© IWA Publishing 2014 **Hydrology Research** | 45.6 | 2014

'Here we have a system in which liquid water is moving; let's just get at the physics of it' (Penman 1965)

Keith Beven

ABSTRACT

This paper is based on the 2012 Penman Lecture delivered at the 11th National Symposium of the British Hydrological Society. The title is taken from a 1965 interview by Howard Penman when hydrological modelling was just really starting. In the period since then, the idea that we might easily move towards 'physically based' representations of hydrological processes has proven problematic. It is argued that this might best be done within a hypotheses testing framework, where the hypotheses are the model representations of processes over some discrete elements of a catchment, integrating the small-scale variability within those elements. This might still require some distribution function to reflect that variability, since the extremes of the variability might be important in controlling the response. Hypotheses will need to be formulated that reflect the interaction of the water flow pathways and the biota. Testing of those hypotheses will require a proper account of the uncertainties inherent in the study of hydrological systems, including a recognition that many sources of uncertainty are epistemic in nature rather than simple random variability. Such uncertainties will only be significantly reduced by the development of new measurement techniques that provide useful information at the element scales of interest.

Key words | epistemic uncertainties, hydrological physics, hypothesis testing, measurement scales, preferential flows, representative elementary watershed (REW)

Keith Beven
Lancaster Environment Centre,
Lancaster University,
Lancaster LA1 4YQ,
UK;
LUVAL,
Uppsala University,
Uppsala,
Sweden;
and
ECHO,
ENAC, EPFL,
Lausanne,
Switzerland
E-mail: *k.beven@lancaster.ac.uk*

INTRODUCTION

It is nearly 50 years since Howard Latimer Penman suggested that we just get at the physics of moving liquid water, but a glance at any undergraduate hydrology text book would suggest that the physics that was available in 1965 is still being taught today (and had already been around for decades). The quotation of the title (from a BBC interview; see Monteith 1986) refers to his approach in looking at the process of evapotranspiration in the 1940s, rather than implying that there was still a lot of work to be done. In fact, the quotation is particularly apt today because it reflects very nicely the classical tension in hydrology (and other environmental sciences) between the desire to do good science and the underlying complexity of the processes. It is only water moving, so should not pose so great a problem in getting at the physics. However, the Penman equation is similar to other physical relationships in hydrology (Darcy, Richards, Manning …) in two respects. Firstly it is based on the analysis of empirical data that is scale, time and space specific and is therefore an inductive physics that is not well-founded on fundamental principles and is therefore not necessarily applicable to other locations (or at least not without some local calibration to reflect local properties). Secondly it ignores biological processes (except in so far as they are reflected in the empirical coefficients) that might lead to feedback mechanisms in the system. But then Penman's work was originally concerned with well-watered short grass sward to provide a relatively well-defined (if not necessarily generally representative) boundary. A similar situation arose with the work of Lorenzo A. Richards on water flow through porous media. Richards applied an air pressure to small samples to study the steady flow rate of water at different

doi: 10.2166/nh.2014.130

degrees of saturation under equilibrium pressure conditions (Richards 1931). Nowhere in his paper does he suggest that this might apply to soil physics in general. Thus it could be suggested that soil physics is based on the wrong experiment for hydrological applications, in that it excluded flow from all the larger pores that might provide the path of least resistance for water flow under natural flow conditions. So if hydrology is just the study of systems in which liquid water is moving: how do we get at the physics of it?

A CRITIQUE OF HYDROLOGICAL PHYSICS

My first real critique of physics-based hydrological models was in my PhD thesis (Beven 1975) and was conditioned by the experience of building a finite element model of hillslope hydrology based on the Darcy–Richards subsurface flow concepts that were the essence of the blueprint of Freeze & Harlan (1969). The model did not simulate the field observations at all well (the story has already been told in Beven (2001)). One of the reasons for that was the neglect of macropores and preferential flow in the Darcy–Richards formulation (see Beven & Germann 1981, 1982), something that is still lacking from most Darcy–Richards-based hydrological models.

Later I was a member of the first SHE (Système Hydrologique Européen) modelling team (Beven *et al.* 1980), an alternative implementation of the Freeze & Harlan blueprint. At that time, computer time was still an enormous constraint (as evidenced by the 4 km grid squares used in SHE for an application in India; Jain *et al.* 1992), but even as computer constraints have decreased, problems in applying this type of model have persisted (see, for example, the review of 30 years of work with SHE by Refsgaard *et al.* (2010); similar considerations apply to other models of this type, see also the history of modelling the R5 catchment at Coshocton reported by Loague *et al.* (2000, 2005)). Experience with this type of model led to the critical review of Beven (1989a), which concentrated on issues of applying nonlinear, small-scale physics at large element scales subject to heterogeneity and preferential flows. A suggestion that a new blueprint for physics-based models was needed followed (Beven 2002), with some ideas about what was needed in a new approach (Beven 2006a) building on the discrete Representative Elementary Watershed approach of Reggiani *et al.* (2000, 2001) and Reggiani & Rientjes (2005) rather than continuum differential equations.

