18]) or from satellite remote sensing (e.g. [26,27,86,119]).

Similarly the inundation area of floods can be viewed in photographs and various satellite sensors including high-resolution synthetic aperture radar (SAR) (e.g. [53]) or high-resolution optical instruments (e.g. LandSat ETM+), or for large areas lower resolution optical instruments (e.g. AVHRR [81]; see also Jensen and Calabresi [60], for examples from a range of platforms). A binary pattern of particular interest to catchment hydrologists is the pattern of saturated areas, however observations of these have been relatively rare, despite their obvious utility [12]. The pioneering work of Dunne [29,30] mapped saturated areas in the field and in very recent times, their utility has been "re-discovered", with a number of mapping projects occurring (e.g. [37,61]). These ideas have been extended by Peschke et al. [82] who mapped the *type* of runoff generation mechanism that occurs for a given catchment state, based on many years of field investigation in the 4.6 km^2 Wernersbach catchment in Germany. The mechanisms identified by Peschke et al. [82] as binary patterns of presence/absence, were Hortonian overland flow, saturation excess overland flow, interflow, recharge and storage.

Binary patterns are a classic example of a trade off between detailed patterns and information in an individual measurement. The detail illustrated in Fig. 2 is clearly of great diagnostic use in comparing model performance, still it says little about the actual volume of a snow pack. Binary patterns are therefore particularly useful when used in conjunction with other data. For example, snow cover patterns are complementary to the measurement of discharge due to snow melt (e.g. [19]). The latter provides a quantitative measure of the change in water storage while the former indicates from

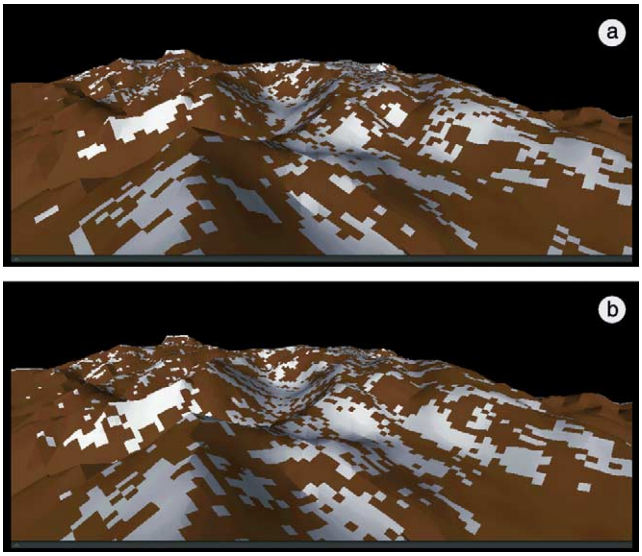


Fig. 2. Snow cover patterns for the Schneealpe region of the Austrian Alps, May 3, 2000. Brown is no snow, white is snow cover, (a) observed from SPOT interpretation and (b) simulated by VUTS.

where it has originated. Similarly saturated areas, and catchment discharge are complementary measures.

3.3. Surrogate patterns

Surrogate patterns are perhaps the most widely available spatial data sources for catchment hydrology. Surrogates are variables that show some (often limited) degree of correlation to the pattern of interest but are much easier to collect in a spatially distributed fashion. Examples of surrogates include terrain [116], soil texture to infer hydraulic properties (e.g. [34,83]) and remote sensing data (where various characteristics of emission or reflectance of radiation of different wavelengths are correlated with features of interest such as vegetation cover, surface temperature, or soil moisture). Terrain-based wetness indices can also be thought of in this context. Strictly speaking, virtually all measurements are surrogates, but here we are referring to those where the correlations are not extremely high. Like binary patterns, there is often a trade off between the detail of the pattern, and the correlation with the variable that is being modelled.

Troch et al. [95] noted that analysis of individual SAR scenes from operational satellites were too poorly correlated with surface soil moisture to be a useful surrogate, but Verhoest et al. [100] and Gineste et al. [41] showed that multitemporal analysis of many SAR images could provide information on areas of the landscape that were continuously wet, providing a useful surrogate for saturated areas. This is an excellent example of where substantial analysis of pattern data was needed to derive a usable surrogate pattern. In other cases, models may need to be reformulated to be more directly comparable to the surrogate measure. An example is land surface schemes where the intention is to compare satellite-derived "skin" soil moisture or temperature with modelled values. Here the models must be formulated with a shallow "skin" layer (e.g. [66]) rather than the more common situation in catchment hydrology of a depth integrated perhaps over the root zone.

The alternative to altering model structure to match the observation is obviously to seek observations that match the model structure. An interesting example of this approach comes from the observations of Tóth [94] as applied to the equilibrium modelling of Salvucci and Entekhabi [88] and Salvucci and Levine [89]. Tóth made field observations of natural time-integrators of subsurface flow conditions (e.g. presence of salt precipitates) to provide a mix of quantitative and observational surrogates for recharge and discharge, enabling a spatial map of long-term recharge and discharge locations to be derived. Such a pattern was ideally suited to comparison with the equilibrium model [89] for which standard measures such as time series of groundwater bores were unsuitable due to limited temporal extents.

4. Characterising structure and heterogeneity

There are a number of ways in which one can generalise the detailed information obtained from experimental studies designed around the collection and analysis of spatial patterns of hydrological response. Ideally one would like to obtain relationships that, ultimately, allow inference of spatial processes in a particular catchment without the need for detailed spatial observations. If it is possible to obtain an accurate description of the spatial structure of, say, soil moisture for a given climate, catchment state, geology, geomorphology, etc. one can then hope to apply this description to other catchments with similar conditions. Potential applications are designing experimental setups, interpolating sparse point data, and defining the structure of dynamic catchment models in terms of the representation of small-scale variability. In this section, we illustrate some recent advances in the characterisation of structure and heterogeneity using data on spatial patterns, in this case focussing on soil moisture.

A starting point is an examination of the spatial *statistical* structure. Very detailed analyses of the TDR soil moisture patterns collected in the 10.5 ha Tarrawarra catchment in south-eastern Australia allowed the derivation of variograms with a high degree of reliability that would not have been possible with a few point data [114]. The data suggested that exponential variogram models, including a nugget, fitted the soil moisture data variograms closely. The geostatistical structure was found to evolve seasonally. High sills (15–25 $(\%V/V)^2$) and low correlation lengths (35–50 m) were observed during the wet winter period. During the dry summer period sills were smaller (5–15 $(\%V/V)^2$) and correlation lengths were longer (50–60 m). This seasonal evolution is explained on the basis of the importance of lateral redistribution of moisture during different seasons. The availability of LOP patterns as opposed to a few point data also enabled an analysis of the reliability of soil moisture variograms. Western et al. [114] sub-sampled transects from typical summer and winter soil moisture patterns containing over 500 individual measurements (Fig. 3a from (i) to (iv)) variograms are based on 44, 86, 164 and 296 points respectively (giving 16, 8, 4 and 2 estimated variograms respectively). i.e. if we had taken just one sample, any of the variograms could have been obtained. These comparisons indicate that very detailed data sets (more than about 150 points in space) are required to reliably estimate variogram parameters that are representative of the landscape. This adds a caveat to the interpretation of variograms used for characterising spatial structure and heterogeneity. It also imposes practical limitations on the use of geostatistical methods such as kriging which hinge on a reliable estimation of

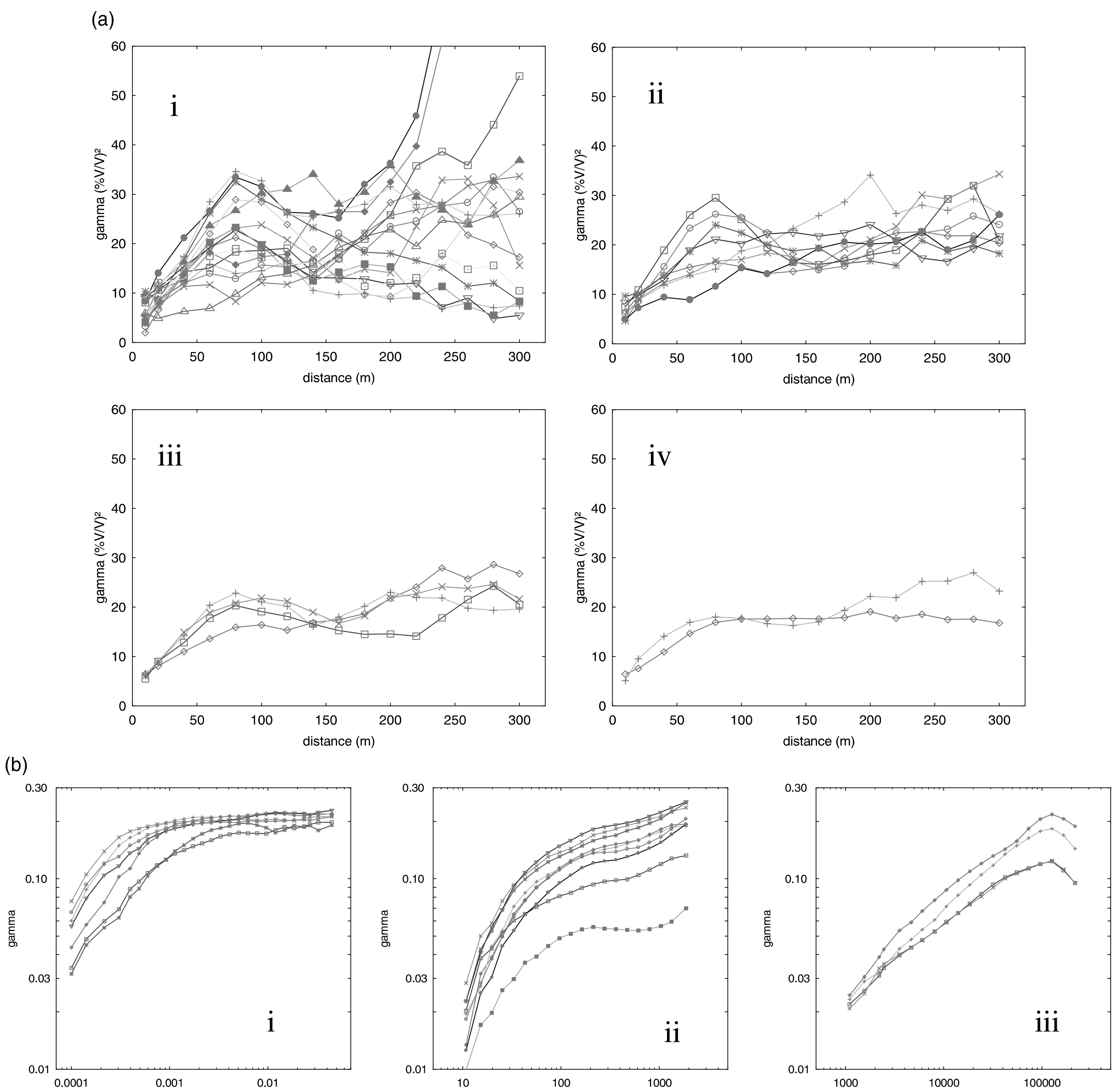


Fig. 3. (a) Effect of sample size on the reliability of soil moisture variograms from Tarrawarra, May 2, 1996. (i) Based on 44 points, (ii) based on 86 points, (iii) based on 164 and (iv) based on 296 data points. (b) Variograms of snow covered area, i.e. a binary variable that is 1 for a snow covered pixel and 0 for a snow free pixel. (i) Thin sections of snow (eight images for different snow types, pixel size is 0.1 mm); (ii) Kühtai aerial photographs (nine scenes in 1989, pixel size is 5 m); (iii) Sierra Nevada AVHRR images (four scenes in 1998, pixel size is 1100 m). (c) Connectivity functions calculated for indicator data from the Tarrawarra soil moisture patterns. (i) Seventy-fifth percentile indicator patterns for two occasions, pixel size is 10m × 20 m (ii) Connectivity functions $\tau(h)$ for each occasion, (iii) normalised variogram for each occasion.

