HYDROLOGICAL PROCESSES
Hydrol. Process. (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/hyp.6734

Multi-Method Global Sensitivity Analysis (MMGSA) for modelling floodplain *hydrological processes*

H. L. Cloke,[1]* F. Pappenberger[2,3] and J.-P. Renaud[4]

[1] *Department of Geography, King's College London, Strand, London WC2R 2LS, UK*
[2] *European Centre for Medium Range Weather Forecasts, Reading RG2 9AX, UK*
[3] *Department of Environmental Science, Lancaster University, Lancaster LA1 4YQ, United Kingdom*
[4] *School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK*

Abstract:

When studying hydrological processes with a numerical model, global sensitivity analysis (GSA) is essential if one is to understand the impact of model parameters and model formulation on results. However, different definitions of sensitivity can lead to a difference in the ranking of *importance* of the different model factors. Here we combine a fuzzy performance function with different methods of calculating global sensitivity to perform a multi-method global sensitivity analysis (MMGSA). We use an application of a finite element subsurface flow model (ESTEL-2D) on a flood inundation event on a floodplain of the River Severn to illustrate this new methodology. We demonstrate the utility of the method for model understanding and show how the prediction of state variables, such as Darcian velocity vectors, can be affected by such a MMGSA. This paper is a first attempt to use GSA with a numerically intensive hydrological model. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS global sensitivity analysis; sensitivity measures; floodplain hydrology; finite element; ESTEL-2D; hydrological processes; fuzzy performance; Darcian velocity

Received 2 August 2006; Accepted 26 February 2007

INTRODUCTION

Global sensitivity analysis (GSA) is emerging as a vital tool in environmental modelling (Hall *et al*., 2005; Pappenberger *et al*., 2006b; van Griensven *et al*., 2006). This is partially because of easier access to high performance computing resources but also because of an increasing appreciation of the complexities of the parameterization and assessment of environmental models. GSA provides a tool to rigorously map the (hyper) space of the possible model predictions, decompose the total uncertainty due to the various model factors (used for configuration of the simulation, such as numerical parameters, boundary and initial conditions, soil algorithm selection, etc.) and, most importantly, understand the influence of the different sensitivities of each of the model factors (Saltelli *et al*., 2004; Pappenberger *et al*., 2006a). Sensitivity analysis can be very useful for model supported decision making and to inform future research and field campaigns.

Cloke *et al*. (2003) note that it remains common practice to define boundary conditions and other configuration decisions from 'previous experience' of other modellers, on the basis of numerical efficiency or merely because it is easy to do so ('the model is already set up like that'). This is not satisfactory when important dependencies may exist between these choices and state variable predictions. For example, Cloke *et al*. (2003) note that an overestimation of Darcian velocity can severely undermine our ability to identify hydrological processes. Sensitivity analysis techniques can be essential in the understanding of the dependencies between parameter/factor/structural configuration of the model and state variable predictions.

We define GSA as being distinct from local sensitivity analysis (LSA). GSA techniques analyse the global impact of factors on model output, i.e. over the whole factor hyperspace at once, whereas LSA concentrates on the local impact of factors on the model output, which is not a robust approach in many cases. For example, because of equifinality, several areas of the parameter space might be equally representative of the natural system (Beven, 2006), and this will not be identified by LSA. GSA is therefore the preferred sensitivity analysis method. However, the issue of computational resources remains a major barrier to many researchers hoping to implement GSA.

In this paper, we combine multi-method GSA with an application of the ESTEL-2D finite element Richards equation model code. We take on the challenge of implementing GSA within a complex modelling framework, with (typically) limited computational resources (i.e. restriction on feasible number of simulations), whilst adequately covering the parameter/factor (hyper) space. We have selected a well-studied flood event on a lowland floodplain of the River Severn, UK, to illustrate the implementation of GSA for hydrological process models. Our objectives are twofold: (i) to demonstrate multi-method GSA for complex distributed subsurface

* Correspondence to: H. L. Cloke, Department of Geography, King's College London, Strand, London WC2R 2LS, UK.
E-mail: hannah.cloke@kcl.ac.uk

models; and (ii) to demonstrate how information on the sensitivity or importance of model factors can inform the investigation of hydrological processes and the prediction of state variables, such as Darcian velocity vectors.

In this article we first describe our GSA methodology, followed by a description of the ESTEL-2D model platform and the example flood event. We then discuss the results of this GSA and the implications of the method for this type of modelling.

MULTI-METHOD GLOBAL SENSITIVITY ANALYSIS (MMGSA)

Several reviews of sensitivity analysis methods have been conducted (e.g. Frey and Patil, 2002; Coyle *et al*., 2003; Helton and Davis, 2003). In hydrological science, GSA investigations of note include those in flood inundation modelling (Hall *et al*., 2005; Pappenberger *et al*., 2006a and 2006b) and rainfall runoff modelling (van Griensven *et al*., 2006). A general methodology has been introduced by Ratto *et al*. (2001), who linked global sensitivity analysis with another technique used in hydrology: the generalized likelihood uncertainty estimation (GLUE) framework. This paper attempts a GSA investigation for the first time on a highly complex, numerically intensive, subsurface model.

Most previous work in this area has been embedded into only one methodology to compute sensitivities, despite the fact that different sensitivity analysis methods can lead to a difference in the ranking of the *importance* of the different model factors. Mokhtari and Frey (2005) have advocated using a decision tree to choose a sensitivity analysis in order to counter this difficulty. However, using a decision tree does not explicitly acknowledge the potential difference in results from using different methods, and thus assumes that the meaning of 'sensitivity' can be defined unambiguously *a priori*. This is not necessarily the situation for many cases of practical importance (e.g. entropy and variance are both equally valid, but different, measures of sensitivity). Instead, we suggest that several different sensitivity measures have to be used in tandem. Here we use three different ways of evaluating sensitivity based on the same sample, termed multi-method global sensitivity analysis (MMGSA). The three different methods are the Sobol (1993), K–L Entropy (Kullback and Leibler, 1951; Liu *et al*., 2004) and Morris (1991) methods. These were selected as a subset of all those available, and represent measures founded on different theoretical bases; see the following section on Description of Sensitivity Analysis Methods for more details.

Within this framework, sensitivity can be quantified as first order, higher order and total sensitivity. First order sensitivity (or main effect) is defined as the sensitivity that a factor has on its own. Higher order sensitivity is due to the interaction of several factors. Total sensitivity is defined by the influence a factor has due to its first order effect together with all the interactions. It is questionable whether higher order effects can be computed due to correlation of the factors and the information content available in the evaluation data. The first order (or main) effect is computed by fixing the factor in question and varying all the others. The total effect is computed by fixing all factors other than the one in question. Therefore although both first order and total effects can be calculated, one cannot necessarily disaggregate these effects into any higher order effects of factors. Therefore, this analysis concentrates on first order and total effects.

In this paper we have used a Replicated Latin Hypercube (RLH) sample, which is generated by replicating a base sample set r times (see Saltelli *et al*., 2004, for more details). The base sample set is generated using the Quasi-Random Lp Tau method (Sobol, 1979). Sensitivity methods have inherent uncertainty due to the limitation of the sampling size or the methodology itself. Therefore, Saltelli and Sobol (1995) developed a methodology to compute uncertainties for the Sobol method. However, in this paper we use a bootstrap approach that allows us to make the uncertainty bounds of different methodologies more comparable. We randomly resample the total sample 1000 times and compute the individual sensitivity measures on each sub-sample. This allows us to obtain a measure of the uncertainty in our computation of the sensitivity measures.