REFINING A PERCEPTUAL MODEL OF A NEW (ECO)HYDROLOGICAL (BIO)(GEO)PHYSICS

The first step in any modelling framework should always be to outline a perceptual model of the processes. The idea of a perceptual model of hillslope and catchment hydrology was first introduced by Beven (1989b, see also 2012b). The important thing about a perceptual model of a system is that it is only qualitative, and can encompass all the complexity that is perceived as being important in the nature of the processes involved. A variety of expressions of the perceptual model have been published, and have evolved even at a single site (see, for example, McDonnell (1990) and recently Nimmo (2012) and Vidon (2012)). Here we will concentrate on a perceptual model of water flows through the soil.

It is clear that the soil is a complex structure, far more complex than the bundle of capillary tubes that is often used in explaining the nature of the soil moisture characteristics. The structure is organised, both in the form of pores, aggregates and peds at the sub-mm to cm scales, but also by other cracks, pipes, root channels and biopores that extend over tens of cm or more. This means that any point scale measurements made in the soil (of soil moisture content, hydraulic conductivity, or tracer concentrations) should be expected to be highly variable. This type of heterogeneity was recognised in soil moisture measurements by Hills & Reynolds (1969), in hydraulic conductivity measurements by Nielsen *et al.* (1973) and in dye tracer experiments by Flury *et al.* (1994), Weiler & Flühler (2004), Blume *et al.* (2009) and many others. Another expectation that arises from this heterogeneity is that even if the flow in the soil was exclusively laminar, then the hydraulic gradients and unsaturated hydraulic conductivities might vary significantly from point to point. Because in the unsaturated case the flux is a nonlinear function of the unsaturated hydraulic conductivity, then we should not expect the flux over larger areas to be represented by some effective Darcian flux. The real local fluxes across any plane through the soil will average linearly, but the product of some local conductivity and local hydraulic gradient will not (Binley *et al.* 1989).

That is the case for unsaturated Darcian flow in the soil but the situation is further complicated by the fingering and other preferential flows that are commonly revealed by tracing experiments. Preferential flows can by-pass parts of the soil matrix and can result in point measurements of water content or capillary potential reacting at depth during infiltration, before points higher in the soil profile. Since it is known that the overall flux is downwards, a back-calculated hydraulic conductivity for that flux will then be negative. This is readily accepted in a perceptual model of a complex soil but is oddly rarely reported in the literature. In the perceptual model we can also recognise that preferential flows are not only associated with macropores, but can occur as laminar films, and are not always laminar, but can be transitional to turbulent once film thicknesses reach about 0.1 mm or more (see Hincapié & Germann 2009).

In the perceptual model, the supply of water to preferential flow pathways does not require the pathways to be directly connected to the surface supply of water. The supply might be the result of local saturation of the matrix resulting in displacement into larger pore spaces, leading to the suggestion that water in preferential flow pathways during wetting may be pre-event water (water already stored in the catchment before an event) displaced into faster flow pathways (e.g. as in the rapid responses of natural ephemeral pipes studied by Sklash *et al.* (1996)). Generation of preferential flows at the surface might be dependent on the redistribution of incoming rainwater by the vegetation through stemflows, and might be dependent on collection of water in surface hollows, with air escaping through larger pores connected to high points (as originally suggested by Robert Horton; see Beven 2004a, b).

Other complicating issues will arise in drying of the soil when the storage of water in the finer pores of the soil matrix becomes important in controlling both evapotranspiration fluxes and water residence times (and chemistry). The physics of water in the soil matrix is generally represented by the Darcy–Buckingham–Richards equation but there is evidence to suggest that in times of water stress, the roots of plants may grow towards stored water faster than the water can move towards the root under capillary gradients. Deeper rooting plants might also have roots that take advantage of water stored at depth, for example by penetrating bedrock fractures or alluvial aquifers in valley bottoms. Such root pathways might also provide routes for preferential infiltration. Indeed, in a particular example of this, the Jarrah trees of Western Australia (*Eucalyptus marginata*) channel water to tap roots that penetrate the impermeable lateritic crust in the soil such that it gets stored close to those roots in the unsaturated subsoil beneath to be available for later use. The biology then provides a boundary condition for the water movement as opposed to being simply reactive to water shortages.

