Searching for the Holy Grail of scientific hydrology: Qt=(S, R, ?t)A as closure

K. Beven

**Hydrology and
Earth System
Sciences**

Searching for the Holy Grail of scientific hydrology: $Q_t = H(S, R, \Delta t)A$ as closure

K. Beven

Environment Science/Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK

Abstract. Representative Elementary Watershed concepts provide a useful scale-independent framework for the representation of hydrological processes. The balance equations that underlie the concepts, however, require the definition of boundary flux closures that should be expected to be scale dependent. The relationship between internal state variables of an REW element and the boundary fluxes will be nonlinear, hysteretic and scale-dependent and may depend on the extremes of the heterogeneities within the REW. Because of the nonlinearities involved, simple averaging of local scale flux relationships are unlikely to produce an adequate decription of the closure problem at the REW scale. Hysteresis in the dynamic response is demonstrated for some small experimental catchments and it is suggested that at least some of this hysteresis can be represented by the use of simple transfer functions. The search for appropriate closure schemes is the second most important problem in hydrology of the 21st Century (the most important is providing the techniques to measure integrated fluxes and storages at useful scales). The closure problem is a scientific Holy Grail: worth searching for even if a general solution might ultimate prove impossible to find.

1 Introduction

Hydrology is many things to different people. In many areas of the world it is the difference between life and death, flood or drought, plenty or famine. Even in the developed world, it underlies constraints on new urban and industrial development in water scarce areas. In the United States, Libya and elsewhere it is the reason why groundwater is being mined for consumption way beyond annual levels of replenishment by recharge. In the European Union, it underlies the new

Water Framework Directive which states that all designated water bodies in the Union should achieve "good ecological status" for "sustainable use" by the year 2015. In China and elsewhere it underlies calculations of water yields for reservoir design, and of the sediment yields that will determine the life expectancy of any new reservoirs.

Scientific hydrology is critical to all these issues and is actually used in virtually none. Instead, we have a wide range of hydrological tools, based on the analysis of data collected in the past and expressed as empirical relationships or calibrated hydrological models. The models may purport to be scientific but, as argued elsewhere, that argument is undermined by all too evident deficiencies in model structures and the need for calibrated effective parameters to compensate for those deficiencies in any particular application (Beven, 2000, 2001a, 2002a, b).

How can hydrology as science help to remedy this situation? As stated in the preamble to the Science Plan for the IAHS PUB (Prediction in Ungauged Basins) project (see http://cee.uiuc.edu/research/pub/, Sivapalan et al., 2003), the development of scientific hydrology to allow prediction of the response of ungauged catchment areas is a major goal for the future. The ungauged catchment problem has long been considered as *the* problem in hydrological prediction but PUB also recognises that it may not be possible to make entirely accurate predictions of the response of ungauged catchment areas, since we will never be able to know the characteristics of those catchments in sufficient detail to allow a full description of the hydrology. PUB is therefore effectively an exercise in using science to constrain uncertainty in the predictions.

This is undoubtedly the right strategy, but raises the question of how best to constrain those uncertainties in model representations and parameters. This is not a problem unique to hydrology. It applies to all areas of environmental science, including the fields of hydraulics and the atmosphere where there is a stronger argument for the application of scientific

Correspondence to: K. Beven
(k.beven@lancaster.ac.uk)

principles without calibration because in these flow domains, at least at larger scales, the flows are self-organising and less affected by poorly defined boundary conditions (see, for example, discussion of Beven and Pappenberger, 2003, and reply by Abbott et al., 2003). In hydrology, dominated by shallow free surface flows over irregular surfaces and subsurface flows in heterogeneous porous and fractured media, and with mass fluxes of precipitation and evapotranspiration that are poorly known, the boundary conditions *are* the science.

This is reflected very nicely in the REW (Representative Elementary Watershed) theory developed by Reggiani et al. (1998, 1999, 2000; Reggiani and Shellekens, 2003). Even if we still cannot close the mass balance equations of hydrology by measurement (see Beven, 2001b), we can theorise that hydrological flows must be consistent with the balance equations for mass, energy and momentum. The REW theory demonstrates how these principles can be applied at any scale in hydrology. All that is required is to then specify the boundary fluxes of mass, energy and momentum between the REW elements at the chosen scale. This is the so-called closure problem.

The science of hydrology depends on this closure problem. It is effectively the "Holy Grail" of scientific hydrology. Finding *the* solution to the closure problem (if such a solution is possible at all) would be the defining act of scientific hydrology (it would not be the *end* of scientific hydrology as there would still be some interesting problems associated with applying that solution to different applications in different places, see, for example, Beven, 2000, 2002).

The closure problem requires that the fluxes across the boundaries of REW elements be specified in some way, either by measurement or by some functional representation. With the exception of surface discharge at a point, we have no measurement techniques for measuring the integral fluxes over an REW boundary. Even our surface discharge measurements are often associated with significant uncertainty. This is why it is so difficult to verify any of the balance equations, including the water balance by measurement. It is also why it will be very difficult to develop functional representations of the fluxes as a function of internal states.

There will also be cases where the boundary fluxes will be dependent on the internal states of contiguous elements. This will be the case for example, in an aquifer where an REW element represents only part of the flow domain. For both confined and unconfined cases, the lateral and vertical fluxes between elements will depend on the head gradient. For a Darcian aquifer, with homogeneous properties and a smooth piezometric surface, we have the theory to represent those boundary fluxes. Even in heterogeneous Darcian cases, the spatial and temporal integration of fluxes will be quasi-linear and developing a functional relationship will be possible. In the case of recharge through a heterogeneous unsaturated zone, and non-Darcian aquifers, however, this will not be the case and, again, it will be very difficult to develop

functional representations of the fluxes as a function of the internal states in the contiguous REW elements.