Models are approximations of real systems and if sensitivity analysis is performed only on the model output then there is an implicit assumption that all the approximations made are correct (or negligible) and that the methodology and model are a valid and reasonable representation of the real physical system (e.g. Pappenberger and Beven, 2006). However, in the face of model uncertainty, such an assumption has to be questioned and thus sensitivity analysis should be embedded in measured data. Moreover, we should question whether it is sensible to compute sensitivity on all model results. Some of the set of factors may lead to model results which are non-behavioural (meaning they cannot be accepted as representation of the system in question). The sensitivity of those simulations can be considered irrelevant for the understanding of the system and the model, since what it represents is unrealistic. A model evaluation measure has to account for total error; this includes the input error, the output error and the model inadequacy (Beven, 2006).

In this paper we adopt two different approaches to quantify this error. The first approach uses the sum of squared errors (traditional approach) and the second uses a fuzzy type approximation. We believe that the latter is a more appropriate representation of all error sources and have only used the more traditional approach for comparative purposes.

Fuzzy type approximation

Each model run in the analysis will predict a time series of a particular variable (in our case, hydraulic head). In a fuzzy type approximation, a 'membership value'

DOI: 10.1002/hyp

is assigned to each model prediction at each timestep. This membership value has to be between 0 and 1, and is assigned based on the trapezoidal function shown in Figure 1, which takes account of two error thresholds a and b. The width z of the trapezoid is an 'error' of a m from the observed value for the inner length and b m for the total length. Any error greater than b will result in a membership value of 0. An error less than a will give a value of 1, and an error between a and b will be a value between 0 and 1. The individual membership values for each timestep are multiplied and a value of 0 defines the model as non-behavioural. For a similar approach, see Pappenberger *et al*. (2005) and Pappenberger and Beven (2004). The widths used in this fuzzy approximation will undoubtedly influence results (the more fuzzy, i.e the greater the widths used, the more behavioural models there will be). However, they have been set conceptually based on known measurement errors and the commensurability error (Beven, 2006).

Although technically our multi-method approach could be used without the fuzzy approximation, we feel that it would not make conceptual sense to use an unrestricted set of models for our multi-method GSA (i.e. not to discriminate by behaviourality). We believe that the fuzzy evaluation criteria are the optimal evaluation criteria for this model as it truly reflects the errors in measurements and commensurability. The implications of using informal likelihood measures and retaining only behavioural simulations is discussed in length by Montanari (2005) and Beven *et al*. (2007).

DESCRIPTION OF SENSITIVITY ANALYSIS METHODS

Sobol. If the uncertainty of input factors can be approximated by independent probability distributions, then sensitivity indices can be related to the decomposition of the total unconditional variance (Sobol, 1993). The decomposition can be shown in an ANOVA-like way as follows:

$$V = \sum_i V_i + \sum_{i<j} V_{ij} + \ldots + V_{1,2,3,\ldots,n} \qquad (1)$$

where V is the total variance, V_i is the variance contribution due to effects of the random variable X_i (e.g. model factor), and V_{ij} is the covariance. Higher order terms show the variance contribution between two or more random variables. Each partial variance is normalized with respect to the total unconditional variance and allows sensitivity indices to be obtained:

$$S_i = \frac{V_i}{V}, 1 \le i \le n$$

$$S_{ij} = \frac{V_{ij}}{V}, 1 \le i < n$$

$$S_{i,i+1,\ldots n} = \frac{V_{i,i+1,\ldots n}}{V} \qquad (2)$$

where S_i is the main effect and S_{ij} is the interaction effect between i and j. It is further possible to define a total effect of i (Total Sensitivity) which incorporates all terms of any order, including variations in X_i.

Further information on this method can be found in Sobol (1993) and Ratto *et al*. (2001).

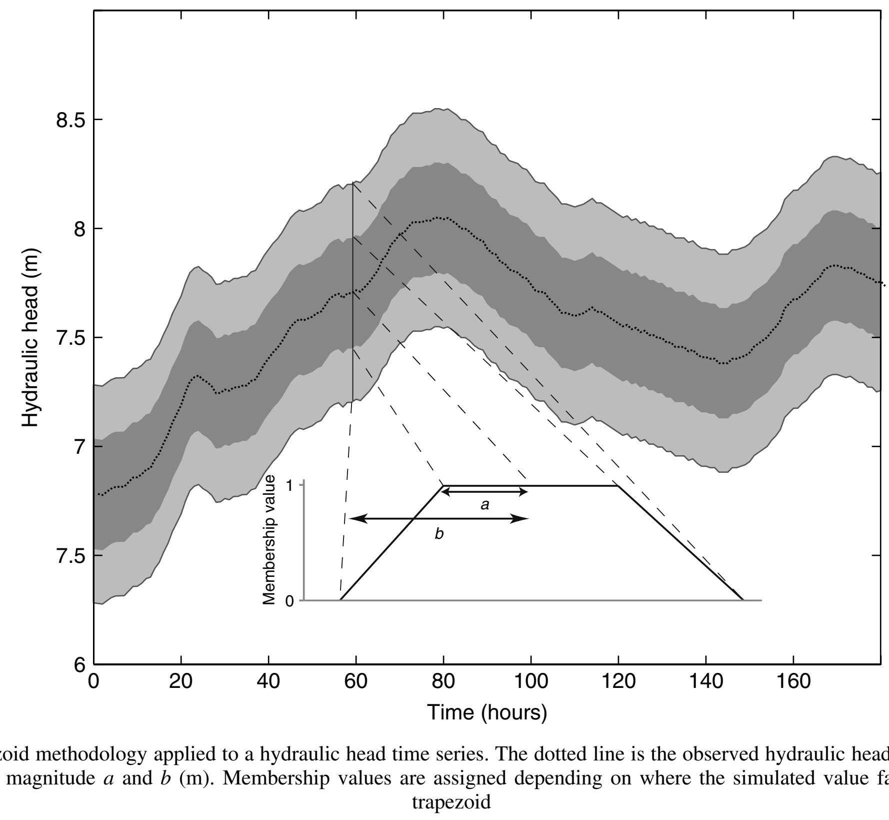

Figure 1. Fuzzy trapezoid methodology applied to a hydraulic head time series. The dotted line is the observed hydraulic head time series, surrounded by two error zones of magnitude a and b (m). Membership values are assigned depending on where the simulated value falls along the top of the trapezoid

DOI: 10.1002/hyp

Kullback–Leibler Entropy. It has been argued that variance-based methods rely too heavily on the assumption that the second moment is sufficient to describe the uncertainties and sensitivities encountered. These assumptions may be invalid if the distribution is highly skewed due to nonlinear functions or inputs (Liu *et al.*, 2004). Krykacz–Hausmann (2001) argued that sensitivity should be measured as the deviation of the factor distribution from being uniform (if the factor has been initially sampled from a uniform distribution). Based on entropy, Liu *et al.* (2004) presented an alternative approach, which is using the Kullback–Leibler Entropy (Kullback and Leibler, 1951):

$$D_{\mathrm{KL}}(pdf_1|pdf_0) = \int_{-\infty}^{\infty} pdf_1(y)\log\frac{pdf_1(y)}{pdf_0(y)}\mathrm{d}y \tag{3}$$

where D_{KL} is the Kullback-Leibler Entropy, y is the model response, pdf_0 is the probability density function of the random response $y = h(\mathbf{X})$ where h is the model and $\mathbf{X}$ is the model factor vector, and pdf_1 is the probability density function of the random response $y = h(\mathbf{X}_1)$ due to changes in $\mathbf{X}_1$.