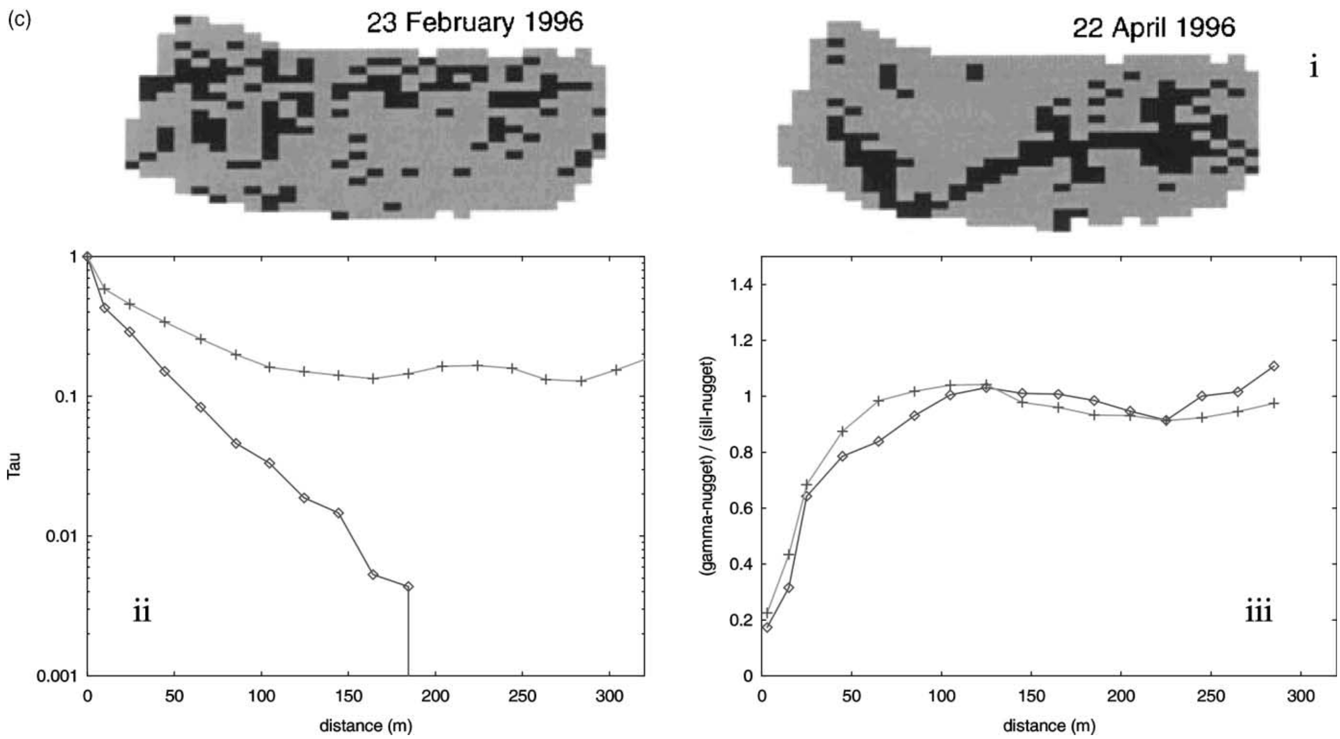


Fig. 3 (*continued*)

the variogram, particularly of its correlation length (see e.g. [57]).

It is interesting to compare the spatial correlation found in the Tarrawarra catchment with correlations obtained in other catchments around the world. Six small catchment studies reviewed in Western et al. [112,114] suggest that values of the correlation length vary between 1 and 600 m and there is a tendency for correlation length to increase with the catchment scale. Studies at much larger scales (50–1000 km) [35,105] from agricultural sites in the Former Soviet Union, Mongolia, China, and the USA have found that soil moisture variation could be represented as a stationary field with a correlation length of about 400–800 km. Vinnikov and Robock [105] and Entin et al. [35] noted the existence of a smaller scale (<50 km) component to the spatial variability that was unresolved by their data.

This scale dependence of the correlation length is illustrated in Fig. 3b for the case of snow patterns from three very different case studies [17]. Different lines in Fig. 3b relate to different dates. The first set of variograms (Fig. 3b(i)) was derived from a number of thin sections obtained in the laboratory by scanning images of snow crystals, i.e. binary images where 1 is ice and 0 is void. The correlation length (i.e. the scale where the variograms flatten out) is on the order of 0.5 mm. The second set of variograms (Fig. 3b(ii)) was derived from aerial photographs of snow cover in the Kühtai catchment, Austria, i.e. binary images where 1 is snow and 0 is no snow [18]. Correlation lengths are on the order of 100 m. The third set of variograms (Fig. 3b(iii)) was derived from snow cover based on AVHRR images in the Sierra Nevada region (K. Elder, personal communication). Correlation lengths are on the order of 30 km. While the variograms in Fig. 3b do not apply to the same date and the same location it is reasonable to assume that their general shape will be similar for other dates and locations. Their main difference then is the scale at which the snow cover data have been collected. This example from snow hydrology clearly illustrates that descriptions of spatial correlation lengths are conditional on the measurement scale. In practice this means that the results from particular studies may be applicable to other studies of similar scale, but cannot be transferred to different scales without correcting for the measurement effects (e.g. by regularisation [109]) and in cases where the scales are very different, this may not be possible.

While spatial correlations are one important way of characterising structure and heterogeneity, it is not necessarily the most meaningful in a hydrological context. Ideally one would like to capture those features that are most relevant to the movement of water on the surface and in the sub-surface. Western et al. [113] argued that connected features such as high conductivity preferred flow paths in aquifers and saturated source areas in drainage lines control the lateral movement of water, so it is the connectedness that needs to be represented by statistical measures of heterogeneity. They tested the utility of connectivity functions of Allard [3], Allard and Group [2], Gould and Tobochnik [44], and Stauffer and Aharony [93] on thirteen observed soil moisture patterns from the Tarrawarra catchment and two synthetic aquifer conductivity patterns. The connectivity function applies to indicator values, Z, which are binary patterns obtained by thresholding the original pattern. $Z = 1$ if the original value is above the threshold, and $Z = 0$ otherwise [115]. This connectivity function is described by

$$\tau(h) = P(x \leftrightarrow x + h | x \in A, x + h \in G) \quad (1)$$

where G is the set of pixels making up the spatial pattern and A is the set of pixels in G with $Z = 1$. Two pixels, x and $x + h$ in A are connected (denoted by $x \leftrightarrow x + h$) if there is a continuous path of neighbouring pixels belonging to A ($x_1, \ldots, x_n \in A$) between them. The connectivity function $\tau(h)$ represents the lag dependent probability that a pixel (x) in A, is connected to any pixel ($x + h$) in G, that is separated from x by the Euclidian distance h.

The potential of this connectivity function is illustrated in Fig. 3c. Fig. 3c(i) shows two soil moisture patterns from Tarrawarra, a disconnected pattern (left, 23 February, 1996) and a connected pattern (right, 22 April, 1996). These are indicator patterns at a 75% threshold, i.e. pixels with soil moisture larger than the 75% percentile are dark (i.e. $Z = 1$) while pixels with soil moisture smaller or equal the 75% percentile are grey (i.e. $Z = 0$). Fig. 3c(ii) gives the connectivity function calculated from these patterns by applying Eq. (1). For the disconnected pattern, the connectivity function drops quickly with distance while for the connected pattern, the connectivity function remains at around 0.2. The connectivity function can be interpreted as the percentage of pixels that are connected over a certain distance. This means that, say at a distance of 150 m, 20% of the pixels are connected in the connected pattern case while only 1% of the pixels are connected in the disconnected pattern case. These differences will clearly be important for lateral flow and transport processes in catchments. For comparison, Fig. 3c(iii) shows the variograms calculated for the same soil moisture patterns, which have been normalised by their sill (i.e. variance). These variograms are very similar to each other, which illustrates that the connectivity functions are able to distinguish between connected and disconnected patterns while variograms cannot. Western et al. [113] provide a wider number of examples for the potential of the connectivity function approach. They discuss potential applications in hydrology such as in interpolation and stochastic simulation and for the derivation of bulk parameters to characterise hydrologically relevant spatial characteristics of patterns.

These advances in the quantification of spatial structure and heterogeneity in soil moisture have been made possible only because LOP field observations were sufficiently dense. As more data sets of observed spatial response become available, we anticipate further advances in this area, and the development and testing of better tools for the representation of spatial structure in hydrological applications.

5. Use of patterns with distributed modelling

There are several uses to which observed spatial patterns can be put in distributed modelling. The most obvious is as inputs such as precipitation, and for model testing such as comparison of simulated and observed snow cover. There is a more recent application in catchment hydrology inspired from the meteorological community, which has a long tradition of using spatial data (albeit at a very large scale) to assimilate into models such as operational weather prediction models. With the rapid increase in remote sensing data of relevance to hydrology, there is great scope for increasing use of data assimilation (DA) methods with hydrological models [55,74,98]. This is already being done at the continental and global scale through projects such as LDAS (Land Data Assimilation Schemes [127]), focussing on use of remote sensing data with models of the land surface–atmosphere interaction. Smaller scale applications do not seem to be operational yet but there appears to be some scope in the near future. DA seeks to update modelled states (and inputs) with observations and so improve the predictive ability of the model by being more sure about initial conditions. For hydrological applications where runoff estimates at the catchment outlet are sought, the complexity of DA techniques may mean that their use will be limited, but for applications where users are interested in output patterns, the reduction in uncertainty of initial conditions may justify their use. Analysis of the degree of updating required can also help identify problems with the model, although this has generally not been the primary objective of DA.

Spatial patterns have been used more directly in the development, calibration and testing of distributed models, particularly in snow modelling, and for research catchments. In an early application, Blöschl et al. [19] used photographs of snow cover to assess the performance of a spatially distributed energy and water balance model of the snowpack. Along similar lines, Wigmosta et al. [119] and Davis et al. [125] have used snow patterns in analyses of alpine hydrology models. Moore and Grayson [126] compared observed saturated source areas to simulations from a distributed parameter model in a small laboratory sand bed, while Whelan and Anderson [118] simulated the spatial variability of throughfall and compared it to measurements from an array of ground collectors. As noted above, recent years have seen an increase in the availability of ground-based pattern data from research catchments and from remote sensing, up to the global scale. This has led to many more examples of using patterns for developing and testing distributed models. For example, in the Peschke et al. [82] study mentioned above, detailed observed patterns of runoff process types were used to test an expert system for the identification of dominant hydrological processes, based on catchment characteristics and state variables as inputs. Kite and Droogers [62] report a number of studies related to observations and modelling of evapotranspiration. Summaries of several recent case studies are reported in Grayson and Blöschl [46,47], where examples of catchments up to 10 km^2 are discussed, and Lakshmi et al. [65], who focus on larger scale applications. It is important to note that in most of these applications, the reason for modelling has not been just in estimating runoff, but rather in estimating the details of the spatial hydrological response of catchments. In the early days of distributed models, many users assumed that good representation of an output hydrograph implied that the distributed model predictions were similarly good [20]. It is now more generally accepted that if we are interested in spatial model output, we need to test the models with spatial information.