It might be possible to develop theoretical representations for such fluxes, if we knew what the internal characteristics of an REW element (and adjacent element if necessary) were. *Hypothetical* studies of fluxes in heterogeneous surface and subsurface flow domains have been widely reported in the literature (e.g. Freeze, 1980; Sivapalan et al., 1987; Binley et al., 1989, 1991; Bashford et al., 2002; Weiler and Mac-Donnell, 2005). The hypothetical studies, of course, have the advantage that everything about the domain is known. This will be impossible in any system of interest, we simply do not have the non-intrusive measurement techniques that would make this possible. There have, in fact, been very few detailed studies of field sites (with the exception of some of the large scale groundwater tracing experiments) where it has been attempted to characterise the statistics of the properties of a flow domain.

So what are the implications of this? The REW approach to scale-dependent modelling seems to be general and the only approach that does not (in principle) involve some ad hoc assumptions. Application of the REW approach, however, requires closure of the balance equations by the representation of the fluxes. These fluxes cannot generally be measured at any scales of real interest, nor can the element characteristics on which they depend, and therefore the introduction of ad hoc assumptions is inevitable. It would appear as if the analogy of the REW closure problem as the Holy Grail of Scientific Hydrology may be apt – the solution is (or might be) out there somewhere in principle, but may be impossible to find.

But the foundations of the REW closure problem are there in *every* hydrological model, lumped or distributed, under one set of approximations or another. The same principles can be extended to *every* water quality or sediment transport model. As with plausible forgeries of the Holy Grail, if the real solution cannot be found, we naturally tend to substitute an approximation to it. Indeed, we are forced into making inadequate approximations since the problems that require solving are very practical and do not go away.

That does not mean that the search for the Holy Grail is not worthwhile. New measurement techniques for any of the integrated boundary fluxes or internal states may get us closure to a reasonable representation of the nonlinearities. An exploration of hypothetical cases or virtual realities may suggest forms for the integrated fluxes that can be parameterised in a relatively simple way, even if they can only be suggestive when lacking detailed information about any particular real system of interest (see for example Bashford et al., 2002; Zehe et al., 2006).

So how should we deal with the closure problem as a "Holy Grail" of Scientific Hydrology? Is it possible to have some searches (model approximations) that are more intelligent than others, particularly in allowing that some parameterisations might be more appropriate in some circumstances

than in others? How best is it possible to constrain the uncertainty in the search, given the current limitations in our measurement techniques? How best is it possible to search for the representation of the boundary fluxes of an ungauged catchment? It would seem that the closure problem is fundamental, critical to the success of PUB and, indeed, to the future of scientific hydrology.

As such it demands that the best hydrological minds address the issues involved. And yet it seems to have so far been largely ignored as an issue. Even recent implementations of the REW concepts, such as in the REWv4.0 model (Fenicia et al., 2005; Reggiani and Rientjes, 2005; Varado et al., 2006), the CREW model (Lee et al., 2006; Zehe et al., 2006), and the REWASH model (Zhang and Savenije, 2005) have used much the same small scale laboratory homogeneous domain theory to represent the integrated fluxes at the much larger scales of hillslope and catchment elements – or simple conceptual representations of fluxes without any strong physical justification for their use at any given scale or across scales (see Zhang et al., 2005; Varado et al., 2006; Lee et al., 2006). This is not adequate; it is searching for the Holy Grail where many others have looked and not found it (in fact, for perfectly good theoretical reasons, see for example, Binley et al., 1989 – so perhaps it was not actually worth looking there, Beven, 2002a).

2 So if we have been searching in the wrong places, how can we change direction?

The REW research programme is predicated on taking a physics-based approach to hydrological science. Hence the use of energy balance and momentum balance constraints in expanding the number of equations in the original papers, even though the energy and momentum constraints do not actually provide strong or useful constraints except in some special cases: the uncertainties in the boundary fluxes and losses in the system are just too high. However, if we continue to apply physical concepts at least in principle to the closure problem, then certain issues become clear (at least in principle).

The first is that the REW approach will not result in a representation that is consistent with continuum mechanics representations at any useful scale. It is a control volume representation, but it does not reduce to a continuum description as the size of the control volume becomes smaller because of the interaction between nonlinearities and heterogeneities in the system. Thus, gradient and divergence terms do not average up in any useful way. Thus, the effects of variability in properties, gradients and divergences at the sub-REW scale will require parameterisation directly at the scale of the REW, and the nature of that representation may change as the REW scale varies. This should be expected as a result of the physics. There is an analogy here with scale dependent closure schemes for the representation of turbulence in fluid dynamics models, but it is not a complete analogy because of the possibility of using continuum representations of pressure and velocity gradients in such models that is not really appropriate for flows through the soil because of the dominance of local boundary conditions over flow dynamics.

The second inference is that, at the REW scale, the boundary fluxes will depend on the processing of inputs within the REW volume and, where the REW scale has important connections to other REWs, on the evolution of storage and fluxes in hydraulically connected REWs. When the hydraulic connections both within and between REWs are dependent on multiple pathways and residence times in the system, these dependencies will be complex and may not necessarily lead to simple functional relationships between average storages or average gradients of potential and the boundary fluxes.