Further information, including the derivation of main and total effects similar to the Sobol methodology, can be found in Liu *et al.* (2004).

Morris. We also apply the screening method by Morris (1991) in this paper. In this methodology, the impact of changing one factor ($\mathbf{X}_1, \ldots, \mathbf{X}_n$) at a time is evaluated in turn. This impact (so called elementary effect of $\mathbf{X}_1, \ldots, \mathbf{X}_n$) is expressed as the gradient of the model response between the original set of factors and the new set of factors. The standard deviation σ_M of all elementary effects and the mean of the absolute of multiple elementary effects $\mu^*{}_M$ are measures of factor sensitivity (Campolongo *et al.*, 1999). A high value of σ_M means that the elementary effect compared to this factor is significantly different. In contrast, a low σ_M indicates very similar values of the elementary effect implying that the effect of the factor is almost independent of the values taken by the other factors. It can be shown that $\mu^*{}_M$ is in fact similar to total sensitivity and, therefore, will be used in this study (Saltelli *et al.*, 2004).

THE ESTEL-2D MODEL PLATFORM

The model platform used here is ESTEL-2D, the subsurface flow component of the TELEMAC modelling system (Hervouet and Bates, 2000) which solves the Richards equation in saturated and unsaturated porous media with the finite element technique. The model has recently been used to study high resolution floodplain hydrological processes (Bates *et al.*, 2000; Claxton *et al.*, 2003; Cloke *et al.*, 2003, 2006a, 2006b). The flow model is described in detail in Renaud *et al.* (2003) and Cloke *et al.* (2006b); only a summary of the model code is given below.

ESTEL-2D solves both the h-based (pressure based) and the mixed (pressure and moisture content based) form of the Richards equation. The mixed form is used preferentially because of its excellent mass conservation properties (Celia *et al.*, 1990). However, the h-based form is generally used as a substitute to the mixed form to counter numerical convergence difficulties (Claxton, 2002; Cloke, 2003). This will not affect results, as long as mass balances are at an acceptable level (approximately less than 1% of the volume of water in the domain) (Claxton, 2002), as the hydrology itself does not vary. The equation system is solved in time and space. The time discretization for the Richards' equation is defined using the modified Picard iterative scheme. This is based on a robust discretization of the mixed form of the Richards equation (Celia *et al.*, 1990). The convergence criterion used is based on Huang *et al.* (1996) which is particularly well suited to the modified Picard scheme. The finite element spatial discretization in ESTEL-2D uses the Galerkin variational formulation.

The governing equations are solved to give the pressure head, from which values of hydraulic head, Darcian velocity and moisture content can subsequently be derived using additional relationships. Various soil moisture algorithms are available including the Brooks–Corey algorithm (Brooks and Corey, 1964) and the van Genuchten algorithm (van Genuchten, 1980), used here together with the Burdine hydraulic conductivity relation (Burdine, 1953). For particular soil hydraulic parameter sets, simulations may suffer from numerical difficulties including long runtimes and slow convergence (require adaptation to a small timestep for solution and many iterations). These parameter sets are notably those associated with a rapid change in moisture content with pressure head, steep soil moisture release curve in general (Cloke, 2003) or those specifically with a van Genuchten n parameter close to 1 (Vogel *et al.*, 2001). For such simulations, the Brooks–Corey algorithm is preferred as it has shallower soil moisture release curves, produces more numerically stable simulations and has faster runtimes (Cloke, 2003). It should be noted that the unstable numerical behaviour is due to the non-linearity in the equation used for the soil hydraulic representation (i.e. the soil moisture release curve itself) and is not specific to a poor numerical solution in ESTEL-2D. Unstable numerical behaviour would also be found in other subsurface numerical models of this type.

In ESTEL-2D it is possible to use an iterative seepage face routine (Claxton, 2002; Renaud *et al.*, 2003) for atmospheric boundary conditions where the saturated zone (actual or perched water table) intersects the boundary at atmospheric pressure, modified from a scheme proposed by Rulon *et al.* (1985). When a seepage face is specified, the whole of the atmospheric boundary (usually the top boundary) becomes a 'system-dependent' boundary condition that is not known *a priori*. Atmospheric pressure is maintained for all nodes along the seepage face and the nodes are treated as Dirichlet nodes with the prescribed pressure: $h = 0$. Unsaturated nodes above the seepage face are specified with a flux (Neumann),

DOI: 10.1002/hyp

which is either zero or positive when rainfall (infiltration) is to be imposed. Multiple seepage faces are possible and a dynamic interaction between the seepage face and the river channel can be simulated by the seepage face scheme, as river stage changes over the course of a simulation.

For each case of interest, a new finite element mesh is generated and the FORTRAN source code must be set up in order to specify the correct time-dependent boundary conditions. Initially, a steady state case is run to generate suitable initial conditions (which match observed spatial distributions of pressure head at $t = 0$). The steady state conditions are tested to determine whether they represent a true steady state field (Cloke *et al*., 2003).

EXAMPLE FLOOD EVENT AND MODEL SETUP

The Leighton hydrology research site is located on the left bank floodplain of the River Severn, near the village of Leighton in Shropshire, UK, at 52°38′N, 02°34′W (Figure 2). Detailed information on the site can be found in Stewart *et al*. (1999), Bates *et al*. (2000), Burt *et al*. (2002) Claxton (2002) and Claxton *et al*. (2003), from which the following is a summary. The river flows through the area as a gravel bed channel that is ~45 m wide. River stage and a network of floodplain subsurface pressures have been continuously monitored since 1997 (with breaks during the foot and mouth crisis) in the area shown on Figure 2 in light grey and schematically in Figure 3. At the location of the piezometer instrumentation, ~40 m above sea level, the floodplain is ~1 km wide in total with the river running in a relatively straight channel ~120 m from the terrace that forms the base of the left hillslope. Alluvial sediments at this location are mainly sandy clay loam soils over gravels, and the soil is relatively homogeneous (because of the straight river section), although some evidence of relic channel structure exists in places. The vegetation at the site is well-grazed grassland pasture. Mean annual rainfall is of the order 1000 mm and the river has inundated the floodplain on average 3 times a year since 1997. Typical bankfull discharge at Leighton is ~180 $m^3 s^{-1}$, with the largest flood peaks exceeding 450 $m^3 s^{-1}$.

The river at the site is a very important control on the riparian hydrological processes, especially during overbank events, and the hyporheic zone is thought to be very large (Bates *et al*., 2000). The hillslope waters also contribute to flow processes and the floodplain itself acts as an important store and conduit of water and of chemicals (Claxton *et al*., 2003). Several flood events at

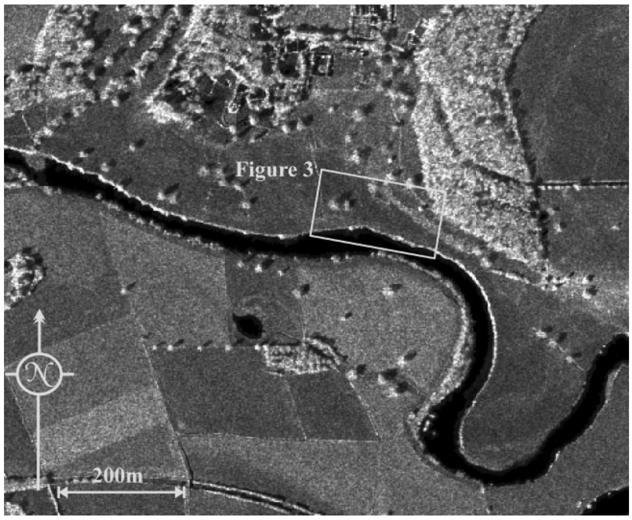

Figure 2. Ortho-rectified radar image of the Leighton hydrology research site on the River Severn at 52°38′N, 02°34′W (courtesy of NERC). The location of the floodplain hydrology sensor network (Figure 3) is outlined in light grey