In this section, we present examples from three case studies where pattern data has been collected and used to develop and test models of distributed catchment response. The studies were chosen to cover different data types (surrogate, binary, LOP) and different hydrological settings (alpine, tropical and temperate). Only brief descriptions of the case studies are provided here since the purpose of this section is to illustrate what can be learned through the use of patterns rather than pursue the nuances of the modelling exercises.

5.1. Snow patterns at Schneealpe

The first example comes from the work of Jansa et al. [58] and Blöschl et al. [21] who applied the Vienna University of Technology Snow model (VUTS) to the 90 km^2 Schneealpe area in the Austrian Alps. VUTS is a terrain based distributed model that couples heat and mass flow within the snowpack using a multilayer model at the grid scale [19]. Atmospheric data used to drive the model include global radiation, air temperature, humidity, wind speed and precipitation on an hourly basis. Horizon shading, and aspect and slope dependence of solar radiation input are accounted for and wind drift is represented by a factor of the form:

$$F = (a + bH)(1 - f(S))(1 + eC) \geqslant 0$$

$$f(S) = \begin{cases} 0 & S < c \\ \frac{S-c}{d-c} & \text{otherwise} \end{cases}$$

where H is elevation, S is slope and C is terrain curvature at the grid scale of the digital elevation model, a and b are factors accounting for the elevation effect on precipitation and derived from analysis of local data, c, d and e are calibration factors derived from analysis of cover patterns [19]. A discussion of this approach is given in Moore et al. [78].

Snow cover patterns for the years 1998–2000 were derived from analysis of SPOT XS based on an unsupervised isodata technique [59] separately for different illumination classes. The classification produced 3-class patterns of snow, no-snow and partial coverage. VUTS was initially run with a standard parameter set and while simulations were generally good, it appeared that improvements could be gained by refinement of the drift factor. The parameters in the drift model are generally calibrated globally, however the availability of detailed spatial patterns made a more sophisticated approach possible. Four SPOT-derived cover patterns were chosen from each of the years 1998 and 1999. These patterns were carefully chosen to span a wide range of overall snow coverages. These eight patterns were compared to the simulated patterns for the same dates to derive a combined error map of cover for these two years (Fig. 4a). This map represents the average percent error in snow cover estimation and it was assumed that this error stemmed largely from the representation of snow drift. The parameters of the snow-drift model were then individually calibrated for each pixel to minimise the error. Fig. 4b shows the error pattern for the calibrated model and illustrates the significant improvement over the pattern in Fig. 4a. Remaining error is due to sources other than wind drift, or that cannot be explained by the structure of the wind drift model in VUTS.

VUTS was then applied to a different snow season (2000) in a classical split sample test. Fig. 4c shows the error map for this year and illustrates that, while there is some deterioration in simulation compared to the fully calibrated year, the revised snow-drift model is a major improvement over the original form (Fig. 4a). Thus it can be concluded that to a large extent, the pattern of wind drift is "time stable" i.e. that from season to season, the pattern of drift is very similar, and that accounting for the spatial structure significantly improves the simulated patterns. Of course it is still likely that some structural errors remain, but the testing would indicate that we can be more confident in the results when using the revised parameter set. This is not always the case when "local" calibration of this sort is done. Lamb et al. [67,68] used spatial patterns of groundwater levels to locally calibrate a version of TOPMODEL but the uncertainty in catchment runoff was little affected due to problems with the structure of the model, i.e. the local calibration partially overcame some of the structural problems with the model to better

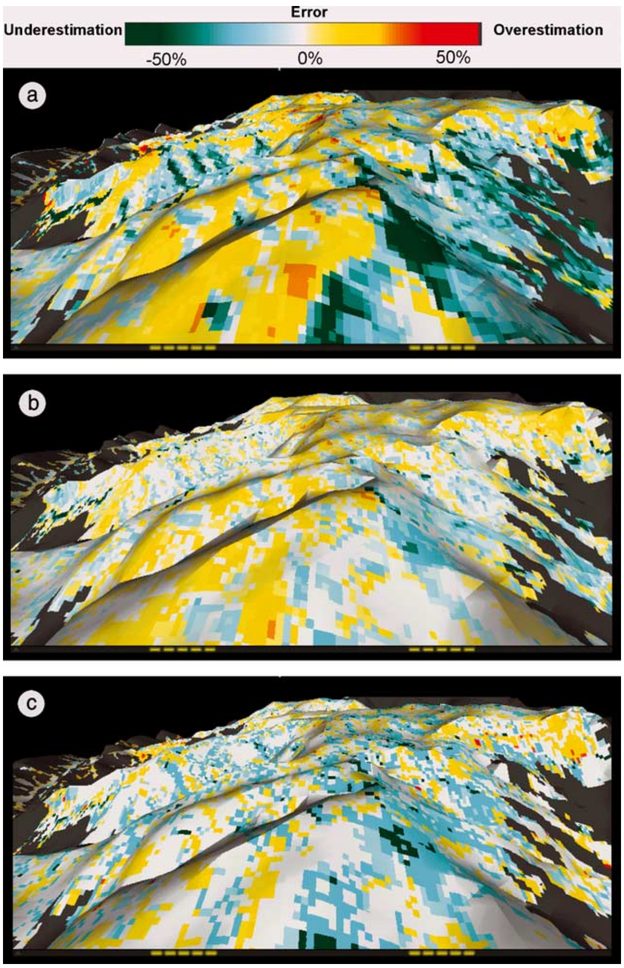


Fig. 4. Bias in snow cover patterns based on residuals of comparisons between observed (SPOT) and simulated (VUTS) patterns for the Schneealpe region. (a) Pre-calibration using eight observed patterns in 1998–1999, (b) post-calibration using eight observed patterns in 1998–1999, (c) validation using four observed patterns in 2000 (after Blöschl et al. [21]).

fit internal patterns, but did not get better runoff estimates. In the Schneealpe, example, it could be similarly argued that the local calibration of parameters partly overcomes a structural problem, but in this case, the main interest was in simulating spatial patterns rather than in bulk catchment runoff. While the use of observed patterns had a minor effect on bulk runoff from the area, it clearly improved the simulations of spatially distributed snow water equivalent and snow melt.

Analysis of the snow patterns also made it possible to separate the effects of individual model parameters. For example it was possible to unravel the effects of albedo and threshold temperature on snow melt processes by analysis of the cover statistics for slopes of different aspects. This separation of effects is not possible to do using just runoff data. It should be noted that a similar approach was taken by Luce et al. [71] in the rangelands of Idaho, where observed snow depth patterns were also shown to be time stable.

5.2. Runoff patterns at La Cuenca

La Cuenca is a small (0.75 ha) catchment in the Rio Pichis valley of central Peru. It is covered by rainforest vegetation and was set up as a research catchment for the study of spatial variability in soil properties and runoff processes in a tropical environment [31–33,128]. Details of the example presented here can be found in Vertessy et al. [103,104] but a brief description is given below.

Over 700 undisturbed soil cores were used for measurements of soil hydraulic properties of the three dominant land units in the catchment. In addition, 72 overland flow detectors were installed across the catchment and checked after more than 180 individual storms, to indicate the spatial pattern of overland flow occurrence. The terrain-based distributed model TOPOG_SBM [28,101,102] was used to simulate runoff and flow occurrence using four different approaches to the representation of saturated hydraulic conductivity over La Cuenca. The first was to assume a uniform value equal to the median of all of the individual measurements. The second was to assume the median value for each of the three land units and apply these values spatially, based on the map of land units. The third was to randomly allocate deciles from the cumulative distribution of the total population of measurements. The fourth was to randomly allocated deciles from the cumulative distribution for each of the land units.

The simulated patterns of runoff occurrence were then compared to the observed pattern from analysis of the flow detectors (Fig. 5). These patterns, along with plots of frequency distributions, clearly showed that the simulations using the fourth of the representations of saturated hydraulic conductivity, which combined a deterministic and a stochastic pattern, was the only one that adequately represented the observed pattern of runoff occurrence. Standard data such as catchment runoff provided complementary information on model performance, but was insufficient to distinguish between the spatial representations.

5.3. Soil moisture patterns at Mahurangi

The Mahurangi catchment in New Zealand is the site of MARVEX (the MAhurangi River Variability Experiment). The total catchment drains 50 km^2 of steep hills and gently rolling lowlands located 70 km north of Auckland [121]. "Satellite Station" is a sub-catchment of approximately 1 km^2 in area, covered largely by pasture and is one of several intensively monitored areas within the MARVEX (see Woods et al. [121] for details of the experimental set up). Six soil moisture patterns

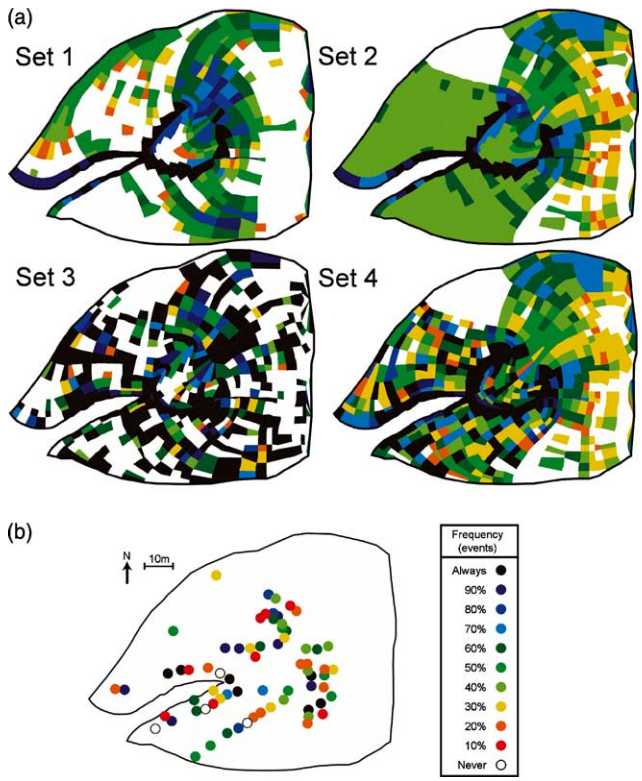


Fig. 5. Spatial patterns of overland flow frequency at La Cuenca, Peru. (a) Simulations using TOPOG and four different representations of saturated hydraulic conductivity: set 1 (uniform), set 2 (organised), set 3 (random), and set 4 (random & organised). (b) Observed patterns from runoff detectors (redrawn with permission from Cambridge University Press).

were collected over the top 30 cm of soil (the root zone for this vegetation) on a 40m $\times$ 40 m grid using the TDR equipped all terrain vehicle discussed above [25,96,110]. In addition, runoff was measured at two points and meteorological data was available locally. The six patterns of soil moisture were compared to simulations as part of the development of a new version of the Thales terrain-based hydrological model, known as NetThales [25]. Fig. 6 shows the observed and simulated patterns for three occasions (Fig. 6a), along with residual maps (Fig. 6b) and cumulative and frequency distributions of the errors (Fig. 6c). Chirico et al. [25] outline in detail how the patterns of soil moisture were used in conjunction with analysis of standard runoff data to develop and test the model structure of NetThales, including a new representation of the sub-grid scale effects of surface and sub-surface flow interaction. Patterns of soil moisture provided information on the size of saturated source area, and the extent to which they were affected by terrain, assisting with the appropriate conceptualisation of grid scale algorithms. Analysis of recessions provided information of the effective celerity of surface and sub-surface flows, and the hydrograph provided data for volume balance comparisons. Further modelling and analysis of the patterns has indicated that soil hydraulic properties appear to show significant temporal variability, particularly during times of rapid change in average soil moisture (e.g. Autumn and Spring transitional periods). This is in line with the conclusions of Western and Grayson [111] who analysed observed and simulated patterns of soil moisture at the Tarrawarra site in Australia, and suggested that dynamic changes in cracks and macropores may explain such temporal changes.