These types of interactions have suggested that natural vegetation may be self-organising in optimising its characteristics in relation to available water (e.g. Eagleson 2002; Rodriguez-Iturbe & Porporato 2005; Schymanski *et al.* 2007, 2008, 2009). The stripes of the tiger bush landscape found in parts of the Sahel of Africa is a commonly cited example as evidence to support the optimality concept and to provide a test of theoretical reasoning (e.g. Eagleson 2002; Esteban & Fairén 2006). This argument can, however, get a bit circular. Optimality (within some limits of uncertainty) is inferred so long as there is not a major change in boundary conditions (the extreme case being a drought or avalanche that causes a sudden dramatic change in the vegetation community). In the Rio Grande valley in New Mexico, a stable (and presumably optimal) natural community of juniper and piñon pine was changed by the extreme drought of 2002 when many of the piñon pine trees died. With this change of boundary conditions, the re-growth appeared to have a much higher net carbon productivity (Huang *et al.* 2010), which must pose questions for the sense in which the original stand might have been optimal. Such examples suggest that there might be a case to be made for the event/relaxation dynamic view of an evolving system rather than for optimality principles, especially when the time scale of the relaxation might be rather long. Similar issues apply to geomorphological change in the catchments. The response of many catchments in the UK is still defined by materials and topography that are a relict of the last ice age, often with some historical modification by man (field drainage, deforestation/afforestation, agricultural practice, urban development, etc.). We can easily perceive that, in these strongly interacting systems, the history of events and consequent relaxation is important in understanding the hydrological, biophysical and geophysical dynamics. The history is, however, necessarily unknowable in detail placing limitations on quantifying such effects in process-based models.

FUNCTIONAL REQUIREMENTS FOR A QUANTITATIVE (ECO)HYDROLOGICAL (BIO)(GEO)PHYSICS

This perceptual model is undoubtedly complex and will require simplifying before the required (eco)hydrological (bio)(geo)physics can be quantified. Based on the discussion above, a functional specification for a new approach can be outlined. As with any functional specification, it is intended to set a goal for the future even if the defined requirements cannot entirely be met.

As an initial suggestion the following components might be required for process-based hydrological modelling in the short term (i.e. neglecting the evolution of hillslopes and channels over longer periods of time):

1. A representation of the interaction between land cover, input variability, (nonstationary) surface characteristics, antecedent conditions, and the disaggregation of the water balance into runoff and actual evapotranspiration.
2. A representation of the propagation and dissipation of heterogeneous water movement (and solutes and sediments), including preferential flows, that distinguishes velocities and celerities in a way that reflects the scale of a calculation element and provides boundary fluxes for that element.
3. A way of reflecting the impacts of man and other biota in modifying the characteristics of the hydrological response.
4. A way of determining the parameters required by these representations for particular applications in particular places.

There are, of course, existing models that try to include one or more of these requirements, either explicitly or in the use of ‘effective’ parameter values. To my knowledge, however, there is none that satisfactorily meets even these simple functional requirements. In particular, existing models do not take an intellectually satisfying account of the heterogeneities of response at sub-element levels nor of the feedbacks within the biophysical system (at least in relation to my perceptual model of catchment processes).

HYPOTHESIS TESTING OF NEW CONCEPTS

Implementing the functional requirements is, in effect, an exercise in hypothesis testing but not in the classical closed system sense of the laboratory experiment, such as that of L. A. Richards. At the scales of interest, the system cannot be easily controlled and will be predominantly driven by uncontrolled input signals. Thus, as is the case with the Richards equation, the unthoughtful application of laboratory scale hypotheses will not necessarily be that useful at the catchment scale. But testing hypotheses at the catchment scale is problematic (Beven 2010, 2012a). Traditionally hypothesis testing has been considered to be a statistical problem and it has been suggested that a Bayes statistical framework might be useful in hydrological hypothesis testing (e.g. Clark *et al.* 2011). But statistical methods require that uncertainties can be treated as aleatory in nature, whereas we should expect in catchment hydrology that the uncertainties are predominantly epistemic in nature (Beven *et al.* 2008, 2012). Much of what is interesting in hydrology takes place under the soil surface where it is very difficult to observe without disturbance to the system so that knowledge about the nature of subsurface flows will be necessarily limited and subject to epistemic uncertainty. There are also epistemic uncertainties in catchment inputs and discharge estimates. Thus, even the water balance equation can be difficult to prove as a hypothesis at the catchment scale without allowing for significant uncertainty (Beven 2001). Significant epistemic uncertainties remain about interactions within the full biophysical functioning of catchment processes.

So what do we actually need in a catchment scale model as hypothesis? We want a tool that will be useful in prediction (or hindcasting as a means of demonstrating understanding). Many existing catchment scale models have been shown to be useful in this sense from an instrumentalist point of view (good performance is sufficient for acceptance). Some researchers view that as the ‘only’ requirement: to capture the dominant modes of response of the catchment system in a way that has an acceptable mechanistic interpretation but which does not necessarily make any attempt to capture the detailed process physics (e.g. Young 2000, 2003). It can also be argued that if we had measurement techniques that could provide time series of the patterns of integrated water storage over a catchment (rather than the much larger scales of the GRACE satellite measurements) then such models might develop to represent the larger scale hydrological laws demanded by Dooge (1986).

Jim Dooge expressed the issues thus:

> *'Those elements of present flood hydrology that are soundly based on deductions from hypotheses confirmed by data either hydrologic or non-hydrologic would contribute in varying degrees to a scientific theory of flood hydrology at the catchment scale … The endeavour to produce such a theory would be well worthwhile. It would improve our understanding of hydrologic phenomena, improve our decision making in relation to water resources and improve our standing among geophysicists'* (Dooge 1986, pp. 56S–57S).