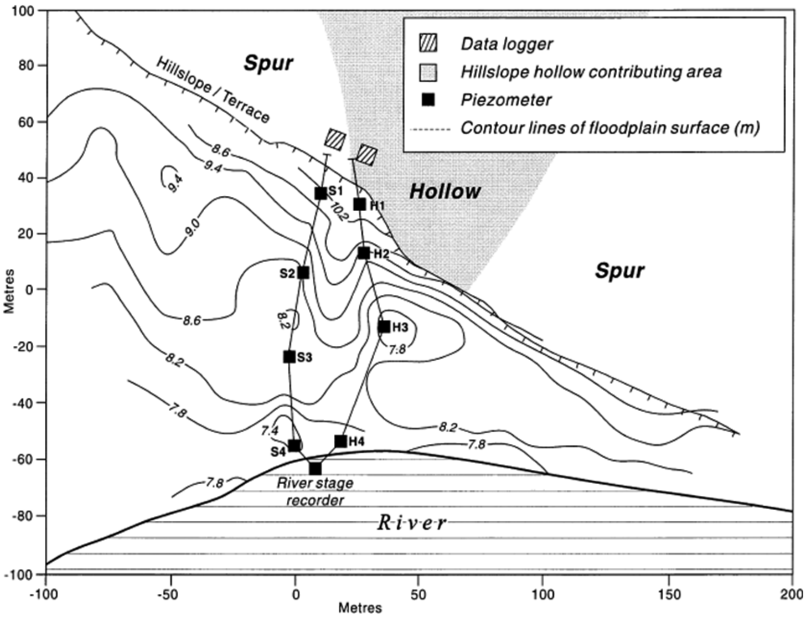

Figure 3. Map of the field site, showing floodplain topography and the location of the original (pre-2005) transects of piezometers (after Burt *et al*., 2002; Claxton *et al*., 2003)

DOI: 10.1002/hyp

the site have been studied in detail (Burt *et al*., 2002). Bates *et al*. (2000) and Claxton *et al*. (2003) have applied ESTEL-2D to vertically aligned cross-sections through the floodplain (Figure 3). During out-of-bank and other high discharge conditions, a reverse groundwater ridge has been found to develop in the floodplain subsurface and results in strong groundwater flux velocities directed towards the base of hillslopes adjoining the floodplain (Bates *et al*., 2000; Burt *et al*., 2002). This prevents hillslope contributions to the floodplain if the flood stage is high, but antecedent conditions, local rainfall and runoff and flood stage all act to complicate this basic pattern (Burt *et al*., 2002), especially for high discharge in-bank events. The case is therefore both useful and well defined in order to demonstrate our MMGSA methodology.

We use ESTEL-2D to simulate a flood event from January 1998 (denoted as event 'D' in Burt *et al*., 2002). The river stage and rainfall characteristics of this event are shown in Figure 4; total rainfall for the event is 96·4 mm over 1101 hours. To represent a vertical cross-section of transect S, we used the finite element mesh and boundary conditions described in Claxton *et al*., (2003). The orientation of the mesh was chosen to be parallel to the dominant groundwater flowpath during flood events. Claxton (2002) and Cloke (2003) report the results of a one-factor-at-a-time sensitivity analysis for some of the soil hydraulic parameters for this case. They found that hydraulic conductivity K_S was by far the most sensitive parameter of all the hydrological variables tested (including the hydraulic head at the piezometers). The head was also sensitive to the variation in air entry pressure H_S. Other parameters were relatively insensitive.

The possible factor values, ranges and distributions for this simulation of ESTEL-2D are shown in Table I. Ranges and distributions of the soil parameter values were originally defined based on data from the original field investigation, including the falling head well test and soil laboratory testing (Stewart, 1998), together with ranges defined in Cloke (2003) modified from Meyer *et al*. (1997). However, initial simulations were run with a Morris screening methodology and a sparse sample to determine the suitability of the values selected, and were changed accordingly. The likely errors associated with the boundary condition specification are considered, as are two different soil moisture algorithms (Cloke, 2003). In previous studies with ESTEL-2D, parameterization for this flood simulation was determined on the basis of trial and error calibration, informed by one-factor-at-a-time sensitivity analysis (Claxton, 2002).

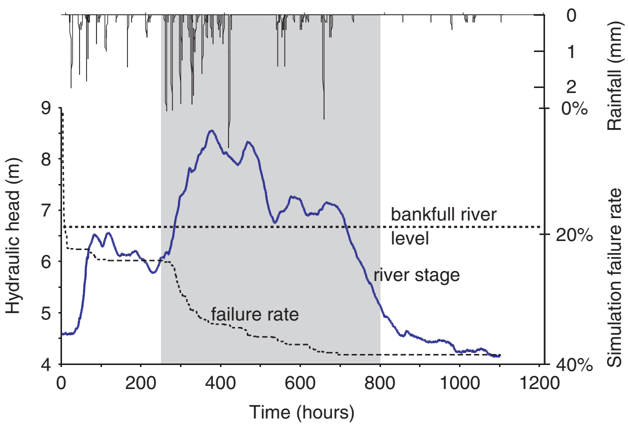

Figure 4. River stage (left ordinate) and rainfall dynamics (top graph) for the January 1998 flood event at Leighton. Level of bankfull for transect S is also depicted. The failure rate of simulations for the BC*h*-based simulation set is plotted as a dashed line (right ordinate). The grey box represents the overbank part of the event, and analysis is focused on the results in this period

In this paper we have sampled the 9-dimensional factor space specified in Table I; sampling was carried out using the RLH method described above. In order to compute the main and total order effects of each factor, two independent samples are required. To compute the main (total) effect, factor 1 is taken from sample 1 and the remainder is taken from sample 2 (and vice versa). The base sample is therefore $9 \times 2 + 2 = 20$ and the

Table I. Specified ranges and distributions of factors. Factors 4 and 5 are exchangeable between the two soil moisture algorithms (Brooks–Corey and van Genuchten)

Factor	Description	Symbol	Unit	Distribution	Mean	Min (0·001 quantile)	Max (0·999 quantile)
1	Saturated moisture content	θ_S	—	Normal ($\sigma = 0{\cdot}09$)	0·41	0·132	0·688
2	Residual moisture content	θ_R	—	Normal ($\sigma = 0{\cdot}01$)	0·0954	0·065	0·125
3	Saturated hydraulic conductivity	K_S	$\mathrm{ms^{-1}}$	Log normal (A = −14·82, B = 1·24)	$9{\cdot}93 \times 10^{-7}$	$1{\cdot}51 \times 10^{-10}$	$1{\cdot}01 \times 10^{-4}$
4a	Brooks–Corey, pore size distribution index	λ	—	Normal ($\sigma = 0{\cdot}1$)	0·318	0·017	0·619
5a	Brooks–Corey, air entry pressure	h_S	m	Log normal (A = −0·382, B = 0·710)	0·880	0·074	6·275
4b	van Genuchten alpha	α	$\mathrm{m^{-1}}$	Log normal (A = −4·22, B = 0·719)	1·9	0·16	13·56
5b	van Genuchten, *n*	n	—	Normal ($\sigma = 0{\cdot}1$)	1·32	1·02	1·62
6	Storage parameter	S	—	Uniform	$0{\cdot}1 \times 10^{-3}$	$0{\cdot}1 \times 10^{-4}$	$0{\cdot}1 \times 10^{-2}$
7	Upslope pressure	UP	m	Uniform	Measured value	−0·5	0·5
8	River stage	R_S	m	Uniform	Measured value	−0·5	0·5
9	Rainfall (precipitation)	PPT	m	Uniform	Measured value	90%	100%

DOI: 10.1002/hyp

total sample has to be a multiple thereof. In this case, we used 1280 samples as at this value the cumulative distribution function of the model responses stabilized. ESTEL-2D setup in the way described requires significant computational resources for these 1280 samples in order to maintain conservation of mass balance, while using an adaptive time stepping strategy and iterative seepage face routine. Individual simulation times varied from 130 min to 52 hours (simulations run on Dell Precision 670 workstations with dual Intel Xeon 2·4 GHz Processors and with a Suse linux 10·0 operating system). The long simulation times reflect the difficulty (and the interest) of the selected test case.