6. Calibration and testing—a role for pattern comparison methods

Calibration and testing of hydrological models has been an active area of research in recent years. The greater use of complex models has increased the problems of balancing data availability, predictive performance and model complexity (Fig. 1), which has led to questioning of the classical calibration paradigm [48]. The last ten years have seen the development of several new methods for the calibration and testing of models using time-series data. Sorooshian and Gupta [92] and McLaughlin and Townley [75] present reviews of approaches to calibration and automated optimisation in surface and groundwater models respectively. They highlight a range of problems such as the presence of multiple optima, strong correlations between parameters, and the subjectivity associated with choice of objective functions. Many researchers have noted that it is common for a large number of parameter sets to give similar fits to observed data. Gupta et al. [48] propose methods that consider the multiobjective nature of calibration, and that allow for explicit consideration of model error, with a recent application provided in Boyle et al. [23]. Mroczkowski et al. [79] have illustrated the value of different types of time-series data for model testing and calibration. Beven and Binley [11] proposed an alternative approach to deal with problems of constraining predictive uncertainty, identifying "behavioural" parameter sets, and defining the value of additional data, all within the framework of Bayesian uncertainty estimation. Their GLUE procedure has been applied to many cases (e.g. [10,38,39,124]) including one where information on saturated area was incorporated with more common runoff estimation to constrain predictive uncertainty [37]. The information on saturated area was obtained from a combined analysis of SAR data and terrain modelling, but in this case, just the total area saturated (rather than information on the spatial pattern) was used. Nevertheless, the value of some information on spatial response was shown to be high in terms of ability to reject non-behavioural parameter sets. The Metropolis algorithm, another

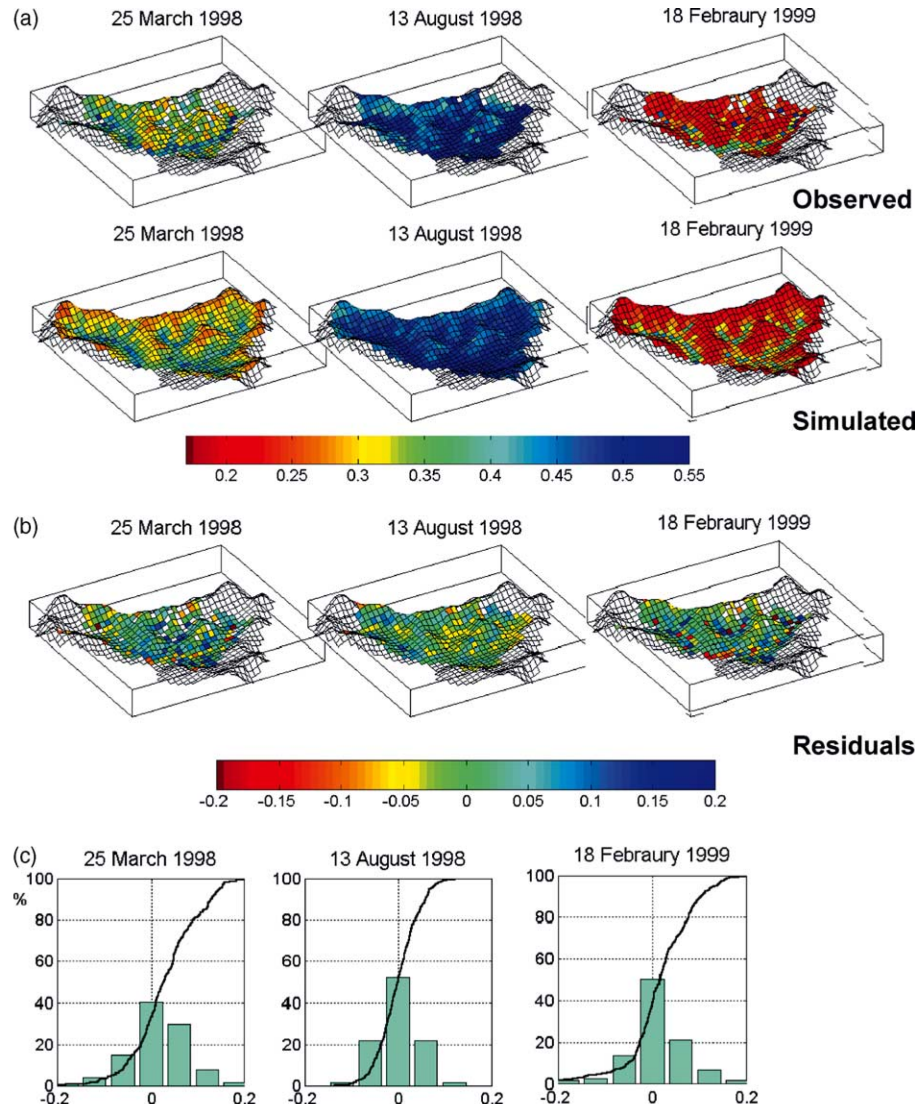


Fig. 6. Patterns of soil moisture in the root zone from the satellite station sub-catchment of the Mahurangi River, (a) observed patterns from TDR and simulated patterns using NetThales for three occasions, (b) residual maps for each occasion, (c) pdf and cdf of errors for each occasion.

Bayesian approach, has also been applied recently for assessing hydrological model uncertainty [64].

One of the main points to have been illustrated in recent work on model testing and calibration is that different types of observed data, and different periods within a given time series, contain information that tests different parts of a model. For example Wagener et al. [106,107] show the value of different parts of a hydrograph in constraining uncertainty in predictions using their DYNIA (DYNamic Identifiability Analysis) approach, and Boyle et al. [23] split the hydrograph into different sections with separate objective functions for each. The same argument applies to the use of spatial patterns. There are particular times and places when the "information content" of a pattern may be greatest. For example, in the Schneealpe example above, Jansa et al. [58] used patterns of snow cover from periods where cover was rapidly changing in their model testing, because these periods highlighted model sensitivity to snow melt. Patterns in early Winter highlighted threshold air temperature effects, while those in late Spring highlighted albedo effects. Similarly Western and Grayson [111] showed that patterns of soil moisture from periods of transition from dry Summer to wet Winter conditions (and vice versa) were useful for characterising effective soil hydraulic properties, while transitions from wet to dry periods were particularly useful for testing evapotranspiration components of their model, where aspect effects were most pronounced.

The recent advances in calibration and testing methodology have highlighted the importance of information in addition to standard runoff data. Grayson and Blöschl [47] illustrate that spatial patterns provide powerful tests for distributed models, have the ability to efficiently identify "behavioural" from "non-behavioural" models, and can provide independent information that is complementary to more traditional data sources. But as yet, pattern information has not been fully used with any of the newer calibration or testing methodologies. This is due, at least in part, to the lack of standard methods for the comparison of observed and simulated patterns of hydrological response. Here we briefly describe the present state of the art and propose some pattern comparison methods that provide information of interest in hydrology.

The methods presented in this section have been selected because the results from them can be interpreted hydrologically, thus enabling identification of model components that are performing poorly and perhaps improvements in hydrological understanding. They are all quite simple, and there is clearly scope for more sophisticated approaches, perhaps drawn from pattern recognition science. Nevertheless, even these simple approaches can be powerful discriminators of model performance and could be used in first attempts to combine pattern information into the newer calibration and testing frameworks.

Visual comparison is the most common method used in studies to date and provides a qualitative assessment of model performance. It is simple and exploits the great capability of the human brain to detect and interpret patterns, and bring a large amount of accumulated knowledge about model and system behaviour to bear. Areas of consistent difference between two patterns can often be identified and qualitative associations can be made with the model components that may be causing errors. The disadvantage of this method is that it does not provide a quantitative measure of model performance, nor is it possible to test hypotheses about specific behaviour (e.g. is a "consistent" difference between two patterns statistically significant?). Thus, it is not possible to extend the method to automated techniques and makes inter-model comparisons rather subjective. Nevertheless, visual assessment is probably the most powerful comparison method, provided the time is available to consider patterns in detail. This may not be possible in modelling involving large areas or comparisons between simulations using a large number of parameter combinations.

Point-by-point comparison methods include scatterplots of simulated and observed "pixel" values, R^2 (coefficient of determination) and E (coefficient of efficiency [80]) values from these plots, and the *mapping* of differences (residuals) between observed and simulated patterns (e.g. Figs. 4a–c and 6b). These techniques provide information about bias (mean error), random simulation errors (error variance) and in the case of residual maps, any spatial organisation that may be present in the errors. The mean error and the error variance are similar to statistics used in traditional model evaluation using time series; however, they can be applied in a spatial context to test internal model predictions. The scatterplots also provide information on how errors are distributed across the range of simulated values. For example soil moisture may be represented well under wet conditions but poorly for low moisture values, indicating problems with evaporation or soils components of the model. R^2 and E values provide quantitative measures for use with automated schemes. For binary patterns, the proportion of pixels correctly and incorrectly identified provides similar information to the scatter plots for continuous variables.

The spatial correlation structure (variogram) of errors can be computed from maps of the residuals, providing information about the spatial scale or correlation length of the errors (e.g. [116]). If the correlation length of the errors is small relative to the model element scale, it can be concluded that the errors are due to either measurement error or to small-scale variability not resolved by the model. Since the model does not aim to represent these features, it can be concluded that the model is performing as well as can be expected (assuming there is no bias and a sufficiently small error variance). If the correlation length of the errors is significantly longer than the model grid scale, it can be concluded that there are patches where the errors are similar; i.e. there is some problem with the structure of the model that causes certain parts of the landscape to be better represented than others. A careful analysis of the simulated response and an understanding of the model structure gives guidance on potential model improvements. For example hillslopes may be consistently too dry compared to gullies, indicating problems with lateral redistribution in the model; or areas of land with particular soils may be consistently in error, indicating problems with the parameterisation of soil hydraulic properties.

An extension of the point-by-point approach, which accounts for measurement error and sub-element variability that is not represented by the model, is to smooth the observed pattern and then compare it with simulations on a point-by-point basis. For the smoothing, geostatistical techniques such as kriging provide a convenient method because properties of the "smoothing filter" can be set to mimic the expected small-scale variability [22] by making the nugget of the variogram equal to the sum of measurement error and sub-element variability that is not represented by the model. Point-by-point comparison between the 'smooth' observed and simulated patterns should then represent the large-scale variability between observed and simulated patterns (e.g. [111]).