Dooge suggested that there might be three ways of developing such hypotheses:

1. By expressing relationships at the point continuum scale with assumptions about the variation of the microscale parameters (a bottom-up approach).
2. By disaggregating the nature of macroscale responses given global scaling relationships for topography, soil, vegetation, rainfall patterns, energy budget and so on (a top-down approach).
3. By adapting relationships from other cognate disciplines where similar scaling problems are encountered (geomorphology, ecology).

All three approaches are currently limited by the measurement techniques available, and there currently is no real prospect of dramatic advances. In the meantime, there would seem to be scope to at least try to incorporate more of the process functionality outlined above into a new generation of hydrological models as hypotheses about catchment response (see Beven 2012b, Ch. 9). The difficulty remains that the only way of assessing a model in terms of whether it will be useful in making accurate predictions is to test its consistency with observations in the past or on some surrogate catchment. We should not expect a model that is not consistent with calibration data to be useful so it should be rejected (assuming we have some belief in the calibration data that are available). Indeed, there may be cases where it is justified to reject all the models tried. Note that rejection, properly justified, is a 'good thing'. It forces a reconsideration of what is causing the failure, which might be either the model hypotheses or disinformative data (Beven & Westerberg 2011; Beven *et al.* 2011). It avoids making the Type I error of accepting a model that is not fit for purpose when that model might then be used to (falsely) inform decision making. Unfortunately there appears to be much misuse of models that are not fit for purpose in this way, particularly in assessing the impacts of climate change (Beven 2011).

But it is even more important to avoid making a Type II error; that is rejecting models that might be useful in prediction simply because of errors in the input data and evaluation observations. Statistical inference aims to identify a true model under assumptions about errors being aleatory, but where epistemic errors do not have aleatory properties then this will not lead to correct inference in this sense. That is one reason why the generalised likelihood uncertainty estimation (GLUE) methodology has continued to be explored (Beven & Binley 1992; Beven 2006b, 2009). In GLUE we can set limits of acceptability prior to running a model based on what we know about the uncertainty associated with the available observations in a way that is not based on any particular model run (e.g. Liu *et al.* 2009; Blazkova & Beven 2009). This can be done down to the level of individual observations. All models as hypotheses can then be considered relative to those limits of acceptability.

THE IMPORTANCE OF MEASUREMENT SCALE IN DEVELOPING (BIO)(GEO)PHYSICAL RELATIONSHIPS AND TESTING HYPOTHESES

It has proven elusive to define catchment or sub-catchment or Representative Elementary Watershed (REW) or hillslope physical laws that satisfy the functional requirements set out above. In one sense, all hydrological models are attempts to produce such laws but made without much justification in the fundamental physics. The REW approach tried to do this by applying the principles of mass, energy and momentum balance, but has not (yet) solved the closure problem of defining the boundary fluxes given the state of the REW (see Beven 2006a, 2012b, Ch. 9). This is because we do not have measurements of those fluxes at the REW scale (except for discharge, when the REW is a gauged catchment, and then with some uncertainty; see Beven *et al.* 2011; Westerberg *et al.* 2011; McMillan *et al.* 2012).

It has also proven elusive to develop better laws at profile or plot scales where more detailed measurements are possible and find ways of applying those laws at the larger scales of practical interest (see Beven 1995, 1996). This is not surprising given the heterogeneity of plot and profile characteristics that is commonly found when replicate measurements are made (e.g. Wendt *et al.* 1986; Mohanty *et al.* 1996; Loague & Kyriakidis 1997). Evidently, providing biophysical laws at equivalent scales will be even more difficult.

So there is a need to develop new measurement techniques that would provide integral measurements over such small-scale heterogeneities. Discharge is clearly an integral measure, but there is as yet no method for continuously measuring discharge in an arbitrary channel that would be accurate enough to get good estimates of the incremental discharges as additional catchment elements are added (say with sufficient accuracy to be able to take differences to better than 10% error). Where dilution gauging has been used to do this at a steady flow, the results have revealed significant heterogeneity in discharges per unit area contributing to channel reaches (e.g. Huff *et al.* 1982; Genereux *et al.* 1993). There are also no adequate methods to estimate the integrated inputs to a catchment (to better accuracy than current radars-raingauge network, say 5%). It would be good to have more accurate estimates of integrated actual evapotranspiration (not just single site eddy correlation estimates or single line laser scintillometer estimates). A very valuable measurement would be to be able to continuously follow the bulk storages in linked REWs (however they might be defined). More detailed environmental tracer data (and artificial tracer measurements where possible) would also allow a better resolution of the celerity/velocity issues in catchment response.