INSTABILITY DIFFICULTIES AND SIMULATION NON-CONVERGENCE

When using a complex model such as ESTEL-2D, there may be some factor sets for which simulations will not converge. This may be different for different soil moisture algorithms (e.g. van Genuchten versus Brooks–Corey) and solution schemes (e.g. mixed versus *h*-based solution of the Richards equation). We reiterate that instability and convergence difficulties are due to non-linearity in the equations of subsurface flow and are not specific to a poor numerical solution in ESTEL-2D. Non-convergence or instability can occur in the first time step for particularly difficult factor sets (for example, parameters combining to produce a very steep soil moisture release curve) or during rapid changes in the internal hydraulics or boundary conditions (such as an intense period of rainfall or flood inundation). It should be noted that although it is possible to mitigate some of the effects of instability and non-convergence by adjusting the numerical parameter options in ESTEL-2D (such as the implication coefficient for the pressure head), the effects would not entirely disappear, and so here we keep numerical parameter options constant throughout to illustrate our case.

Table II shows the different solution schemes and soil moisture algorithms for the number of simulations that reached the end of the time series without numerical difficulties. This is also shown as a failure percentage of all simulations attempted. It can be seen that the van Genuchten–*h*-based (VG*h*) combination has the lowest failure percentage, with the Brooks–Corey-mixed (BCm) performing the worst, with 100% failure. For each simulation set there was no obvious pattern in the factor space to explain this failure, as sampled in Table I. Of course, a sudden change in the magnitude of the boundary conditions in time (including factors precipitation PPT and upslope pressure UP) can cause convergence failure. It should be noted that just because a simulation reaches the end of the time series does not mean that the simulation is behavioural (simulated values are close to observed values).

The simulation failure rate for the BC*h* simulation set is shown on Figure 4. The pattern of failure was similar for the VGm and VG*h* simulation sets. Around 20% of simulations fail in the first few time steps, probably due to a combination of factors resulting in highly non-linear equations to solve. The failure rate is high as the river stage moves beyond bankfull and the domain floods; this is probably because of the rapidly changing top boundary condition and especially the part of the boundary where the dynamic seepage face is acting. It is the experience of the authors that this type of failure is to be expected for dynamic flood simulations of this type.

The question of what this means for our global sensitivity analysis arises. The convergence failure creates 'holes' in the response surface. We do not know whether the surface in these holes would be sufficiently smooth from one edge of the gap to the other, so that surface interpolators such as Bayesian interpolators (O'Hagan, 2006) can be applied. Moreover, we do not know whether good representations of the system lie within this unknown space. This problem is distinct from undersampling a factor space as we cannot re-sample that space; although different numerical schemes should converge to the same solution for the same input this is not guaranteed. This means that our sensitivity analysis can never be truly global and one must be careful of making assumptions about the unknown space. However, all of the measures that we use assume that the overall result would not change if the factor sets which failed would have been successful.

RESULTS OF THE GLOBAL SENSITIVITY ANALYSIS

The generalized results of behavioural simulations

Although each simulation ran for ~46 days (the entire time series), for simplicity the following results are displayed for the period 250–500 hours. This time period covers the overbank section of this flood event (period shaded in grey on Figure 4), and is the most hydrologically active and therefore important time period.

Table II. Number of simulations that reached the end of the time series without numerical difficulties and failure (%) of all simulations attempted

Numerical solution scheme	Soil moisture algorithm	Abbreviation	Number of simulations that reached end of time series (out of 1280 in total)	Failure (%) of simulations attempted (to 3SF)
mixed	van Genuchten	VGm	667	47·9
mixed	Brooks–Corey	BCm	0	100
h-based	van Genuchten	VG*h*	942	26·4
h-based	Brooks–Corey	BC*h*	611	52·3

DOI: 10.1002/hyp

Simulations with fuzzy membership values above 0 are classified as behavioural, as explained previously. Table III shows the generalized results of behavioural simulations for piezometers S2 and S3 for the BC*h* simulation set. For the other simulation sets (with the van Genuchten soil moisture algorithm) no behavioural simulations were found. In Table III it can be seen that with wider fuzzy thresholds ($a = 0{\cdot}25$ m and $b = 0{\cdot}50$ m) more behavioural simulations are evident. This is as would be expected (more fuzzy), and is in agreement with the findings of Claxton *et al*. (2003). It can also be seen that there are many more behavioural simulations for piezometer S3 than for piezometer S2, also reflected in higher Nash Sutcliffe values which we have included as a comparison to traditional goodness-of-fit methods. This is perhaps more surprising, as the observed hydraulic heads at these piezometers exhibit similar dynamics to each other. It suggests therefore that either only a limited number of factor sets can truly represent the flow behaviour at piezometer S2, or more likely that there is 3-dimensional flow evident in the floodplain at this point, even during overbank conditions when 2-dimensional flow has been assumed to dominate.

It could be argued that a model of this complexity should be able to reproduce all measurements, and perhaps should be discarded if it does not do so. Young *et al*. (2004) have argued that a model should be as simple as possible, fit the data and have a physical explanation. Although all of this is desirable, it would be difficult to use such a model for hypothesis testing of vector flowpaths, as is the application in this case. However, as ESTEL-2D cannot represent all measurements sufficiently we are forced to reject the model as a completely valid representation of the overall modelled system. Such a rejectionist approach allows us to redefine measurement campaigns based not only on sensitivity but also on model failure (e.g. installing more piezometers in the failing area as we want to know what is going on there).

For the remainder of this analysis we will concentrate on behavioural simulations calculated with the wider fuzzy thresholds for piezometer S3. Figure 5 shows scatter (dotty) plots of the factor values against membership values from the fuzzy performance function for the overbank time period. The high membership values (∼1) are distributed across the ranges for each factor

Table III. Number of behavioural simulations (calculated with fuzzy methodology) and their associated Nash–Sutcliffe values for piezometers S2 and S3

Piezometer	Number of simulations over Fuzzy Behavioural threshold values (m) of:		Number of simulations over Nash–Sutcliffe values of:			
	(a, b) = (0·10, 0·15)	(a, b) = (0·25, 0·50)	0·6	0·7	0·8	0·9
S2	0	9	96	35	13	3
S3	14	114	356	269	156	54