To gain further insight into which hydrological process representations may or should be improved in the model, errors can be analysed to ascertain whether there is any *relationship with topographic or other spatial* *variables (e.g. soil type, vegetation)* by plotting the observed and simulated values (or residuals) against such variables for each location. These relationships may provide hydrological insight into the cause of the errors. For example, consistent errors associated with topographic aspect may imply problems with components of the model influenced by radiation exposure.

A limitation of the methods above (except visual comparison) is that they do not provide any information on lateral shifts—i.e. where the basic shape of patterns is correct but their location is shifted. Because water flows through the landscape along pathways that are dominated by the topography, pattern shifts may be associated with particular terrain features. *Transects* of simulated and observed variables can be examined to search for shifts between the simulated and observed patterns. The transects can be placed such that they follow surface flow trajectories, elevation contours, or other directional features where shifts might be expected due to the model structure and physics of the problem. For example in a snow model, observed and simulated snow depth plotted for a transect down a hillslope may provide information on how well temperature lapse rates are represented or how well sloughing or avalanching is simulated.

Transects provide information about shifts in one dimension only. Methods developed for pattern recognition applications may also be applied to determine the lateral shifts in two dimensions. Optimal local alignment (OLA) is a method that provides information on space shifts between two patterns. The method is a form of particle image velocimetry (e.g. [40,120]) and is based on cross-correlations between "windows" within the overall domain. A field of shift vectors can be calculated in the following way. Initially the whole domain is divided into sub-areas (or windows). Then correlation coefficients between point-by-point comparisons of the observed versus simulated patterns are calculated for corresponding sub-areas. The relative position of the two corresponding sub-areas is then changed (i.e. shifted) and correlations are again calculated (i.e. a cross-correlation analysis). This is repeated over a defined range of shifts in each direction. The optimum shift (i.e. optimum alignment) is where the correlation is highest. This approach has the potential to identify model mismatches in the direction of the hillslopes as well as other shifts such as those associated with biases due to aspect, the way soil parameters were imposed, or georeferencing problems. To use this method successfully it is necessary to have small sub-areas but there is a tradeoff between having sufficient points in each sub-area for reliable estimation of correlations and obtaining detailed spatial information (high resolution) in the resulting vector field. In applications in image particle velocimetry, the size of the data sets can be such that the approach described above becomes computationally intractable. Willert and Gharib [120] describe a method using fast fourier transforms that significantly speeds up the computation of correlations.

It is important to note that shift vectors can be computed only when there is variability apparent in both the observed and simulated fields. This means that for binary patterns, vectors can be computed only for the edge areas. Also if model predictions are uniform for a particular period, the method will fail.

7. Application of pattern comparison methods

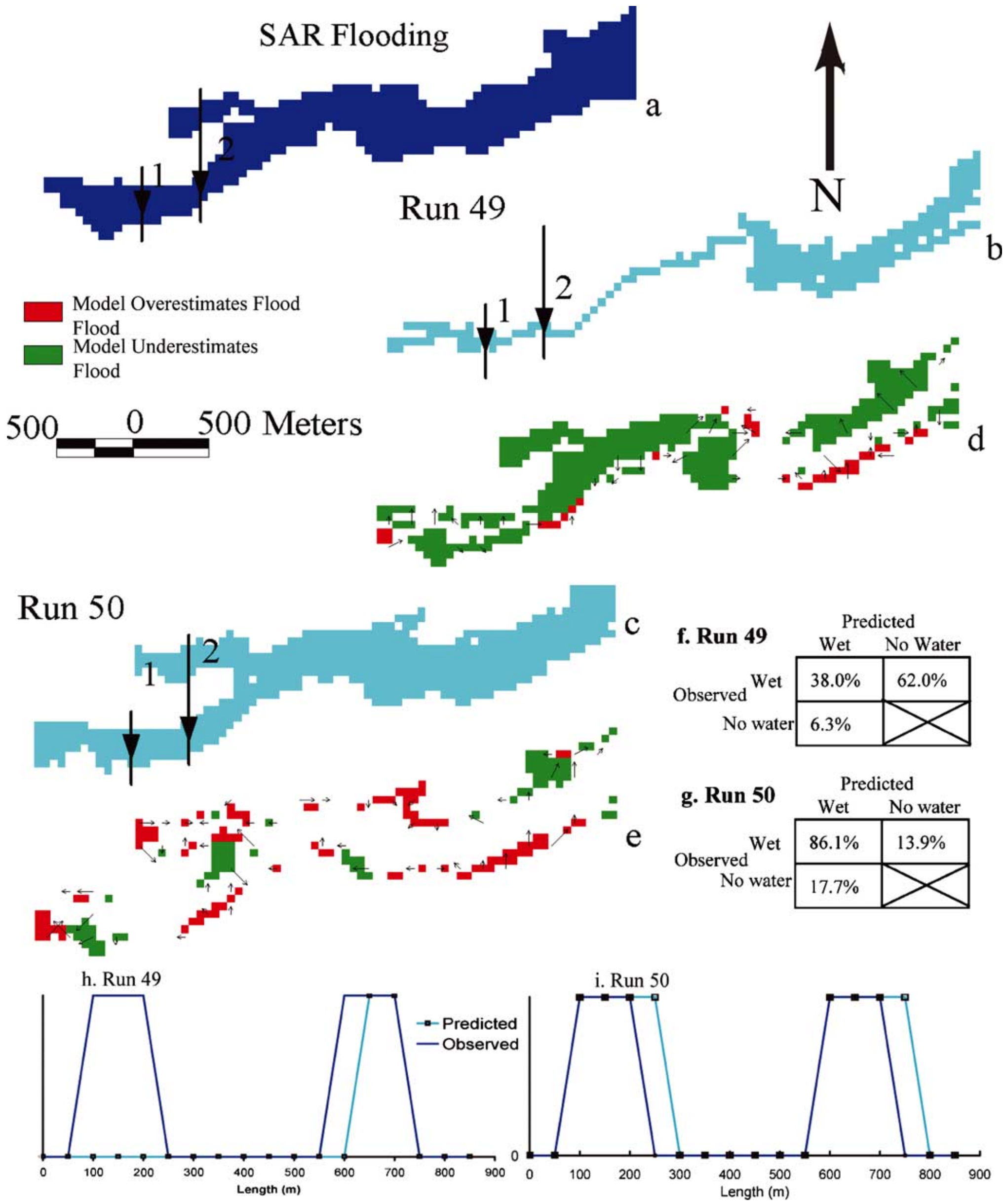
To illustrate some of the points made above, the pattern comparison methods are applied to two different data sets where the methods can be interpreted from a hydrological perspective [117].

7.1. Data sets

The data sets were chosen to be representative of different types of data that may be encountered in hydrological modelling. The descriptions are brief since the comparisons are simply for illustrative purposes and are not intended to formerly test the models used for the simulations. The first data set is of a binary variable, flood inundation, while the second is a continuous (bounded) variable, root zone soil moisture.

The patterns of flood inundation are taken from the work of Horritt and Bates [53]. Measured inundation is derived from analysis of SAR images taken over the Thames River [15,54]. The simulated inundation is produced by the model of Bates and De Roo [5], which combines a one-dimensional kinematic wave channel flow model with a two-dimensional diffusion wave solution for floodplain flow. The resolution of the SAR data was higher than that of the model so the former was aggregated to match the model element scales. Two simulated patterns were chosen from an ensemble of model results for illustrative purposes (P. Bates, personal communication). Fig. 7a shows the observed inundated area resulting from the SAR analysis while Fig. 7b and c show the simulated inundation from the two model runs.

The measured root zone soil moisture data sets used here are described in detail by Western and Grayson [110] and are of soil moisture measured in the top 30 cm of soil over a 10 ha, undulating pasture catchment in south eastern Australia referred to previously. One pattern was chosen from a relatively wet period (October, 1996) to illustrate the methods. Fig. 8a shows the observed pattern, topography and the location of transects used in the analysis. The simulations used for comparison were produced by the Thales model and are described in detail in Western and Grayson [111]. Thales uses a contour based element network, which does not



f. Run 49

		Predicted Wet	Predicted No Water
Observed	Wet	38.0%	62.0%
	No water	6.3%	

g. Run 50

		Predicted Wet	Predicted No water
Observed	Wet	86.1%	13.9%
	No water	17.7%	

Fig. 7. Pattern comparisons of flood inundation on the Thames River, England. (a) Observed pattern from SAR analysis, (b) simulated inundation for run 49, (c) simulated inundation for run 50, (d) residual maps and OLA vectors from run 49, (e) residual maps and OLA vectors from run 50, (f) table of errors for run 49, (g) table of errors for run 50, (h) transect 2 for run 49, (i) transect 2 for run 50.

match the measurement grid (10 m $\times$ 20 m), hence simulated results were interpolated onto the measurement grid for the comparisons below (Fig 8b).

7.2. Results and discussion

7.2.1. Flood inundation example

Fig. 7a can be visually compared to Fig. 7b and c. Clearly the inundated area is underestimated in Fig. 7b, particularly for the western half of the reach. Fig. 7c shows some minor underestimation in the north-eastern section but does well at identifying the small off shoot from the main river channel. The general impression of this comparison is quantified in the results in Fig. 7f and g. These tables show that, normalised by flooded area, run 50 successfully identifies 86% of the flooded pixels whereas run 49 correctly identifies only 38%. Neither run simulates many pixels as being wet that are actually dry, with most of the erroneous pixels being predicted as dry when in fact they were wet—i.e. both runs underpredicted total inundated area.

The maps of residuals (Fig. 7d and e for runs 49 and 50 respectively) show how these errors are distributed spatially. Most of the underprediction of run 49 occurs to the north of the river, except in the far western end where errors occur on both sides of the river. The overall