With the currently available measurement techniques, it is not easy to 'just get at the physics of it' in assessing hydrological and biogeophysical processes. Small-scale physics is too easily shown to be inadequate when applied at larger scales and larger-scale physics remains measurement technique limited. The search for hydrological laws as advocated by Jim Dooge (1986) has been elusive and it seems that we will need better measurements at REW scales to develop new (bio)(geo)physical hypotheses. Such measurement techniques need to integrate over smaller scales of variability to avoid the need to consider the details of the variability. This was the original intention of the Representative Elementary Area (REA) concepts of Wood *et al.* (1990). They used the Topmodel of Beven & Kirkby (1979) as a way of representing sub-element variability of fluxes and showed that above some scale of catchment (the REA) the 'spatial pattern' of variability could be neglected as long as the statistical distribution was taken into account.

There were certainly issues in these hypothetical REA studies about whether the random realisations of heterogeneities were realistically structured, but Topmodel and its variants remain an interesting example of a representation of sub-REW processes (at least where the necessary assumptions are appropriate). Topographic heterogeneity is represented as a distribution function of a topographic index. Soil variability (in so far as it is reflected in a downslope transmissivity at soil saturation) can also be represented as a distribution function. Fast runoff will be produced first on the parts of the catchment represented by the extremes of those distributions. Introducing a more explicit representation of the unsaturated zone can also introduce length scale dependent hysteresis into the saturated area/storage relationship as shown in the Topkapi model study of Martina *et al.* (2011). Such studies demonstrate that simple but useful REW element scale process representations might be developed but they are not yet complete. Different types of assumptions will be required for different process regimes (Topmodel, for example, has often been misapplied where it is clear that there is not a continuous saturated zone quasi-parallel to the soil surface; see discussion in Beven (2012b), Ch. 6). Heterogeneity of soil properties and preferential flows can be considered to be implicit in a realistic downslope transmissivity profile (perhaps inferred from the drainage of bulk storage as in Nyberg (1995)), but this might not be adequate to represent residence times of water and tracers in the system (see the alternative Multiple Interacting Pathways (MIPs) modelling framework of Davies & Beven (2010, 2012) and Davies *et al.* (2011)). Interactions between flow pathways, vegetation cover and actual evapotranspiration are still not properly integrated. In this respect some ideas might be borrowed from the optimality concepts in Schymanski *et al.* (2007, 2008, 2009) but, as suggested earlier, implemented within an event/relaxation framework.

I have suggested before that a practical implementation of the functional requirements outlined above remains the 'Holy Grail' of hydrological science (Beven 2006a), a metaphor that applies in the sense of an aim that might never be achieved but for which the search is still worthwhile. Unfortunately there just seems to be too much historical inertia associated with classical hydrological process representations and model structures to result in much serious thinking about what a new generation of models should look like (but see Beven 2012b, Ch. 9). This is, in part, the direct consequence of current hydrological training which does not depart much from classical theory even though, as we have seen above, there are good reasons for demanding a paradigm shift in process representations. In teaching hydrological theory to undergraduates I have faced this dilemma of having what appears to be perfectly respectable theory that can be presented as a basis for a science of hydrology while knowing that it has rather fundamental limitations when applied at scales of practical import. In Shaw *et al.* (2010), we did, at least, add some qualifying comment in presenting the Richards equation but this is hardly satisfactory.

I suspect that hydrology is still a subject awaiting the stimulus of new measurement techniques at appropriate scales to progress further theoretically. In the meantime it is difficult to see how a paradigm shift will be instigated. Adding conceptual ideas (such as the type of self-organisational and optimality principles referred to above) does not appear to be sufficient, given the large uncertainties in the basic hydrological variables in the water balance equation in most situations (e.g. Beven *et al.* 2011; McMillan *et al.* 2012). It might be possible to make some advances in this respect by working at the scale of the REW as a small zero-order or first-order sub-catchment where precipitation inputs can be estimated with reasonable accuracy and (at least for some such sub-catchments) a reasonable estimate of discharges can be obtained by a surface water measurement. Independent estimates of actual evapotranspiration would also be useful to provide a check (albeit uncertain) on bulk storage changes. If these data are combined with whole environmental or artificial sub-catchment tracer data at fine scale (sub-hydrograph) time resolution, then at least we could demand of a model that it reproduce both hydrological and tracer data adequately at the sub-catchment scale. Variability in such relationships across sub-catchments would then be particularly interesting to investigate (see, for example, McGuire *et al.* (2005) and Tetzlaff *et al.* (2009, 2011), albeit working with less frequent isotope sampling at larger scales).

SUMMARY

There was a time, including when I started my PhD thesis many years ago, that hydrological modellers were confident about developing laws and models based on the physics of water flows. This is expressed nicely in the quotation from Penman used as the title of this piece. Over time, that confidence has been undermined by a recognition of the complexity of hydrological systems, by the difficulty of scaling up laboratory theory to larger scales, by the additional complexity introduced by the biota including the impact of man, and by measurement limitations in observing hydrological responses and inferring parameter values. The tension between the desire to produce (bio)(geo)physical theory and this complexity is illustrated very nicely by the framework of the Representative Elementary Watershed. This discrete representation of the elements of a catchment can be expressed deductively by the (bio)(geo)physical balances of mass, energy and momentum, but each of the equations requires closure fluxes to be specified which are much more difficult to derive theoretically. Indeed, implementations of the REW concepts to date have mostly relied on inductive inference to obtain parameter values for some conceptual representations of those closure fluxes.