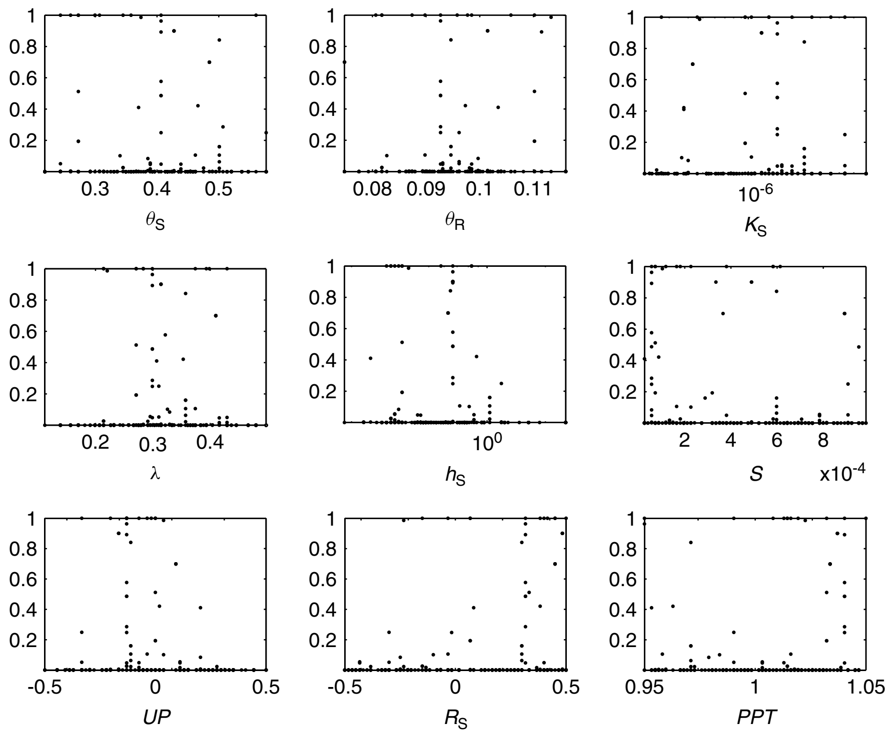

Figure 5. Scatter plots of parameter versus membership values for the overbank time period for piezometer S3 and the BC*h* simulation set

DOI: 10.1002/hyp

and there is no obvious pattern in factor values controlling the membership values. Although one would expect that pressure head against the hydraulic conductivity K_s would have some structure, in Figure 5 this structure is not evident because the nature of the fuzzy performance function means that a range of K_s and pressure head values can lead to behavioural simulations. Although the prior sampling distribution is not fully reflected in the abscissae, the tendency of the sensitivities can already be derived from these plots. For example, the air entry value H_s has a drop in performance at higher values. However, these are only initial conclusions and need to be confirmed through further analysis.

Figure 6 shows a comparison of the results of the Sobol analysis for total and first order sensitivities. These box and whisker plots are shown for our fuzzy performance function (restriction in that only behavioural simulations are analysed) but also for the sum of squared errors (SSE), which incorporates the entire response surface. We have included SSE for comparative purposes as it is often used for GSA (van Griensven *et al*., 2006). A comparison of the SSE and Sobol results immediately shows the effect of taking into account the whole response surface versus our behavioural fuzzy performance approach. We would like to reiterate that there are only a limited number of behavioural simulations for the case we have shown, and these results should ideally be confirmed by the use of a greater number of simulations (although this may be difficult to achieve given run time and stability restrictions).

For the total sensitivity, the fuzzy method shows the highest Sobol value (with the 100% sample value at $\sim$0·9) and therefore that of the highest importance to be the Upslope Pressure *UP* factor. The other factors are much less important, with the next important factor being the hydraulic conductivity K_S with a Sobol value of $\sim$0·35. The saturated moisture content θ_S, residual moisture content θ_R, Brooks–Corey pore size distribution index λ and the rainfall *PPT* are all somewhat important with Sobol values of $\sim$0·25. The other factors are unimportant. For the first order sensitivities, the fuzzy method shows only the Upslope Pressure *UP* to be of importance. However, if we look at the sum of squared errors, the results are very different. In the total sensitivity plot, the most important factor in this case is the Brooks–Corey air entry pressure H_s with an importance of $\sim$1, followed closely by the hydraulic conductivity K_s ($\sim$0·95). The Upslope Pressure *UP* is still important with a Sobol value of $\sim$0·75. Many of the other factors have Sobol values of over $\sim$0·4. The same differences are evident in the first order sensitivity plot, with *UP* definitely not as important as for the fuzzy method, and with some importance attached to many of the parameter. The highest first order importance is attached to the Brooks–Corey pore size distribution index λ parameter.

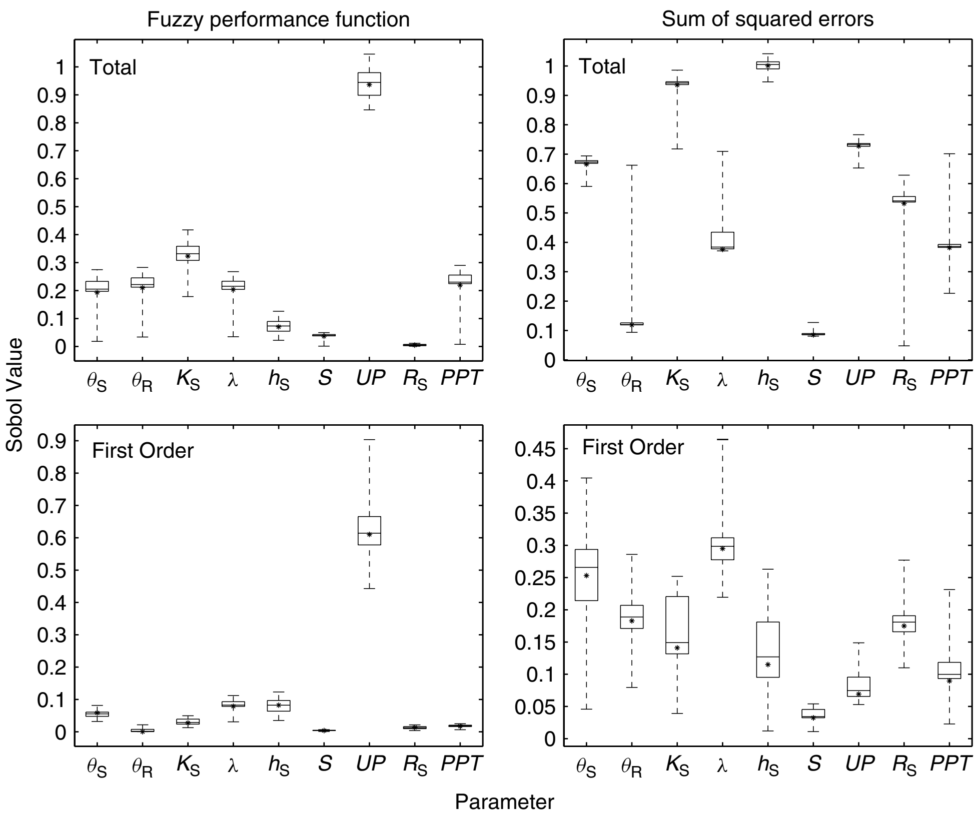

Figure 6. Box and whisker representations of the total and first order sensitivity results of the Sobol method for the BC*h* simulation set and piezometer S3. Fuzzy performance function results are shown on the left hand side and sum of squared errors results are on the right. Whiskers show the 5 and 95% quantiles, and the boxes show the 25 and 75% quantiles. The central line shows the 50% quantile and the star indicates the value calculated with 100% of the sample

DOI: 10.1002/hyp

We have seen that Claxton (2002) and Cloke (2003) found K_s to be the most sensitive parameter for this case, with H_s exhibiting some sensitivity. The sum of squared errors method (SSE) results agree with such previous findings. If we compare the hydraulic conductivity H_s results for the fuzzy method and the sum of squared errors method the major differences are illustrated. If we take into account the entire response SSE then hydraulic conductivity does indeed have a high total importance. However, restricting the response surface to just behavioural results in the hydraulic conductivity H_s dropping significantly in importance. This has significant implications for model decision making, and we advocate using some kind of behavioural restriction in order to truly determine which factors one might concentrate on for data collection etc. (in this case, more instrumentation at the upslope pressure boundary would be a useful idea).