If, within a hydrological control volume there is enough variation in hydraulic gradients and permeabilities to cause significant local spatial and temporal variability in flux rates within the volume, then the small scale equations (such as the Darcy-Richards equation for unsaturated flow) will not integrate to give the same equation at the control volume scale (Beven, 1989, 2002a). This has not been properly appreciated in the development of physically-based models (almost certainly because no other equations have been developed to replace the point scale equations). It will therefore be necessary to search for another way of representing the fluxes based on state variables for the area.

To take a simple example, consider a complete hillslope hydrological unit as the REW. Allow that surface and discharges from this slope could be collected and measured. Start with an initial condition of the soil being just at saturation and allow to drain and cover the surface to reduce evapotranspiration to negligible values. Many hillslope experiments of this type have been carried out in the past (e.g. Hewlett and Hibbert, 1963). A drainage characteristic curve could then be measured over days or weeks, analogous to the primary drying curve measured on a small soil sample in determining the soil moisture characteristic curves.

Once a sufficient period of drainage has elapsed, start to wet the hillslope unit (initially at least at below the infiltration capacity of the soil). In wetting, the discharge at a given storage level will not be the same as for the primary drying curve. There will be hysteresis in the storage-discharge relationship. This will be due to hysteresis in the small scale matrix soil characteristics, the possibility of changing vertical and downslope connectivities of flow pathways as the soil dries and rewets, the possibility of by-passing of available matrix storage by preferential flows and fingering during wetting, the possibility of threshold effects in local flux storage relationships, the development of the inter-unit patterns of antecedent wetness, dynamics of contributing areas, spatial structure in the soil depths and permeabilities that might lead to perched saturation or other complex flow pathways, and the effects of routing delays within the unit. We can also

envisage, therefore, a primary wetting curve for the hillslope unit (although this might vary with the rate at which the unit is wetted, see below), and, during a sequence of storms the possible secondary scanning curves for the relationship between storage and discharge flux.

Hysteresis has already, of course, arisen in soil physics (since at least the work of Haines, 1930). In applying the the Darcy-Richards equation it was found that the soil moisture characteristic curves that link soil moisture content, soil water potential and hydraulic conductivity, were not single valued but were rather different for wetting and drying (see the review of concepts and models by Jaynes, 1990). Thus, the soil moisture characteristics must be considered to be hysteretic, even though this is very often forgotten in applying the Darcy-Richards equation because there is little information available on the nature of the hysteresis for different soils (what there is has been collected in the GRIZZLY database of Haverkamp et al., 2002), multiple physical hypotheses about the causes of hysteresis, and no consensus about how it should be parameterised (see the recent discussions of O'Kane, 2004; Flynn et al., 2005).

In fact, the need to introduce hysteresis into soil physics is itself an indication of the failure of the continuum mechanics hypothesis already at these small ("representative elementary volume") scales. Allowing for hysteretic soil moisture characteristic functions is a fix for this failure. Use of mono-valued characteristic functions (as in nearly all profile or hillslope scale hydrological models, including my own) is already a departure from our understanding of the physical principles that underlie hydrology.

So, why are such departures accepted so easily, even at small scales? Is it just because hysteretic soil moisture characteristics for different soils are not readily available from the literature and are time consuming and costly to measure, even for small samples? Or is it because we believe that the effects of heterogeneity of soil properties or preferential flow pathways are more important than hysteresis at scales of interest (but where is the evidence?)? Or do we believe that at larger REW scales the effects of hysteresis can be averaged out by using effective parameter values in the same way assumed to allow for heterogeneities and preferential flow pathways (but where is the evidence?)?

Consider the hillslope scale REW again. The input to the hillslope will induce a response in the pattern of storage and, with some delay, a response in the boundary output flux (whether by surface or subsurface flow pathways). Predicting this boundary flux as a function of the evolution of the storage is the essence of the closure problem (remembering that in some circumstances, this may also depend on boundary fluxes from upslope REWs and patterns of storage in downslope REWs).

Looked at in this way, ALL hydrological models proposed in the past are conceptual contenders for the representation of the closure problem. By analogy with the earliest of these models (Mulvaney's rational method, see Beven, 2001), we could define the closure problem in the form:

$$Q_t = H(\underline{S}, \underline{R}, \Delta t)A$$

where Q_t is the boundary flux from an REW (or REW subsystem), A is the area of the REW, $\underline{S}$ is the past trajectory of REW storage, $\underline{R}$ is the past pattern of rainfall (and other) inputs, Δt is the length of time step considered and $H()$ is a nonlinear hysteretic function.

3   Hysteretic flux closure in practice

We can easily appreciate that any function $H(\underline{S}, \underline{R}, \Delta t)$ will be complex. This is one reason why we have so many competing hydrological models (another is the difficulty of identifying an appropriate function given the error and uncertainties in the measured inputs and output data at the catchment scales at which these models are applied, Beven, 2001d). The (surface or subsurface) runoff coefficient for rainfall over an REW area (regardless of how it is defined) is nonlinearly and complexly dependent on antecedent conditions, and the pattern of rainfall intensities in time and space. Similar considerations hold for evapotranspiration fluxes, which may depend on patterns of soil depth, the effects of downslope fluxes on available fluxes, boundary layer affects where there are changes in vegetation characteristics, and so on. Deriving functional forms for these nonlinearities is difficult, because it will depend on the heterogeneities of the catchment area, as expressed in the distribution of antecedent storage, soil, vegetation, rainfall and other characteristics and their arrangement in space.