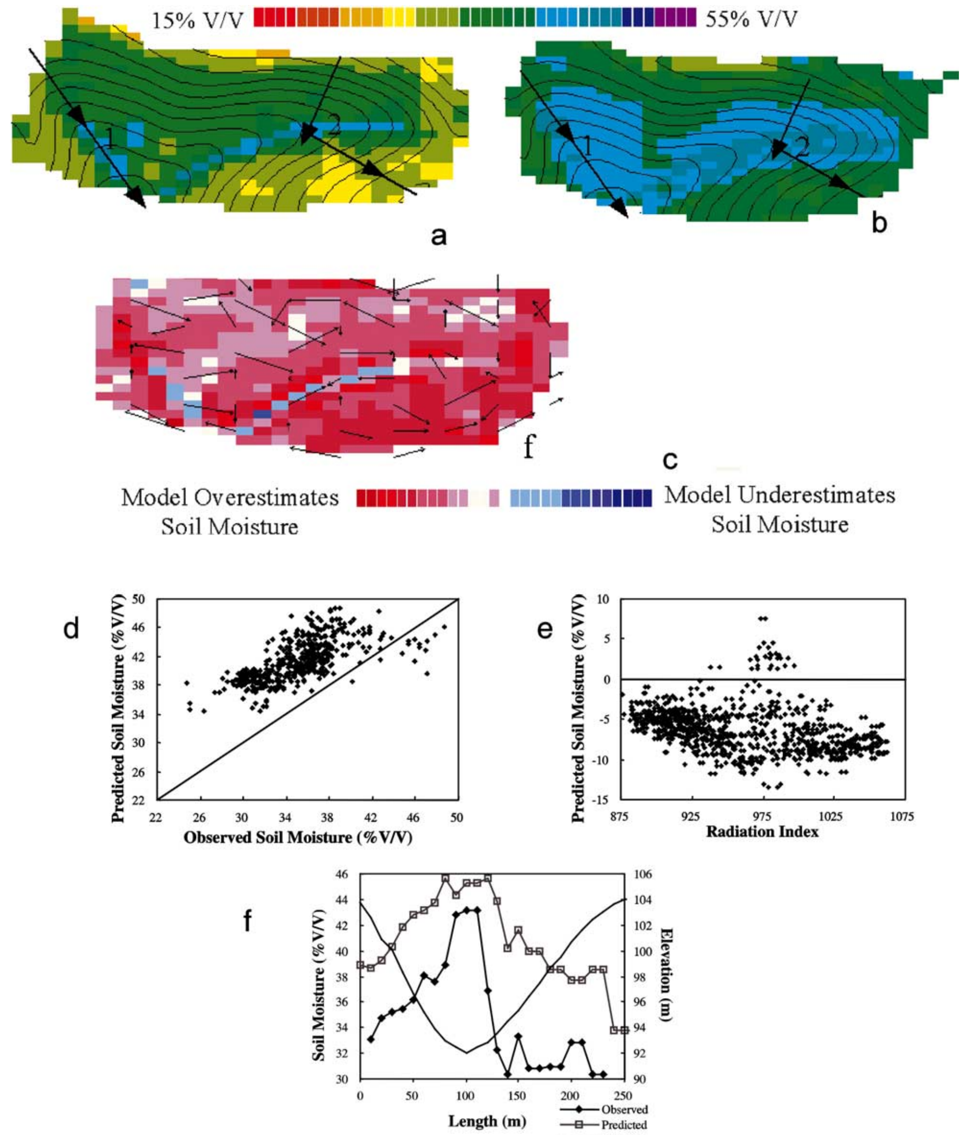


Fig. 8. Pattern comparisons of soil moisture at Tarrawarra. (a) Observed root zone soil moisture from TDR, (b) simulated root zone soil moisture from Thales, (c) residual map and OLA vectors, (d) point-by-point scatterplot, (e) residuals versus radiation index, (f) transect 2 soil moisture and elevation.

underestimation of run 50 is quite evenly distributed spatially.

The transect 2 comparison (Fig. 7h and i) provide little information in addition to the residual maps for this simple binary pattern, although perhaps makes the magnitude of the errors easier to identify.

The simple patterns of flood inundation are amenable to OLA analysis and the resulting shift vectors are shown in Fig. 7d and e. Areas where the model underestimates inundated area have shift vectors pointing out onto the floodplain, and vice versa. Fig. 7e shows that for the eastern end of the reach, the simulated area lies too far south although the small magnitude of the shifts indicate overall good model performance. As mentioned above, shift vectors can be computed only where there is variability in both the modelled and measured pixels in the windows. This is apparent in the large areas where the model underestimates inundated area in run 49 and no shift vectors are computed.

For the patterns of inundated area, visual comparisons of simulated and observed patterns and residual maps are simple enough to interpret visually. The shift vectors from OLA are useful in indicating the direction in which the model is in error. With simple patterns, this may be determined visually, but as the patterns get more complex, the information conveyed by the shift vectors in conjunction with the residual map becomes more useful. The quantification of the number of pixels in error is probably the best quantitative measure for automated optimisation, while the residual maps and OLA vectors indicate more about the structural problems in the model.

7.3. Soil moisture example

Visual comparison of Fig. 8a (observed soil moisture) and b (simulated) show that the simulated soil moisture is generally higher and less variable than that observed. The general pattern of a dryer south-east quarter is captured by the model but the magnitude of the difference is less than observed, and other dryer areas around the northern and western boundary are not simulated. The strong aspect effect of the dryer south-eastern (north facing) hillslope is not well simulated, although the overall wetter gully is represented, but the extent is overestimated. With respect to the performance of the model, the apparently 'too wet' simulation may imply problems with the evaporation estimation, or else drainage (either surface or sub-surface) is not well modelled.

The point-by-point comparison is shown in Fig. 8d and supports the notion that the simulated pattern is too wet overall, except for the wettest observations, which are underestimated. The variance appears to be similar for all except the highest values.

The pattern of residuals is shown in Fig. 8c and reinforces the likely problems with modelled drainage (or evapotranspiration), particularly for the south eastern hillslope. Fig. 8e shows the residuals plotted against radiation index (a function of topographic aspect; [116]) and highlights the problem with poor prediction on the north facing hillslopes in the south-east (radiation index > 1). The overly dry gullies are apparent around radiation index values of 0.975.

Transect 2 is identified in Fig. 8a and the observed and simulated data plotted in Fig. 8f. The results show the "too wet" simulation but generally match the pattern of variability, although there appears to be shift in pattern at the right end of the transect.

Results from the OLA (Fig. 8d) generally show random shift vectors, indicating no consistent errors in the model. The exception is in the south-east quarter where there is a consistent "patch" of shift vectors upslope, matching the shift error apparent in the transect data. This is likely to be due to errors in the lateral redistribution of water in this area, or problems with the evapotranspiration estimates or both. The ability of OLA to detect this spatial shift problem is encouraging, and the "condensed information" contained in the shift vectors would make analysis of large and complex patterns more tractable, making it easier to detect areas of consistent model problems compared to visual analysis. More work is needed on how best to quantify the information in OLA analysis.

8. Conclusions and outlook

In the preceding sections we have illustrated the advances made in the use of spatial patterns of hydrological response in relation to three main areas (i) characterising the structure and heterogeneity of hydrological variables, (ii) using patterns (LOPs, binary and surrogate) to test and develop distributed models, and (iii) specific techniques for comparing patterns in a manner that assesses hydrologically important characteristics of model performance. While the use of observed patterns in catchment hydrology is relatively new, the progress made to date augers well for future developments. In this section we propose several areas where we believe emphasis should be placed in the coming years.

Geostatistical methods have been shown to be useful in characterising some aspects of the structure and heterogeneity of observed patterns, such as soil moisture and snow cover, but there are important hydrological features that are not captured by conventional methods. In particular, the connectivity of patterns (e.g. of wet areas connected to a catchment outlet, or of high conductivity flow paths in the sub-surface) has been shown to be extremely important to simulating hydrological response, yet methods for its characterisation are only just being tested. In distributed modelling it is often appropriate to stochastically generate patterns of parameters or other input information that preserve the structure of an observed pattern but enable sensitivity to alternate realisations to be assessed. While this is possible with standard geostatistical characterisations, methods that preserve connectivity measures are not yet available. More generally, we may draw some confidence for the future from the work done in precipitation where the availability of detailed spatial observations (from radar) has enabled new methods for generating space–time patterns of rainfall that preserve the key features of these intricate patterns (e.g. [90,99]). As more data becomes available on other variables of hydrological interest, we anticipate similar advances to occur.

There is also the potential for increased use of surrogate patterns that identify the functional behaviour of catchments. These provide rigorous tests of the "behavioural" nature of simulations of catchment response and should minimise the problems of what Klemeš [63] described as the *right results for wrong reasons* (e.g. [97]). Although the use of observed patterns of catchment dynamics does not necessarily improve streamflow simulations at the outlet as there is the danger of overfitting [67,68], it will improve the spatial estimates of the hydrologic quantities within catchments. These spatial estimates are increasingly of practical interest. Observed patterns of catchment dynamics include saturated source areas and their dynamics, observations of runoff occurrence, and more detailed patterns of runoff process types such as those of Peschke et al. [82]. This type of information generally requires extensive field work and its utility is enhanced by innovative methods of comparison that enable the inclusion of categorical information.

The pattern comparison methods presented are simple and there is clearly scope for more sophisticated approaches. Nevertheless, even these relatively simple methods are yet to become part of routine testing, and are yet to be implemented in the more recent approaches to model calibration and uncertainty prediction described above. This is partly due to the newness of the methods, partly to data availability and partly to computational limitations. In time, each of these constraints will lessen and the ability of patterns to rigorously test and improve models will be more fully exploited.

An area where there has already been extensive use of spatial data (largely from remote sensing instruments) is in data assimilation. There are a number of specific techniques used in data assimilation [74], but in applications to date, the information used is just the pixel value (and some error estimate for the observation), rather than any measure of the pattern structure. There may be value in exploring different approaches to assimilation of spatial data, particularly where there is a chance of data errors in absolute values (such as may occur if atmospheric corrections to remotely sensed images are biased), while the underlying structure of the pattern may be quite useful. For some applications it may be more effective to assimilate, for example, the observed and simulated pattern each normalised by their means, enabling corrections to the spatial patterns to be made, while maintaining the modelled means. Data assimilation for catchment hydrological applications is relatively new (e.g. [98]) but there appears to be the potential for significant advances in our ability to understand and predict spatial hydrological response that will arise from the application of data assimilation methods.

Finally, we believe that there is a need for better quality archiving and accessing of data, particularly from research catchment studies, to better facilitate comparative studies. The remote sensing community has led the way with efficient archiving, access and retrieval approaches, as have many operational meteorological and hydrographic organisations. But this has generally not been the case for data from individual research catchments. Progress has been made in some major programs such as the Southern Great Plains and similar experiments, and the research community is better recognising the value of data via developments such as the "data notes" in *Water Resources Research*. Nevertheless, there are enormous amounts of data, including detailed spatio-temporal patterns of hydrological response, that have been collected as part of field projects and post-graduate study programs that would be useful to the wider hydrological community but are largely inaccessible. We encourage everyone involved in or planning such programs to build into their activities the resources needed to make their information more widely available and use vehicles such as data notes to publicise their work.

As was made clear in the quote from Hornberger and Boyer at the start of this paper, and we hope we have illustrated through the examples discussed above, the combined use of data and modelling is a key to progress in hydrological science. The increased availability of spatial data, both from ground-based studies and remote sensing, as well as developments in modelling approaches and our deeper understanding of the role and limitations of modelling, has heralded advances in our ability to understand and predict spatial hydrological behaviour. We hope this trend continues apace and foresee a central role for the use of observed patterns of hydrological response.

Acknowledgements

Our research work has been supported financially by the Cooperative Research Centre for Catchment Hydrology, Australian Research Council (projects C39804872, A39801842), University of Melbourne Collaborative Grant's scheme, Austrian Academy of Sciences project HOE18, and the Austrian Federal Ministry of Education, Science and Culture. We would also like to thank Josef Jansa who provided the SPOT snow classifications, Robert Kirnbauer for his collaboration on the snow modelling projects, Rob Vertessy and Helmut Elsenbeer whose work on runoff occurrence was used in the examples, Paul Bates who provided the patterns of inundation, and Kim Mueller who undertook the programming related to the pattern comparison methods. We would also like to thank Keith Beven for his constructive review of the manuscript.

References

[1] Abbott MB, Refsgaard JC., Distributed hydrological modelling, Kluwer Academic Publishers. 321p.

[2] Allard D, Group H. On the connectivity of two random set models: the truncated Gaussian and the Boolean. In: Soares A, editor. Geostatistics Tróia '92. Quantitative geology and geostatistics. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1993. p. 467–78.