Results of comparison of sensitivity analysis methods

Figure 7 shows that the *UP* remains an important parameter for all of the different sensitivity analysis methods, however, there are also other differences in results between the methods (all using the fuzzy performance function this time). For the Entropy method, the total sensitivity of the air entry pressure H_s is also an important factor. The first order sensitivity analysis for the Entropy method is a little more interesting. These results show that for this calculation method *UP* has little first order sensitivity, and therefore suggests that this factor is only important in conjunction with the other factors. This is the opposite to what was found with the Sobol method, where *UP* had a very high first order sensitivity. The entropy measure is more sensitive to changes in the point distribution than to variations in the response surface, and so such a difference in first order sensitivity is due to this

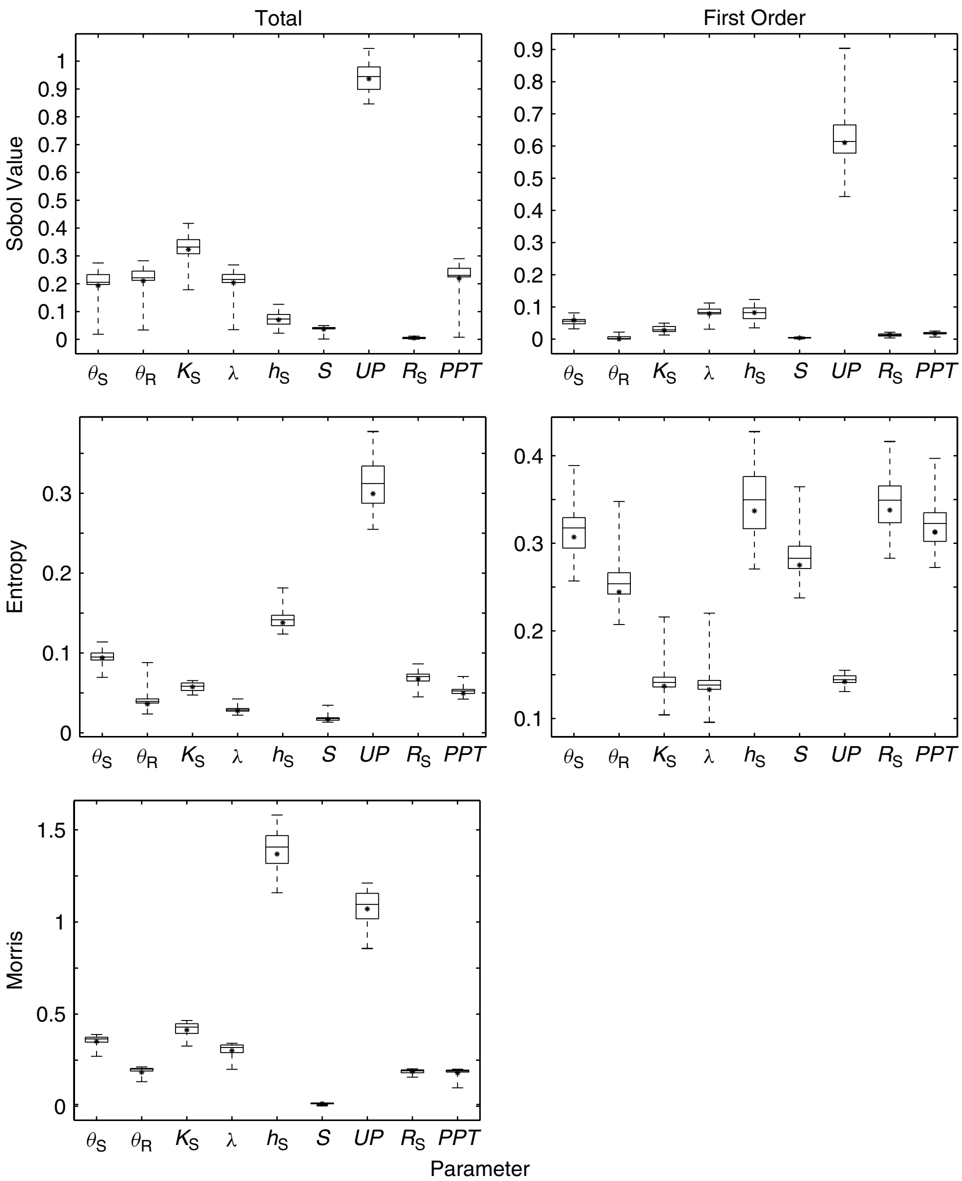

Figure 7. Box and whisker representations of the total and first order sensitivity results of the Sobol, Entropy and Morris methods for the BC*h* simulation set and piezometer S3. Whiskers show the 5 and 95% quantiles, and the boxes show the 25 and 75% quantiles. The central line shows the 50% quantile and the star indicates the value calculated with 100% of the sample

DOI: 10.1002/hyp

calculation difference. Neither measure is more correct than the other, so we are left with an ambiguous result although it is clear that *UP* is sensitive. For the Entropy method, several other factors have medium importance for the first order sensitivity results, namely H_s, R_s, *PPT*, θ_S, θ_R and *S* (storage). Although individually sensitive, there is some correlation between the factors which leads to an effective cancelling out of these sensitivities. For the Morris method, only the H_s and *UP* factors are important.

We therefore have a set of sensitivity analysis results that show some confirmation of sensitivities and some ambiguities. To aid in the interpretation of results, Table IV shows the overall rankings calculated from the different methods. Although only a relative ranking, this provides a guideline on where to begin interpretation, especially if such relative rankings are confirmed by an analysis of the sensitivity plots themselves. In this case, it can be confirmed that the air entry pressure parameter H_s is a sensitive parameter for this case. This has strong implications for the estimation of effective soil parameters. In the original analysis in Claxton (2002), the mean value of H_s for the measured soil type was selected from published literature tables. However, this is not acceptable practice when there are clear dependencies between H_s and the results. Any future simulations of this nature should (i) confirm the sensitivity of H_s, (ii) use a range of reasonable values within an uncertainty analysis framework, and (iii) evaluate the effects of spatial variability on the sensitivity of this parameter.

Table IV also confirms *UP* as a factor on which we should concentrate further constraining field measurements. Whether this is an interactive factor or not is not clear from the results, however, as it achieves high values for all three total sensitivity type measures it is safe to assume so. An interaction effect could be explained by our physical understanding of the system: the internal behaviour of the flow domain is necessarily controlled by the boundary conditions. However, this upslope pressure boundary does seem to have a large effect on the results of the model and therefore future simulations of this nature should focus on the definition of, and the data for, this boundary. Recommendations include a further nest of piezometers at this hillslope–floodplain boundary to further constrain the exact behaviour of the pressure and to confirm 3D flow in the area. These data could lead to a different specification of the boundary at this point (see Cloke *et al*., 2003).

The saturated moisture content θ_s and hydraulic conductivity K_s were also important overall. However, the importance of the K_s parameter has perhaps been overstated in previous sensitivity analyses of this type. This is because (as demonstrated by the SSE method), without a behavioural restriction, K_s is one of the most sensitive factors involved. However, with a behavioural restriction it drops in importance. This does not mean that there is no sensitivity exhibited and it still remains prudent to constrain the K_s parameter as much as possible. In addition, the insensitivity of the factor does not mean that the factor can be fixed in a prediction mode as it may be important in non-linear interactions with different boundary conditions and under different circumstances.