An important point, emphasised in the "representative elementary area" concept (Beven et al., 1988; Wood et al., 1988) is that the extremes of these distributions may dominate the runoff response, e.g. the lowest infiltration capacities and highest rainfall intensities in the production of infiltration excess runoff, the extremes of topographic controlled contributing areas in the production of saturation excess runoff, the extremes of hydraulic conductivities (in percolines and fractures) in the production of subsurface stormflow. These extremes may (or may not) be related in a useful functional way to the mean or modal characteristics of the relevant characteristics. It can be argued that there will never be enough information about a catchment to specify the distributions of these characteristics in a way adequate to allow the prediction of fluxes (Beven, 2000, 2004). Certainly, as noted above, the responses over the distribution of characteristics will not average in a linear way, so that the scaling of responses based on small scale measurements to fluxes at the boundaries of larger area will be very difficult, if not impossible. Averages of variable dynamic quantities over the REW volume, such as capillary potentials, may have little relevance to the closure relationship.

So far, this hysteretic representation has dealt only with a storage-flux relationship at the REW scale. The way in which local characteristics at the sub-unit level might produce this hysteresis has not been considered, nor has the possibility of fluxes being controlled by the patterns of storage in adjacent units. The latter might be important in the case of fluxes in a deeper saturated zone (noting that flow and dispersion parameters for fluxes between elements may still be scale dependent). However for shallow partially saturated systems at hillslope scales, and where the water table maintains a characteristic shape in wetting and drying, treating REW elements as independent of surrounding elements might not be an important problem. The challenge then is to find appropriate functional forms for representing the hysteretic storage-discharge relationship given (generally) very little information about the internal characteristics of the unit, very little observable data in the way of storage or discharge measurements at the hillslope unit scale, and no theoretical framework on which to base such a representation.

For the hillslope REW case, we have neither theory or measurements for the general case. We need, however, to start somewhere so following the example of other applications of the REW concepts which have used sub-catchment discretisations as the basis for defining REW units, we will start with a unit that is a watertight representative elementary watershed (REW) draining to a first order stream where we can measure the outflow as discharge in the stream. We will also assume that we can have good estimates of inputs and evapotranspiration from the REW by having sufficient rain and throughfall gauges and eddy correlation flux measurement sites. In principle, we can then derive the changes in bulk storage over time within the REW by mass balance (as long as we do not wish to break down the subsurface storages into saturated and unsaturated zone, accepting that there will inevitably be some error associated with each of the estimations of the fluxes). We can then plot each of the elements of the water balance to check if there are any unexpected or unphysical trends in the storage term, and also plot storage against discharge to see if there is any consistent hysteretic behaviour.

These plots are demonstrated for several winter and spring periods (when evapotranspiration effects should be small) for different small catchments treated as REW elements in Figs. 1 and 2. Figures 1a and 2a show the hydrographs for the periods analysed; Figs. 1b and 2b show plots of relative storage and relative runoff for the two catchments; Figs. 1c and 2c show another way of looking at potential hysteresis in the data by using dynamic maps (here using a 3-h time delay). Note that we consider only the combined storage changes in unsaturated and saturated zone sub-REW components in assessing the relative storages in these plots. The Greenholes Beck catchment (Fig. 1), on Caton Moor near Lancaster in the UK is a catchment dominated by podzolic soils that stay close to saturation in the winter. The Slapton Wood catchment (Fig. 2) has generally more permeable soils,

underlain by fractured Devonian shales, and shows an important subsurface contribution to runoff with superimposed small fast runoff peaks. These systems demonstrate hysteresis loops, with some apparent consistency of timing if not in magnitude, associated with individual storms. However, for Greenholes Beck, in particular, it appears that the range of the hysteresis for each storm is small relative to the range of antecedent storage magnitude that develops between each storm. This is also apparent in the dynamic maps (Figs. 1c and 2c). Given the nature of the data period considered, this is unlikely to be entirely due to a poor estimate of the rainfall and evapotranspiration boundary conditions used in the calculation of the evolving relative storage.

Interestingly, if we consider a unit hydrograph model as a way of representing the time transformation of an "effective rainfall" at the REW scale, it provides a very simple form of hysteresis in the storage-discharge relationship. Recent models based on linear transfer function routing methods have been shown to successfully reproduce the response of hydrological systems after an appropriate nonlinear transformation of the inputs to account for the effects of antecedent conditions (e.g. Young and Beven, 1994; Young, 2001, 2003; Young et al., 2004) (see Fig. 3). The transfer function itself does not, of course, help determine the amount of any given rainfall input which contributes to the "effective rainfall". This evident in Fig. 3 where, although the shape of the storage-discharge relationship is similar for both the observed and modelled data, the relative storage values are quite different (both being initialised to relative storage values of zero at the start of the analysis period). Here then is more evidence that the effect of antecedent conditions on the quantity of outputs might dominate the hysteresis effects of routing within a REW. Does this then suggest that it might be possible to develop simple REW scale dependent closure schemes by concentrating on the nonlinearity of the storage-discharge relationship, with hysteresis as a second order effect?