In fact, it is not simple even to verify mass, energy and momentum balances for any discrete catchment element because of a lack of appropriate measurement techniques. This has led to a greater appreciation for the uncertainties that are intrinsic to the study and prediction of catchment systems and, gradually, to a greater recognition that not all uncertainties can be easily treated using formal statistical methods because they stem from a lack of knowledge rather than random variability. It remains a subject of research as to how such epistemic uncertainties might best be treated in the modelling process and in hypothesis testing for process representations.

Hydrological modelling was just really starting when Penman was interviewed in 1965. We have made significant progress in understanding the nature of the problem in the nearly 50 years since then, but incorporating the physics has not proven to be so simple and there remains much to be done in both process representations and handling the uncertainties in the modelling process.

ACKNOWLEDGEMENTS

It is always an honour to be invited to give a named lecture, particularly the Penman Lecture as a long time member of BHS. The ideas presented here have developed over a long period of time and have been influenced by discussions with Mike Kirkby, George Hornberger, Peter Germann, Peter Young, Bruno Ambroise, Jeff McDonnell, Jim Freer, Rob Lamb, Stewart Franks, Sarka Blazkova, Paul Smith, Lenny Smith, Andrea Rinaldo and, of course, many others. This can now be considered to be a contribution to the CREDIBLE consortium project funded by the UK Natural Environment Research Council (Grant NE/J017299/1).

REFERENCES

Beven, K. J. 1975 A Deterministic Spatially Distributed Model of Catchment Hydrology. unpublished PhD Thesis, University of East Anglia, Norwich, UK.

Beven, K. J. 1989a Changing ideas in hydrology: the case of physically based models. *J. Hydrol.* **105**, 157–172.

Beven, K. J. 1989b Interflow. In: *Proceedings of NATO ARW on Unsaturated Flow in Hydrological Modelling* (H. J. Morel-Seytoux, ed.). Kluwer, Dordrecht, pp. 191–219.

Beven, K. J. 1995 Linking parameters across scales: sub-grid parameterisations and scale dependent hydrological models. *Hydrol. Process.* **9**, 507–526.

Beven, K. J. 1996 The limits of splitting: hydrology. *Sci. Total Environ.* **183**, 89–97.

Beven, K. J. 2001 On hypothesis testing in hydrology. *Hydrol. Process.* **15**, 1655–1657.

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

Beven, K. J. 2004a Surface runoff at the Horton Hydrologic Laboratory (or not?). *J. Hydrol.* **293**, 219–234.

Beven, K. J. 2004b Robert Horton's perceptual model of infiltration. *Hydrol. Process.* **18**, 3447–3460.

Beven, K. J. 2006a The Holy Grail of Scientific Hydrology: $Q_t = H(S, R, \Delta t)A$ as closure. *Hydrol. and Earth Systems Science* **10**, 609–618.

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

Beven, K. J. 2009 *Environmental Modelling: An Uncertain Future?* Routledge, London.

Beven, K. J. 2010 Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models. *Hydrol. Process.* **24**, 1537–1547.

Beven, K. J. 2011 I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrol. Process.* **25**, 1517–1520.

Beven, K. J. 2012a Causal Models As Multiple Working Hypotheses about Environmental Processes. Comptes Rendus Geoscience, Académie de Sciences, Paris.

Beven, K. J. 2012b *Rainfall-Runoff Modelling: The Primer*. 2nd edn, Wiley-Blackwell, Chichester, UK.

Beven, K. J. & Binley, A. M. 1992 The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* **6**, 279–298.

Beven, K. J. & Germann, P. F. 1981 Water flow in soil macropores, II. A combined flow model. *J. Soil Sci.* **32**, 15–29.

Beven, K. J. & Germann, P. F. 1982 Macropores and water flow in soils. *Water Resour. Res.* **18** (5), 1311–1325.

Beven, K. J. & Kirkby, M. J. 1979 A physically-based variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* **24** (1), 43–69.

Beven, K. J. & Westerberg, I. 2011 On red herrings and real herrings: disinformation and information in hydrological inference. *Hydrol. Process.* **25**, 1676–1680.

Beven, K. J., Smith, P. J. & Freer, J. 2008 So just why would a modeller choose to be incoherent? *J. Hydrol.* **354**, 15–32.

Beven, K., Smith, P., Westerberg, I. & Freer, J. 2012 Comment on 'Pursuing the method of multiple working hypotheses for hydrological modeling' by P. Clark *et al.* *Water Resour. Res.* **48**, W11801, doi:10.1029/2012WR012282.

Beven, K. J., Smith, P. J. & Wood, A. 2011 On the colour and spin of epistemic error (and what we might do about it). *Hydrol. Earth Syst. Sci.* **15**, 3123–3133.