[3] Allard D. Simulating a geological lithofacies with respect to connectivity information using the truncated Gaussian model. In: Armstrong M, Down PA, editors. Geostatistical simulations Proceedings of the geostatistical simulation workshop, Fontainbleau, France, 27–28 May 1993. Quantitative geology and geostatistics. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1994. p. 197–211.

[4] Bárdossy A, Lehmann W. Spatial distribution of soil moisture in a small catchment. Part 1: geostatistical analysis. J Hydrol 1998;206:1–15.
[5] Bates PD, De Roo APJ. A simple raster-based model for flood inundation simulation. J Hydrol 2000;236:54–77.
[6] Beckie R. Measurement scale, network sampling scale, and groundwater model parameters. Water Resour Res 1996;32:65–76.
[7] Beven KJ. Uniqueness of place and process representations in hydrological modelling. Hydrol Earth Syst Sci 2000;4(2):203–13.
[8] Beven KJ. Rainfall runoff modelling—the primer. John Wiley and Sons; 2001. 360p.
[9] Beven KJ. How far can we go with distributed hydrological modelling? Hydrol Earth Syst Sci 2001;5(1):1–12.
[10] Beven KJ. Calibration, validation and equifinality in hydrological modelling. In: Anderson MD, Bates PD, editors. Model validation in the hydrological sciences. Chichester: Wiley; 2001. p. 43–55.
[11] Beven KJ, Binley AM. The future of distributed models: model calibration and uncertainty prediction. Hydrol Proc 1992;6:279–98.
[12] Beven K. Changing ideas in hydrology—the case of physically based models. J Hydrol 1989;105:157–72.
[13] Beven KJ. Prophecy, reality and uncertainty in distributed hydrological modelling. Adv Water Resour 1993;16:41–51.
[14] Beven KJ, Moore ID, editors. Terrain analysis and distributed modelling in hydrology. New York: John Wiley and Sons; 1992. 249p.
[15] Biggin DS, Blyth K. A comparison of ERS-1 satellite radar and aerial photography for river flood mapping. J Chartered Inst Water Environ Manage 1996;10:59–64.
[16] Blöschl G, Sivapalan M. Scale issues in hydrological modelling—a review. Hydrol Process 1995;9:251–90.
[17] Blöschl G. Scaling issues in snow hydrology. Hydrol Process 1999;13:2149–75.
[18] Blöschl G, Kirnbauer R. An analysis of snow cover patterns in a small Alpine catchment. Hydrol Process 1992;6:99–109.
[19] Blöschl G, Kirnbauer R, Gutknecht D. Distributed snowmelt simulations in an Alpine catchment. 1. Model evaluation on the basis of snow cover patterns. Water Resour Res 1991;27:3171–9.
[20] Blöschl G, Gutknecht D, Kirnbauer R. On the evaluation of distributed hydrologic models. In: Rosso R, Peano A, Becchi I, Bemporad GA, editors. Advances in distributed hydrology, Proceedings of a workshop held in Bergamo, Italy, June 1992. Highlands Ranch, CO, USA: Water Resources Publications; 1994. p. 397–412.
[21] Blöschl G, Kirnbauer R, Jansa J, Kraus K, Kuschnig G, Gutknecht D, Reszler Ch. Einsatz von Fernerkundungsmethoden zur Eichung und Verifikation eines flächendetaillierten Schneemodells (Using remote sensing methods for calibrating and verifying a spatially distributed snow model). Österr Wasser u Abfallwirtsch 2002;54:1–16.
[22] Blöschl G, Grayson R. Spatial observations and interpolation. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 17–50 [Chapter 2].
[23] Boyle DP, Gupta HV, Sorooshian S, Koren V, Zhang Z, Smith M. Towards improved streamflow forecasts: the value of semi-distributed modelling. Water Resour Res 2000;36(12):3663–74.
[24] Bronstert A, Plate E. Modelling of runoff generation and soil moisture dynamics for hillslopes and micro-catchments. J Hydrol 1997;198:177–95.
[25] Chirico GB, Grayson RB, Western AW. Lateral flow processes in a small experimental catchment: analysis and distributed modelling. Hydrol Process, in press.
[26] Cline DW, Bales RC, Dozier J. Estimating the spatial distribution of snow in mountain basins using remote sensing and energy balance modeling. Water Resour Res 1998;34:1275–85.
[27] Davis RE, McKenzie JC, Jordan R. Distributed snow process modelling: an image processing approach. Hydrol Process 1995;9: 865–75.
[28] Dawes WR, Zhang L, Hatton TJ, Reece PH, Beale GTH, Packer I. Evaluation of a distributed parameter ecohydrological model (Topog_IRM) on a small cropping rotation catchment. J Hydrol 1997;191:64–86.
[29] Dunne T, Black RD. Partial area contributions to storm runoff in a small New England watershed. Water Resour Res 1970;6(5): 1296–311.
[30] Dunne T, Moore TR, Taylor CH. Recognition and prediction of runoff-producing zones in humid regions. Hydrol Sci Bull 1975;20:305–27.
[31] Elsenbeer H, Cassel K, Castro J. Spatial analysis of soil hydraulic conductivity in a tropical rainforest catchment. Water Resour Res 1992;28(12):3201–14.
[32] Elsenbeer H, Lack A. Hydrometric and hydrochemical evidence for fast flowpaths at La Cuenca, Western Amazonia. J Hydrol 1996;180:237–50.
[33] Elsenbeer H, Vertessy RA. Stormflow generation and flowpath characteristics in an Amazonian rainforest catchment. Hydrol Processes 2000;14:2367–81.
[34] Elsenbeer H. Pedotransfer functions in hydrology. J Hydrol 2001;251(3–4):121–221 [special issue].
[35] Entin JK, Robock A, Vinnikov KY, Hollinger SE, Liu SX, Namkhai A. Temporal and spatial scales of observed soil moisture variations in the extratropics. J Geophys Res—Atmos 2000;105: 11865–77.
[36] Federico VD, Neuman SP. Scaling of random fields by means of truncated power variograms and associated spectra. Water Resour Res 1997;33:1075–85.
[37] Franks SW, Gineste Ph, Beven KJ, Merot Ph. On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process. Water Resour Res 1998;34(4):787–97.
[38] Franks S, Beven KJ. Bayesian estimation of uncertainty in land surface–atmosphere flux predictions. J Geophys Res 1997; 102(D20):23991–9.
[39] Freer J, Beven KJ, Ambroise B. Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resour Res 1996;32(7):2161–73.
[40] Gharib M, Daribi D. Digital image velocimetry. In: Smith, Lim, editors. Flow visualisation: techniques and examples. Imperial College Press; 2000 [chapter 6].
[41] Gineste P, Puech C, Mérot P. Radar remote sensing of the source areas from the Coët-Dan catchment. Hydrol Process 1998;12: 267–84.
[42] Goovaerts P. Geostatistics for natural resources evaluation. In: Applied Geostatistics Series. Oxford University Press; 1997. 483p.
[43] Goovaerts P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J Hydrol 2000;228: 113–29.
[44] Gould H, Tobochnik J. An introduction to computer simulation methods. Part 2. Reading, MA: Addison-Wesley; 1988. 695p.
[45] Grayson RB, Blöschl G, Barling RD, Moore ID. Process, scale and constraints to hydrologic modelling in GIS. In: HydroGIS93, Proc. Vienna Symp., April 1993, vol. 211. IAHS Publ.; 1993. p. 83–92.
[46] Grayson R, Blöschl G. Spatial modelling of catchment dynamics. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 51–81 [chapter 3].

[47] Grayson R, Blöschl G. Summary of pattern comparison and concluding remarks. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 355–67 [chapter 14].
[48] Gupta HV, Sorooshian S, Yapo PO. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resour Res 1998;34(4):751–63.
[49] Hillel D. Modelling in soil physics: a critical review. In Future Developments in Soil Science Research. A collection of Soil Sci. Soc. Am. Golden Anniversary contributions presented at Annual Meeting, New Orleans. 1986. p. 35–42.
[50] Hornberger GM, Boyer EW. Recent advances in watershed modelling. US National report to the IUGG 1991–1994. Rev Geophys Supplement 1995:949–57.
[51] Hornberger GM, Spear RC. An approach to the preliminary analysis of environmental systems. J Environ Manage 1981;12: 7–18.
[52] Hornberger GM, Spear RC. Eutrophication in Peel Inlet, I, The problem-defining behaviour and a mathematical model for the phosphorus scenario. Water Res 1980;14:29–42.
[53] Horritt MS, Bates PD. Predicting floodplain inundation: raster-based modelling versus the finite-element approach. Hydrol Process 2001;15:825–42.
[54] Horritt MS, Mason DC, Luckman AJ. Flood boundary delineation from synthetic aperture radar imagery using a statistical active contour model. Int J Remote Sensing 2001;22(13):2489–507.
[55] Houser PR, Shuttleworth WJ, Gupta HV, Famiglietti JS, Syed KH, Goodrich DC. Integration of soil moisture remote sensing and hydrologic modelling using data assimilation. Water Resour Res 1998;34(12):3405–20.
[56] Houser PR, Goodrich D, Syed K. Runoff, precipitation, and soil moisture at Walnut Gulch. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 125–57 [chapter 6].
[57] Hutchinson MF. On thin plate splines and kriging. In: Tarter ME, Lock MD, editors. Computing science and statistics, vol. 25. Interface Foundation of North America, University of California, Berkeley, 1993. p. 55–62.
[58] Jansa J, Kraus K, Blöschl G, Kirnbauer R, Kuschnig G. Modelling snow melt processes in alpine areas. International Archives of Photogrammetry and Remote Sensing, vol. XXXIII, 2000. Amsterdam, CD-ROM.
[59] Jensen JR. Introductory digital image processing a remote sensing perspective. 2nd ed. NJ: Prentice-Hall; 1996.
[60] Jensen W, Calabresi G, editors. Flood monitoring workshop. Hydrol Process 1997;11(10).
[61] Kirnbauer R, Blöschl G, Haas P, Müller G, Merz B. Space–time patterns of runoff generation in the Löhnersbach catchment. In: Leibundgut Ch, Uhlenbrook S, McDonnell J, editors. Runoff Generation and Implications for River Basin Modelling. Freiburger Schriften zur Hydrologie, band 13, Institute of Hydrology, University of Freiburg i.Br./Germany, 2001. p. 37–45.
[62] Kite G, Droogers P, editors. Comparing actual evapotranspiration from satellite data, hydrological models and field data. J Hydrol 2000;229(1–2):1–100 [special issue].
[63] Klemeš V. Dilettantism in hydrology: transition or destiny. Water Resour Res 1986;22:177S–88S.
[64] Kuczera G, Parent E. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. J Hydrol 1998;211:69–85.
[65] Lakshmi V, Albertson JD, Schaake J, editors. Observations and modelling of the land surface hydrological processes. Water science and application, vol. 3. American Geophysical Union; 2001. 246p.
[66] Lakshmi V, Wood EF, Choudhury BJ. A soil–canopy–atmosphere model for use in satellite microwave remote sensing. J Geophys Res 1997;102(D6):6911–27.
[67] Lamb R, Beven K, Myrabø S. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. Adv Water Resour 1998;22(4):305–17.
[68] Lamb R, Beven K, Myrabø S. Shallow groundwater response at Minifelt. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 272–303 [chapter 11].
[69] Lehmann W. Anwendung geostatistischer Verfahren auf die Bodenfeuchte in ländlichen Einzugsgebieten. PhD thesis, Universität Karlsruhe, Karlsruhe, 1995.
[70] Loague K. Soil water content at R-5. Part 1. Spatial and temporal variability. J Hydrol 1992;139:233–51.
[71] Luce CH, Tarboton DG, Cooley KR. The influence of the spatial distribution of snow on basin-averaged snowmelt. Hydrol Process 1998;12(10-11):1671–83.
[72] Luce CH, Tarboton DG, Cooley KR. Subgrid parameterization of snow distribution for an energy and mass balance snow cover model. Hydrol Process 1999;13:1921–33.
[73] McDonnell JJ, Freer J, Hooper R, Kendall C, Burns D, Beven K, Peters J. A new method developed for studying flow on hillslopes. EOS, Trans Am Geophys Union 1996;77(47):465, 472.
[74] McLaughlin D. Recent developments in hydrologic data assimilation. US National report to the IUGG 1991–1994. Rev Geophys Supplement 1995:977–84.
[75] McLaughlin D, Townley LR. A reassessment of the groundwater inverse problem. Water Resour Res 1996;32(4):1131–61.
[76] Meyles E, Williams A, Dowd J, Ternan L. Soil moisture patterns and runoff generation in a small Dartmoor catchment, south-west England. In: Leibundgut Ch, Uhlenbrook S, McDonnell J, editors. Runoff generation and implications for river basin modelling. Freiburger Schriften zur Hydrologie, band 13, Institute of Hydrology, University of Freiburg i.Br./Germany, 2001. p. 28–36.
[77] Mohanty BP, Skaggs TH, Famiglietti JS. Analysis and mapping of field-scale soil moisture variability using high-resolution, ground-based data during the Southern Great Plains 1997 (SGP97) hydrology experiment. Water Resour Res 2000;36(4): 1023–32.
[78] Moore RJ, Harding RJ, Austin RM, Bell VA, Lewis DR. Development of improved methods for snowmelt forecasting. R&D Note 402, Research Report, Institute of Hydrology, Wallingford, 1996.
[79] Mroczkowski M, Raper GP, Kuczera G. The quest for more powerful validation of conceptual catchment models. Water Resour Res 1997;33(10):2325–35.
[80] Nash IE, Sutcliffe IV. River flow forecasting through conceptual models, part I. J Hydrol 1970;10:282–90.
[81] Nyborg L, Sandholt I. NOAA-AVHRR based Flood Monitoring, IGARSS 2001. The Institute of Electrical and Electronics Engineers, University of New South Wales, Sydney, Australia, 2001.
[82] Peschke G, Etzenberg C, Müller G, Töpfer J, Zimmermann S. Das wissensbasierte System FLAB—ein Instrument zur rechnergestützten Bestimmung von Landschaftseinheiten mit gleicher Abflußbildung. IHI-Schriften, H.10, Internat. Hochschulinstitut Zittau, 1999.
[83] Rawls WJ, Brakensiek DL, Miller N. Green-Ampt infiltration parameters from soils data. J Hydraul Eng 1983;109:62–70.
[84] Refsgaard JC. Parameterisation, calibration and validation of distributed hydrological models. J Hydrol 1997;198:69–97.
[85] Refsgaard JC. Towards a formal approach to calibration and validation of models using spatial data. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 329–54 [chapter 13].