4 Taking advantage of science: closure as multiple competing hypotheses

Such a conclusion is, evidently, a simple reinterpretation of the old, old hydrological modelling problem of how to parameterise the effect of antecedent conditions on runoff generation. Looking back over the history of hydrological modelling, this provides us with a very wide range of choices of conceptual parameterisations as multiple competing hypotheses for the closure problem (see, for example, Beven, 2001c). This suggests an approach to searching for a solution to the closure problem as an evaluation of these multiple working hypothesis (see discussion of Beven, 2002a, b). If we consider all the possible parameterisations of boundary fluxes that are consistent with our (qualitative) perceptual model of the processes controlling a particular boundary
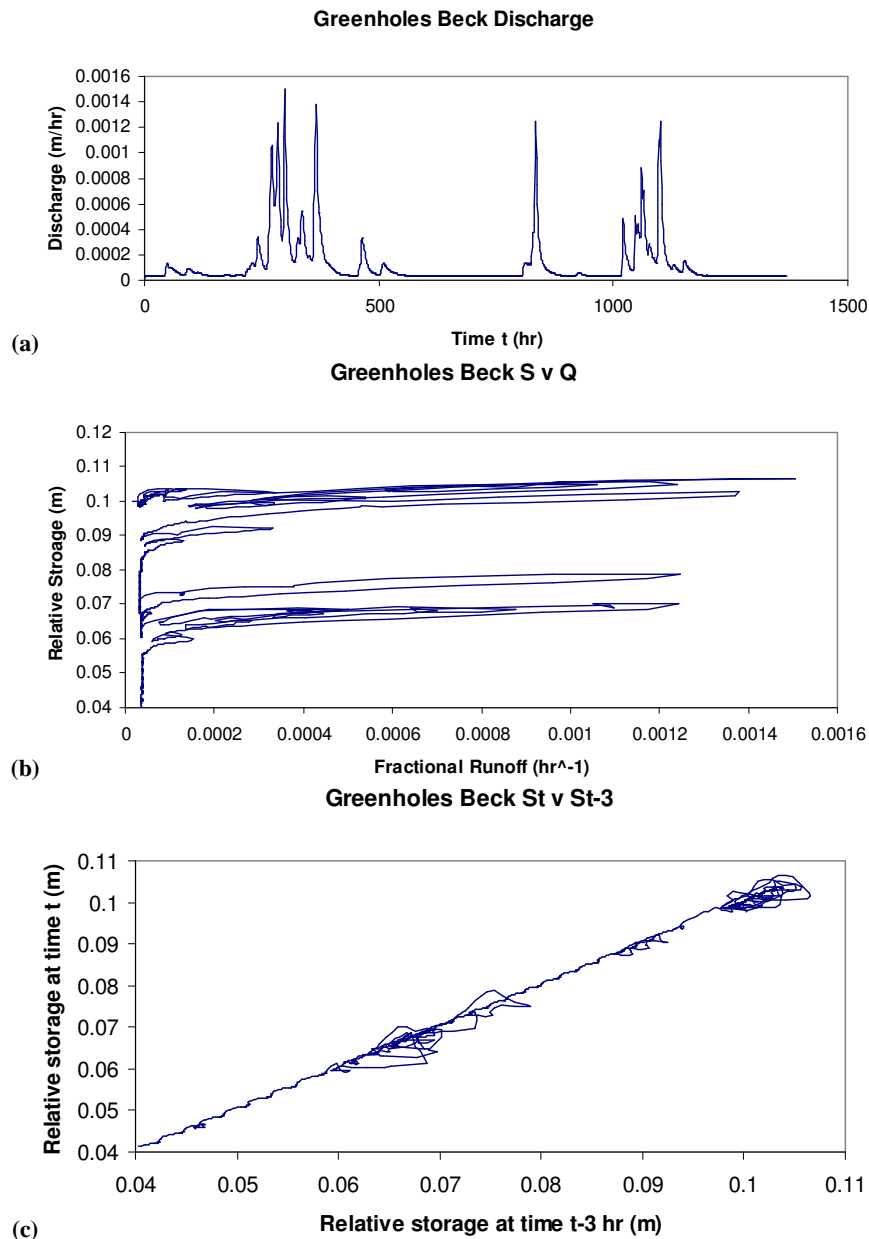
Fig. 1. Plots of **(a)** – catchment hydrograph; **(b)** – relative catchment discharge (as a fraction of input volume) vs. relative catchment storage deficit; **(c)** – relative catchment storage at time t vs. relative catchment storage at time t–3 h; as determined from water balance closure for winter periods assuming actual evapotranspiration at estimated potential rates for the Greenholes Beck catchment, UK (3 km^2).

flux in this way, how far could we evaluate those hypotheses on an ungauged catchment by either a priori evaluation or by taking certain critical measurements? Could we design a programme of measurements that would allow an increasing number of hypotheses to be eliminated? Or must we admit that the search for a unique closure parameterisation is ultimately doomed to failure given the limitations of current measurement techniques – including the very fundamental inherent error in estimating the components of water balance?