Beven, K. J., Warren, R. & Zaoui, J. 1980 SHE: towards a methodology for physically-based, distributed forecasting in hydrology. In *International Conference on Hydrological Forecasting*, Oxford, UK, April 1980. Int. Assoc. Sci. Hydrol. Pubn, No. 129, pp. 133–137.

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

Blazkova, S. & Beven, K. J. 2009 A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resour. Res.* **45**, W00B16.

Blume, T., Zehe, E. & Bronstert, A. 2009 Use of soil moisture dynamics and patterns at different spatio-temporal scales for the investigation of subsurface flow processes. *Hydrol. Earth System Sciences* **13**, 1215–1234.

Clark, M., Kavetski, D. & Fenicia, F. 2011 Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resour. Res.* **47**, W09301.

Davies, J. & Beven, K. J. 2010 Scaling implications of a new multiple interacting pathways hillslope model. BHS International Hydrology Symposium 2010.

Davies, J. & Beven, K. J. 2012 Comparison of a multiple interacting pathways model with a classical kinematic wave subsurface flow solution. *Hydrol. Sci. J.* **57**, 203–216.

Davies, J., Beven, K. J., Nyberg, L. & Rodhe, A. 2011 A discrete particle representation of hillslope hydrology: hypothesis testing in reproducing a tracer experiment at Gårdsjön, Sweden. *Hydrol. Process.* **25**, 3602–3612.

Dooge, J. C. I. 1986 Looking for hydrologic laws. *Water Resour. Res.* **22**, 46S–58S.

Eagleson, P. S. 2002 *Ecohydrology: Darwinian Expression of Vegetation Form and Function*. Cambridge University Press, Cambridge.

Esteban, J. & Fairén, V. 2006 Self-organized formation of banded vegetation patterns in semi-arid regions: a model. *Ecolog. Complexity* **3** (2), 109–118.

Flury, M., Flühler, H., Jury, W. A. & Leuenberger, J. 1994 Susceptibility of soils to preferential flow of water: A field study. *Water Resour. Res.* **30** (7), 1945–1954.

Freeze, R. A. & Harlan, R. L. 1969 Blueprint for a physically-based, digitally-simulated hydrologic response model. *J. Hydrol.* **9**, 237–258.

Genereux, D. P., Hemond, H. F. & Mulholland, P. J. 1993 Spatial and temporal variability in streamflow generation on the West Fork of Walker Branch Watershed. *J. Hydrol.* **142** (1), 137–166.

Hills, R. C. & Reynolds, S. G. 1969 Illustrations of soil moisture variability in selected areas and plots of different sizes. *J. Hydrol.* **8**, 27–47.

Hincapié, I. & Germann, P. F. 2009 Impact of initial and boundary conditions on preferential flow. *J. Contam. Hydrol.* **104**, 67–73.

Huang, C.-Y., Asner, G. P., Barger, N. N., Neff, J. C. & Floyd, M. L. 2010 Regional aboveground live carbon losses due to drought-induced tree dieback in piñon juniper ecosystems. *Remote Sens. Environ.* **114**, 1471–1479.

Huff, D. D., O'Neill, R. V., Emanuel, W. R., Elwood, J. W. & Newbold, J. D. 1982 Flow variability and hillslope hydrology. *Earth Surf. Process. Landf.* **7** (1), 91–94.

Jain, S. K., Storm, B., Bathurst, J. C., Refsgaard, J. C. & Singh, R. D. 1992 Application of the SHE to catchments in India: Part 2. Field experiments and simulation studies with the SHE on the Kolar subcatchment of the Narmada River. *J. Hydrol.* **140**, 25–47.

Liu, Y., Freer, J. E., Beven, K. J. & Matgen, P. 2009 Towards a limits of acceptability approach to the calibration of hydrological models: extending observation error. *J. Hydrol.* **367**, 93–103.

Loague, K. M. & Kyriakidis, P. C. 1997 Spatial and temporal variability in the R-5 infiltration data set: Déjà vu and rainfall–runoff simulations. *Water Resour. Res.* **33**, 2883–2896.

Loague, K. M., Gander, G. E., VanderKwaak, J. E., Abrams, R. H. & Kyriakidis, P. C. 2000 Simulating hydrologic response for the R-5 catchment: a never ending story. *J. Floodplain Manage.* **1**, 57–83.

Loague, K. M., Heppner, C. S., Abrams, R. H., Carr, A. E., VanderKwaak, J. E. & Ebel, B. A. 2005 Further testing of the integrated hydrology model (InHM): Event-based simulations for a small rangeland catchment located near Chickasha, Oklahoma. *Hydrol. Process.* **19**, 1375–1395.

Martina, M. L. V., Todini, E. & Liu, Z. 2011 Preserving the dominant physical processes in a lumped hydrological model. *J. Hydrol.* **399**, 121–131.

McDonnell, J. J. 1990 A rationale for old water discharge through macropores in a steep, humid catchment. *Water Resour. Res.* **26**, 2821–2832.

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M. & Seibert, J. 2005 The role of topography on catchment-scale water residence time. *Water Resour. Res.* **41** (5), W05002.