[86] Rosenthal W, Dozier J. Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper. Water Resour Res 1996;32(1):115–30.

[87] Rosso R. An introduction to spatially distributed modelling of basin response. In: Rosso R, Peano A, Becchi I, Bemporad GA, editors. Advances in distributed hydrology. Water Resources Publications; 1994. p. 3–30.

[88] Salvucci GD, Entekhabi D. Hillslope and climatic controls on hydrologic fluxes. Water Resour Res 1995;31(7):1725–39.

[89] Salvucci G, Levine J. Groundwater–vadose zone interactions at Trochu. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 304–25 [chapter 12].

[90] Seed AW, Srikanthan R, Menabde M. A space–time model for design rainfall. J Geophys Res 1999;104(D24):31623–30.

[91] Sharma ML, Sander GA, Hunt CG. Spatial variability of infiltration in a watershed. J Hydrol 1980;45:101–22.

[92] Sorooshian S, Gupta VK. Model calibration. In: Singh VP, editor. Computer models of watershed hydrology. Highlands Ranch, CO: Water Resources Publications; 1995. p. 23–63 [chapter 2].

[93] Stauffer D, Aharony A. Introduction to Percolation theory. London: Taylor and Francis; 1991.

[94] Tóth J. Mapping and interpretation of field phenomena for groundwater reconnaissance in a prairie environment, Alberta, Canada. IASH Bull 1966;11(2):20–68.

[95] Troch P, Verhoest N, Gineste P, Paniconi C, Merot P. Variable source areas, soil moisture and active microwave observations at Zwalmbeek and Coët-Dan. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 187–208 [chapter 8].

[96] Tyndale-Biscoe JP, Moore GA, Western AW. A system for collecting spatially variable terrain data. Comput Electron Agric 1998;19(2):113–28.

[97] VanderKwaak JE, Loague K. Hydrologic-response simulations for the R5 catchment with a comprehensive physics-based model. Water Resour Res 2001;37(4):999–1013.

[98] van Loon EE, Troch PA, editors. Proceedings of the International Workshop on Catchment scale Hydrologic Modelling and Data Assimilation, Wageningen, September 3–5, 2001. 151p.

[99] Venugopal V, Foufoula-Georgiou E, Sapozhnikov V. A space–time downscaling model for rainfall. J Geophys Res 1999; 104(D16):705–21.

[100] Verhoest NEC, Troch PA, Paniconi C, De Troch FP. Mapping basin scale variable source areas from multitemporal remotely sensed observations of soil moisture behavior. Water Resour Res 1998;34(12):3235–44.

[101] Vertessy RA, Hatton TJ, O'Shaughnessy PJ, Jayasuriya MDA. Predicting water yield from a mountain ash forest catchment using a terrain analysis-based catchment model. J Hydrol 1993;150:665–700.

[102] Vertessy RA, Hatton TJ, Benyon RJ, Dawes WR. Long term growth and water balance predictions for a mountain ash (E. regnans) forest subject to clearfelling and regeneration. Tree Physiol 1996;16:221–32.

[103] Vertessy RA, Elsenbeer H. Distributed modeling of stormflow generation in an Amazonian rainforest catchment: Effects of model parametrization. Water Resour Res 1999;35(7):2173–87.

[104] Vertessy R, Elsenbeer H, Bessard Y, Lack A. Storm runoff generation at La Cuenca. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology: observations and modelling. Cambridge: Cambridge University Press; 2000. p. 247–71 [chapter 10].

[105] Vinnikov KY, Robock A. Scales of temporal and spatial variability of midlatitude soil moisture. J Geophys Res 1996; 101:7163–74.

[106] Wagener T, Wheater HS, Gupta HV. A dynamic identifiability approach to the evaluation of hydrological models. In: van Loon EE, Troch PA, editors. Proceedings of the International Workshop on Catchment Scale Hydrologic Modelling and Data Assimilation, Wageningen, September 3–5, 2001. p. 73.

[107] Wagener T, Lees MJ, Wheater HS. A framework for the development and application of parsimonious hydrological models. In: Singh, Frevert, Meyer, editors. Mathematical models of small watershed hydrology, vol. 2. Water Resources Publications LLC; 2001.

[108] Walker JP, Willgoose GR, Kalma JD. The Nerrigundah data set: soil moisture patterns, soil characteristics and hydrological flux measurements. Water Resour Res 2001;37(11):2653–8.

[109] Western AW, Blöschl G. On the spatial scaling of soil moisture. J Hydrol 1999;217:203–24.

[110] Western AW, Grayson RB. The Tarrawarra data set: soil moisture patterns, soil characteristics and hydrological flux measurements. Water Resour Res 1998;34(10):2765–8.

[111] Western A, Grayson R. Soil moisture and runoff processes at Tarrawarra. In: Grayson R, Blöschl G, editors. Spatial patterns in catchment hydrology observations and modelling. Cambridge: Cambridge University Press; 2000. p. 209–46 [chapter 9].

[112] Western AW, Grayson RB, Blöschl G. Scaling of soil moisture—a hydrological perspective. Ann Rev Earth Planetary Sci 2002;30:149–80.

[113] Western AW, Blöschl G, Grayson RB. Towards capturing hydrologically significant connectivity in spatial patterns. Water Resour Res 2001;37(1):83–97.

[114] Western AW, Blöschl G, Grayson RB. Geostatistical characterisation of soil moisture patterns in the Tarrawarra Catchment. J Hydrol 1998;205:20–37.

[115] Western AW, Blöschl G, Grayson RB. How well do indicator variograms capture spatial connectivity of soil moisture? Hydrol Process 1998;12:1851–68.

[116] Western AW, Grayson RB, Blöschl G, Willgoose G, McMahon TA. Observed spatial organisation of soil moisture and relation to terrain indices. Water Resour Res 1999;35(3):797–810.

[117] Western AW, Mueller K, Grayson RB. Methods for the comparison of observed and simulated patterns. CRC Catchment Hydrol Rep, in preparation.

[118] Whelan MJ, Anderson JM. Modelling spatial patterns of throughfall and interception loss in a Norway spruce (*Picea abies*) plantation at the plot scale. J Hydrol 1996;186:335–54.

[119] Wigmosta MS, Vail LW, Lettenmaier DP. A distributed hydrology–vegetation model for complex terrain. Water Resour Res 1994;30(6):1665–79.

[120] Willert CE, Gharib M. Digital particle image velocimetry. Exp Fluids 1991;10:181–93.

[121] Woods RA, Grayson RB, Western AW, Duncan M, Wilson DJ, Young RI, Ibbitt R, Henderson R, McMahon TA. Experimental design and initial results from the Mahurangi river variability experiment: MARVEX. In: Lakshmi V, Albertson JD, Schaake J, editors. Observations and modelling of the land surface hydrological processes. Water science and application, vol. 3. American Geophysical Union; 2001. p. 201–13.

[122] Yang S, Woo M. Representativeness of local snow data for large scale hydrologic investigations. Hydrol Process 1999;13:1978–88.

[123] Young KL, Woo M, Edlund SA. Influence of local topography, soils and vegetation on microclimate and hydrology at a high arctic site, Ellesmere Island, Canada. Arctic Alpine Res 1997; 29(3):270–84.

[124] Zak S, Beven KJ, Reynolds B. Uncertainty in the estimation of critical loads: a practical methodology. Soil, Water Air Pollut 1997;98:297–316.

[125] Davis RE, McKenzie JC, Jordan R. Distributed snow process modelling: an image processing approach. Hydrol Process 1995; 9:865–75.

[126] Moore ID, Grayson RB. Terrain based prediction of runoff with vector elevation data. Water Resour Res 1991;27(6):1177–91.

[127] Houser PR. Challenges and progress in operations land surface data assimilation. In: van Loon EE, Troch PA, editors, Proceedings of the International Workshop on Catchment Scale Hydrologic Modeling and Data Assimilation, Wageningen, September 3–5, 2001, p. 134.

[128] Elsenbeer H, Vertessy RA. Stormflow generation and flowpath characteristics in an Amazonian rainforest catchment. Hydrol Process 2000;14:2367–81.