It will come as no surprise that I suspect the latter (see, for example, Beven, 2006a), at least for the foreseeable future (all could be changed by the appearance of a reliable measurement technique for integrated fluxes, but that still seems a long way off – point surface discharge measurements and scintillometer measurements of evapotranspiration measurements are the nearest we have, but utility of the former still depends on upstream mass balance closure and the latter depends on an energy balance closure that may also be subject to significant error). But failure to close the water and energy
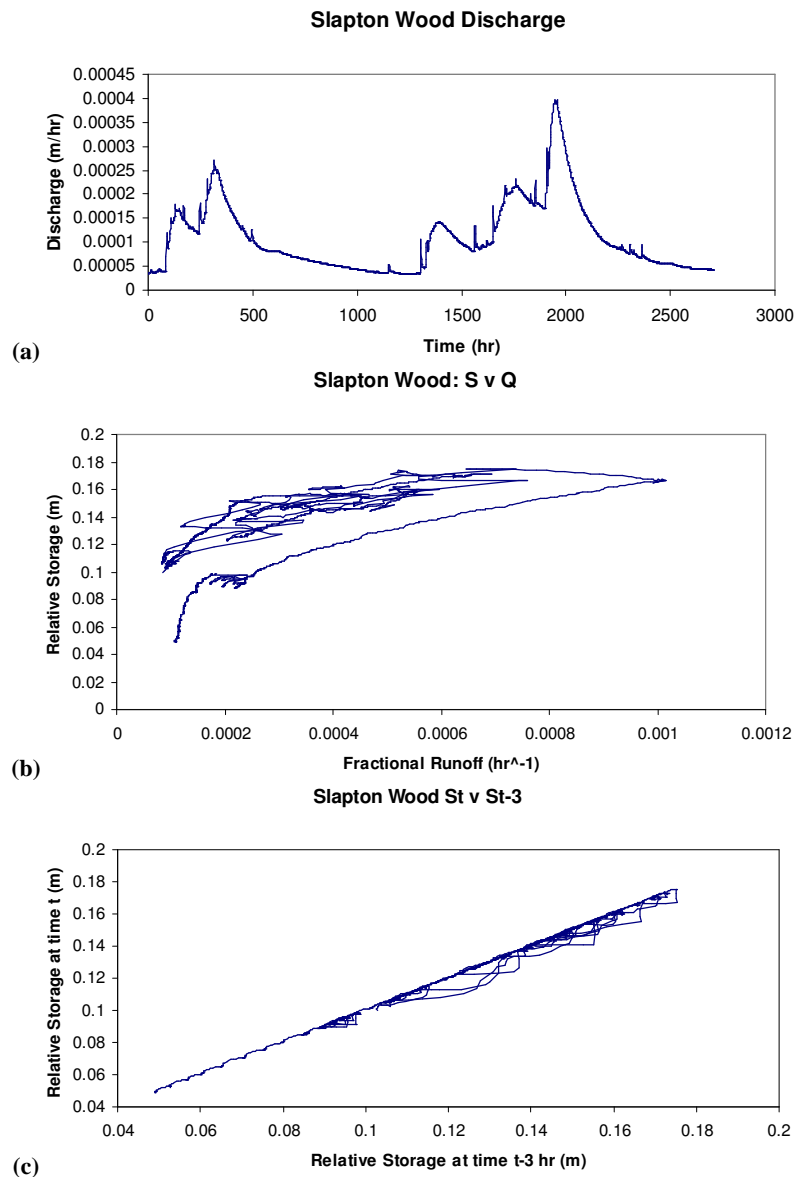
Slapton Wood Discharge

(a)

Slapton Wood: S v Q

(b)

Slapton Wood St v St-3

(c)

Fig. 2. Plots of **(a)** – catchment hydrograph; **(b)** – relative catchment discharge (as a fraction of input volume) vs. relative catchment storage deficit; **(c)** – Relative catchment storage at time t vs. relative catchment storage at time t–3 h; as determined from water balance closure for winter periods assuming actual evapotranspiration at estimated potential rates for the Slapton Wood catchment, UK (1 km^2).

balances by measurement also has implications for hydrological science and hydrological methodology that might be no easier to work out in practice. If, in a given application site, we cannot unequivocally decide what form of parameterisation of the boundary fluxes is appropriate then it will be necessary to retain multiple possibilities (hypotheses) about the form of the parameterisation and the values of the parameters (the equifinality thesis) and look for ways of trying to differentiate between these different hypotheses.

There is also one important additional complicating issue in defining REW scale flux closures that is sometimes forgotten: the difference between the celerity of the discharge response and the velocity of water (and conservative tracer) particles. This difference is one explanation for the results of tracer experiments that suggest that in many small catchments, the hydrograph is dominated by the displacement of pre-event water (see discussion of the history of these ideas in Beven, 2006b). The difference can be illustrated simply within a simplified kinematic wave description of the flow processes (e.g. Beven, 1981; 2001c) but in reality is much more complex because of the effects of heterogeneities, immobile storage, fingering and preferential flows.
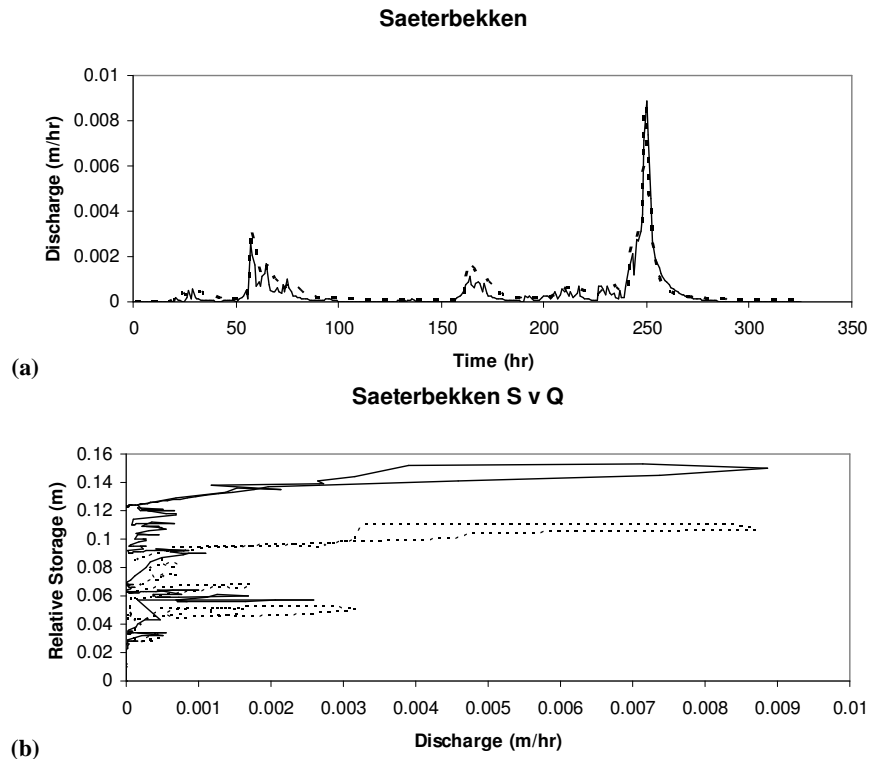
Saeterbekken

(a)

Saeterbekken S v Q

(b)

Fig. 3. Plot of **(a)** – observed and predicted discharges and **(b)** – observed and predicted discharge vs. calibrated storage for a calibrated transfer function of the form of a nonlinear rainfall filter in series with two parallel linear stores for the Saeterbekken catchment, Norway $(0.0075 \, km^2)$. Dotted lines: observed data; solid line: modelled data.