McMillan, H., Krueger, K. & Freer, J. 2012 Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrol. Process.* **26** (26), 4078–4111.

Mohanty, B. P., Horton, R. & Ankeny, M. D. 1996 Infiltration and macroporosity under a row crop agricultural field in a glacial till soil. *Soil Sci.* **161**, 205–213.

Monteith, J. J. 1986 Howard Latimer Penman, 10 April 1909–13 October 1984. *Biogr. Mems. Fell. R. Soc. Lond.* **32**, 378–404.

Nielsen, D. R., Biggar, J. W. & Erh, K. T. 1973 Spatial variability of field-measured soil-water properties. *Hilgardia* **42**, 215–259.

Nimmo, J. R. 2012 Preferential flow occurs in unsaturated conditions. *Hydrol. Process.* **26**, 786–789.

Nyberg, L. 1995 Soil and Groundwater Distribution, Flowpaths, and Transit Times in a Small Till Catchment. Unpublished PhD Thesis, Department of Hydrology, Uppsala University, Sweden.

Refsgaard, J. C., Storm, B. & Clausen, T. 2010 Système Hydrologique Europeén (SHE): Review and perspectives after 30 years development in distributed physically based hydrological modelling. *Hydrol. Res.* **45**, 355–377.

Reggiani, P. & Rientjes, T. H. M. 2005 Flux parameterization in the representative elementary watershed approach: Application to a natural basin. *Water Resour. Res.* **41**, W04013.

Reggiani, P., Sivapalan, M. & Hassanizadeh, S. M. 2000 Conservation equations governing hillslope responses: Exploring the physical basis of water balance. *Water Resour. Res.* **36** (7), 1845–1863.

Reggiani, P., Sivapalan, M., Hassanizadeh, S. M. & Gray, W. G. 2001 Coupled equations for mass and momentum balance in a stream network: theoretical derivation and computational experiments. *Proc. Roy. Soc. Lond.* **A457**, 157.

Richards, L. A. 1931 Capillary conduction of liquids through porous mediums. *Physics* **1**, 318–333.

Rodriguez-Iturbe, I. & Porporato, A. 2005 *Ecohydrology of Water-Controlled Ecosystems: Soil Moisture and Plant Dynamics.* Cambridge University Press, Cambridge.

Schymanski, S. J., Roderick, M. L., Sivapalan, M., Hutley, L. B. & Beringer, J. 2007 A canopy scale test of the optimal water use hypothesis. *Plant, Cell Environ.* **31**, 97–111.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J. & Hutley, L. B. 2008 An optimality-based model of the coupled soil moisture and root dynamics. *Hydrol. Earth Syst. Sci.* **12**, 913–932.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Hutley, L. B. & Beringer, J. 2009 An optimality based model of the dynamic feedbacks between natural vegetation and the water balance. *Water Resour. Res.* **45**, W01412.

Shaw, E. M., Beven, K. J., Chappell, N. A. & Lamb, R. 2010 *Hydrology in Practice*, 4th edn. Spon, London.

Sklash, M. G., Beven, K. J., Gilman, K. & Darling, W. G. 1996 Isotope Studies of pipeflow at Plynlimon, Wales, UK. *Hydrol. Process.* **10**, 921–944.

Tetzlaff, D., Seibert, J., McGuire, K. J., Laudon, H., Burns, D. A., Dunn, S. M. & Soulsby, C. 2009 How does landscape structure influence catchment transit time across different geomorphic provinces? *Hydrol. Process.* **23** (6), 945–953.

Tetzlaff, D., Soulsby, C., Hrachowitz, M. & Speed, M. 2011 Relative influence of upland and lowland headwaters on the isotope hydrology and transit times of larger catchments. *J. Hydrol.* **400** (3), 438–447.

Vidon, P. 2012 Towards a better understanding of riparian zone water table response to precipitation: surface water infiltration, hillslope contribution or pressure wave processes? *Hydrol. Process.* **26** (21), 3207–3215.

Weiler, M. & Flühler, H. 2004 Inferring flow types from dye patterns in macroporous soils. *Geoderma* **120**, 137–153.

Wendt, R. C., Alberts, E. E. & Hjelmfelt, A. T. 1986 Variability of runoff and soil loss from fallow experimental plots. *Soil Sci. Soc. Amer. J.* **50**, 73–76.

Westerberg, I., Guerrero, J.-L., Seibert, J., Beven, K. J. & Halldin, S. 2011 Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras. *Hydrol Process.* **25**, 603–613.

Wood, E. F., Sivapalan, M. & Beven, K. J. 1990 Similarity and scale in catchment storm response. *Rev. Geophys.* **28**, 1–18.

Young, P. C. 2000 Data-based mechanistic modelling and validation of rainfall-flow models. In: *Validation of Hydrological Models* (M.G. Anderson & P.D. Bates, eds). John Wiley & Sons, Chichester, UK.

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

First received 19 August 2012; accepted in revised form 16 October 2013. Available online 11 February 2014