The storage-discharge response will be governed primarily by the celerities with which pressure effects are transmitted through the system. We still have much to learn about the details of this, particularly in unsaturated soils. We still also have much to learn about the velocity distributions and residence time characteristics of the water particles (see, for example, Kirchner et al., 2000, 2001). It is clear, however, that the two types of response must be differentiated.

It must be stressed that these arguments apply to all hydrological models (and to a much wider class of environmental models, Beven, 2002b). The lack of thoughtful approaches to the Holy Grail of the closure problem is one of the reasons why we now have so many different lumped and distributed model formulations to represent catchment responses and why we have mostly been satisfied to find parameterisations acceptable if they work after calibration. But this, surely, is not real Hydrological Science. The REW research programme needs to be completed by the development of appropriate closure schemes that properly consider the potential for the effects of nonlinearities and heterogeneities on the integrated fluxes. The types of storage-discharge and dynamic maps demonstrated here would appear to be a useful first step in thinking about this issue and evaluating the performance of different models. A methodology for applications based on the evaluation of multiple hypotheses for the flux closures needs to be developed, even where this means that multiple hypotheses have to be retained for the present.

These questions are the second most important problem in hydrology of the 21st Century (the most important is providing the techniques to measure integrated fluxes and storages at useful scales). At a time when the CUAHSI initiative in the United States has resulted in a possibility of significant investment in multiscale hydrological research programs (at least at some time in the future), it is absolutely essential that these problems be addressed in a coherent way (rather than simply assuming that looking across scales will eventually allow "bottom-up" physics-based theorising to be successful). Resolving the closure problem (or even trying to do so) will mean that we will then have a hydrological science that we can be proud of, even if the Holy Grail is still, ultimately (as in the "Conte de Graal" or "Morte d'Arthur"), unattainable. We should still contemplate the search as a matter of scientific honour.

References

Abbott, M. B., Babovic, V. M., and Cunge, J. A.: Response to Beven and Pappenberger, J. Hydraul. R., 41(3), 331–333, 2003.

Bashford, K. E., Beven, K. J., and Young, P. C.: Observational data and scale dependent parameterisations: explorations using a virtual hydrological reality, Hydrol. Process., 16(2), 293–312, 2002.

Beven, K. J.: Kinematic subsurface stormflow, Water Resour. Res., 17(5), 1419–1424, 1981.

Beven, K. J.: Uniqueness of place and process representations in hydrological modelling, Hydrol. Earth Syst. Sci., 4, 203–213, 2000,
http://www.hydrol-earth-syst-sci.net/4/203/2000/.

Beven, K. J.: How far can we go in distributed hydrological modelling?, Hydrol. Earth Syst. Sci., 5, 1–12, 2001a.

Beven, K. J.: On hypothesis testing in hydrology, Hydrol. Processes (HPToday), 15, 1655–1657, 2001b.

Beven, K. J.: Rainfall-runoff modelling – the primer, Wiley: Chichester, 2001c.

Beven, K. J.: On hypothesis testing in hydrology, Hydrolog. Processes (HPToday), 15, 1655–1657, 2001d.

Beven, K. J.: Towards an alternative blueprint for a physically-based digitally simulated hydrologic response modelling system, Hydrol. Processes, 16(2), 189–206, 2002a.

Beven, K. J.: Towards a coherent philosophy for environmental modelling, Proc. Roy. Soc. Lond. A, 458, 2465–2484, 2002b.

Beven, K. J.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, 2006a.

Beven, K. J.: Benchmark papers in Streamflow Generation Processes, IAHS Press, Wallingford, UK, 2006b.

Beven K. J., Wood, E. F., and Sivapalan, M.: On hydrological heterogeneity: catchment morphology and catchment response, J. Hydrol., 100, 353–375, 1988.

Beven, K. J. and Pappenberger, F.: Discussion of "Towards the hydraulics of the hydroinformatics era" by Abbott, M. B., Babovic, V. M., and Cunge, J. A., J. Hydraul. Res., 41(3), 331–336, 2003.

Binley, A. M., Beven, K. J., and Elgy, J.: A physically-based model of heterogeneous hillslopes. II. Effective hydraulic conductivities, Water Resour. Res., 25(6), 1227–1233, 1989.

Binley, A. M. and Beven, K. J.: Three-dimensional modelling of hillslope hydrology, Hydrol. Processes, 6, 347–359, 1992.

Fenicia, F., Zhang, G. P., Rientjes, T., Hoffman, L., Pfister, L., and Savenije, H. H. G.: Numerical simulations of the ruonn generation with surface water – groundwater interactions in the Alzette river alluvial plain, Phys. Chem. Earth, 30(4–5 Spec. Iss.), 277–284, 2005.

Flynn, D., McNamara, H., O'Kane, P., and Pokrovskii, A.: Application of the Preisach model to soil-water hysteresis, in: The Science of Hysteresis, edited by: Bertotti, G. and Mayergoyz, I., Elsevier: Amsterdam, 681–736, 2005.

Freeze, R. A.: A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope, Water Resour. Res., 16, 391–408, 1980.

Haines, W. B.: Studies in the physical properties of soil, V. The hysteretis effect in capillary properties, and the modes of moisture distribution associated therewith, J. Agric. Sci., 20, 97–116, 1930.

Haverkamp, R., Reggiani, P., Ross, P. J., and Parlange, J.-Y.: Soil water hysteresis model based on theory and geometric scaling, in: Environmental Mechanics:Water Mass and Energy Transfer in the Biosphere, American Geophysical Union, Washington D.C., 213–246, 2002.

Jaynes, D. B.: Soil water hysteresis: models and implications, in: Process Studies in Hillslope Hydrology, edited by: Anderson, M. G. and Burt, T. P., Wiley: Chichester, 93–126, 1990.

Kirchner, J. W., Feng, X., and Neal, C.: Fractal stream chemistry and its implications for contaminant transport in catchments, Nature, 403, 524–527, 2000.

Kirchner, J. W., Feng, X., and Neal, C.: Catchment-scale advection and dispersion as a mechanism for fractal scaling in stream tracer concentrations, J. Hydrol., 254, 82–101, 2001.

Lee, H., Sivapalan, M., and Zehe, E.: Predictions of rainfall runoff and soil moisture dynamics in a microscale catchment using the CREW model, Hydrol. Earth Syst. Sci. Discuss., 3, 1667–1743, 2006,
http://www.hydrol-earth-syst-sci-discuss.net/3/1667/2006/.

O'Kane, P.: Hysteresis in hydrology, Acta Geophysica Polonica, 53, 373–383, 2005.

Reggiani, P., Sivapalan, M., and Hassanizadeh, S. M.: A unifying framework of watershed thermodynamics: balance equations for mass, momentum, energy and entropy and the second law of thermodynamics, Adv. Water Res., 22, 367–398, 1998.

Reggiani, P., Sivapalan, M., Hassanizadeh, S. M., and Gray, W. G.: A unifying framework of watershed thermodynamics: constitutive relationships, Adv. Water Res., 23, 15–39, 1999.

Reggiani, P., Sivapalan, M., and Hassanizadeh, S. M.: Conservation equations governing hillslope responses: physical basis of water balance, Water Resour. Res., 38, 1845–1863, 2000.

Reggiani, P. and Schellekens, J.: Modelling of hydrological responses: the representative elementary watershed approach as an alternative blueprint for watershed modelling, Hydrol. Processes, 17, 3785–3789, 2003.

Reggiani, P. and Rientjes, T. H. M.: Flux parameterisation in the Representative Watershed (REW) approach: application to a natural basin, Water Resour. Res., W04013, doi:1029/2004WR003693, 2005.

Sivapalan, M., Beven, K. J., and Wood, E. F.: On hydrologic similarity 2. A scaled model of storm runoff production, Water Resour. Res., 23(12), 2266–2278, 1987.

Sivalapalan, M., Takeuchi, K., Franks, S. W., et al.: IAHS Decade on predictions in ungauged basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences, Hydrol. Sci. J., 48(6), 857–880, 2003.

Varado, N., Braud, I., Galle, S., Le Lay, M., Séguis, B., Kamagate, B., and Depraetere, C.: Multi-criteria assessment of the Representative Elementary Watershed approach on the Donga catchment (Benin) using a downward approach of model complexity, Hydrol. Earth Syst. Sci., 10, 427–442,, 2006.

Weiler, M. and McDonnell, J. J.: Virtual experiments: a new approach for improving process conceptualisation in hillslope hydrology, J. Hydrol., 285, 3–18, 2004.

Wood, E. F., Sivapalan, M., Beven, K. J., and Band, L.: Effects

of spatial variability and scale with implications to hydrologic modelling, J. Hydrology, 102, 29–47, 1988.

Young, P. C.: Data-based mechanistic modelling and validation of rainfall-flow processes, in: Model Validation: Perspectives in Hydrological Science, edited by: Anderson, M. G. and Bates, P. D., Wiley, Chichester, 117–161, 2001.

Young, P. C.: Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale, Hydrol. Processes, 17, 2195–2217, 2003.

Young, P. C. and Beven, K. J.: Data-based mechanistic modelling and the rainfall-flow non-linearity, Environmetrics, 5, 335–363, 1994.

Young, P. C., Chotai, A., and Beven, K. J.: Data-Based Mechanistic Modelling and the Simplification of Environmental Systems, in: Environmental Modelling: Finding Simplicity in Complexity, edited by: Wainwright, J. and Mulligan, M., Wiley, Chichester, 371–388, 2004.

Zehe, E., Lee, H., and Sivapalan, M.: Dynamical process upscaling for deriving catchment scale state measures and constitutive relations for meso-scale process models, Hydrol. Earth Syst. Sci. Discuss., 3, 1629–1665, 2006,
http://www.hydrol-earth-syst-sci-discuss.net/3/1629/2006/.

Zhang, G. P. and Sanenije, H. H. G.: Rainfall-runoff modelling in a catchment with complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach, Hydrol. Earth Syst. Sci., 9, 243–261, 2005,
http://www.hydrol-earth-syst-sci.net/9/243/2005/.

Hydrol. Earth Syst. Sci., 10, 609–618, 2006

www.hydrol-earth-syst-sci.net/10/609/2006/