Temporal variation in factor importance

Figure 8 demonstrates how the factors vary in importance through time. For example, for the Sobol method, the Upslope Pressure *UP* and hydraulic conductivity K_s are important during the flood peak ($\sim$350 – 500 hrs). Other factors are important before then, in particular H_s and θ_s. Also of importance in all methods are the peaks in sensitivity; for example the peak at $\sim$460 hours coincides with a dip at the top of the hydrograph and follows an intense period of rainfall. For all methods, many factors in addition to *UP* and K_s become important. This highlights the fact that when different hydrological processes are operating, the model is differentially sensitive to different parameters. During times of flooding (which applies to most of this event), *UP* and K_s are important. However, during wetting and drying of the floodplain, other factors, become important, especially H_s. This is as would be expected, as the wetting and drying behaviour of the soil is controlled by the air entry value parameter. Once the soil is totally saturated, then other factors will necessarily become more important. This demonstrates that the overall sensitivity of different model factors will depend on the length of the time series and the different hydrological regimes involved.

Table IV. Ranking of factors (1 most sensitive/important to 9 least sensitive important) for the Sobol, Entropy, Morris and SSE methods for the BC*h* simulation set and piezometer S3. Rankings are calculated for the 100% sample set value (denoted by a star on Figures 6 and 7) and hence do not take into account uncertainties. The Overall ranking was calculated by multiplying the other rankings and rank transforming the results (the same ranked result was obtained by adding the other rankings)

Factor name	Factor symbol	Sobol total	Sobol 1st order	Entropy total	Entropy 1st order	Morris	SSE total	SSE 1st order	Overall
Saturated moisture content	θ_S	6	4	3	4	4	4	2	3
Residual moisture content	θ_R	4	8	7	6	6	8	3	8
Saturated hydraulic conductivity	K_S	2	5	5	8	3	2	5	4
Pore size distribution index	λ	5	3	8	9	5	6	1	5
Air entry pressure	h_S	7	2	2	1	1	1	6	1
Storage	*S*	8	9	9	5	9	9	9	9
Upslope pressure	*UP*	1	1	1	7	2	3	8	2
River Stage	R_S	9	7	4	2	7	5	4	6
Rainfall	*PPT*	3	6	6	3	8	7	7	7

DOI: 10.1002/hyp

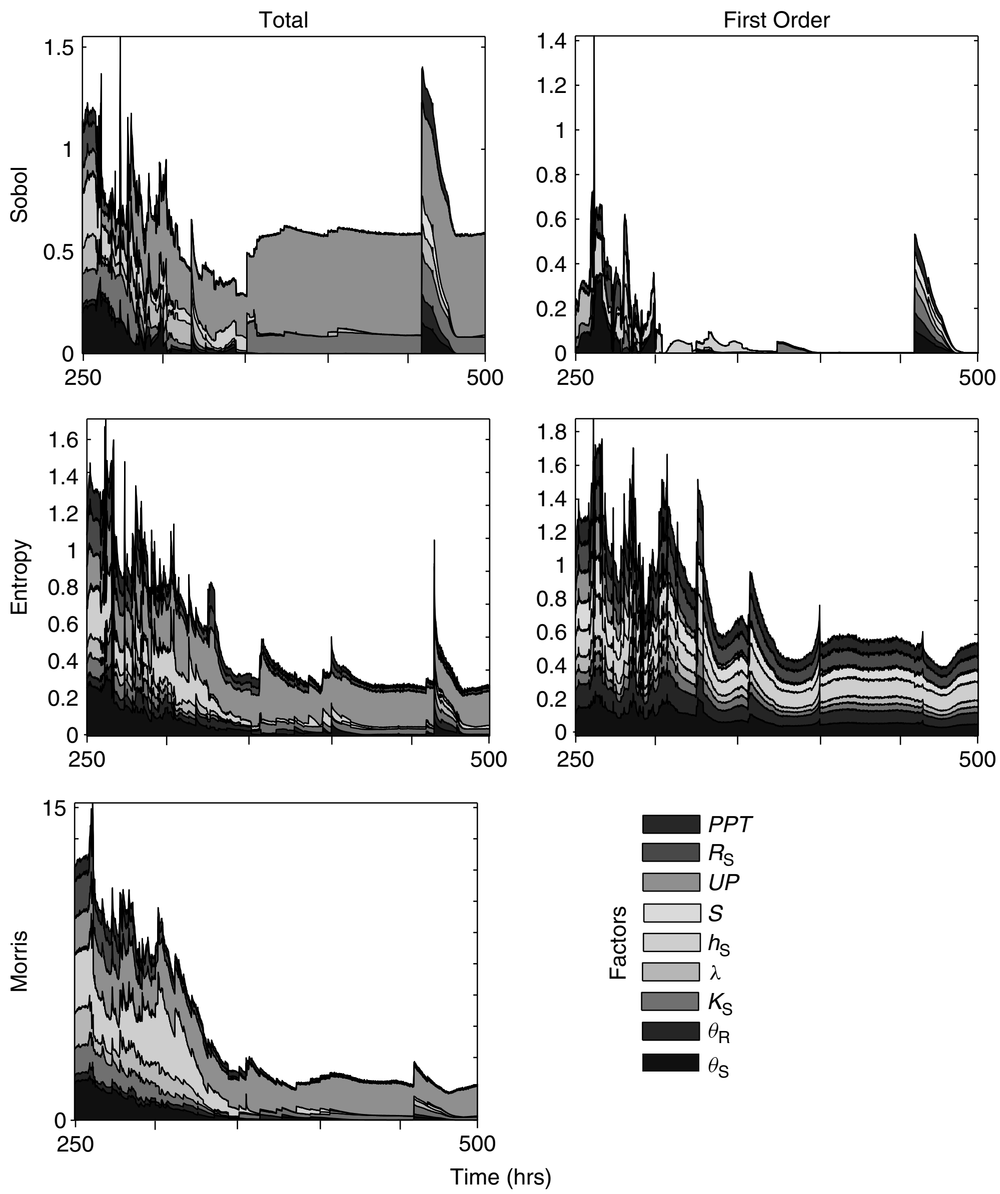

Figure 8. Factor importance variations through time for the Sobol, Entropy and Morris methods for the BC*h* simulation set and piezometer S3. The sequence of grey colours shown in the legend remains the same order for the graphs themselves

Sensitivity of the Darcian velocity and implications for hypothesis testing

We now take the behavioural parameter sets and look at the behaviour of the x component of the Darcian velocity at a location on the floodplain between piezometer S4 and S3. This location is at a distance $x = 5$ m away from piezometer S3, and has been used in this test-case before as a reference point as it is flooded relatively early on in the event. The x component of the Darcian velocity is a key state variable as this provides information on the flux of water into the back of the floodplain, and hence the hillslope, during the event. This velocity can also be used to calculate water residence time in the floodplain, and hence has implications for transport and residence of solutes within the floodplain groundwaters (see Claxton *et al.*, 2003 for more details). It is therefore very important to get an accurate prediction of the x-component Darcian velocity. Figure 9 shows scatter plots of the Darcian velocity vector magnitude (x component) at this point against the K_s, *UP* and H_s factors at a time of 370 hours when the flood is nearly at its full extent. We have seen in Figure 8 that during this time K_s and *UP* are particularly sensitive and this is again reflected in Figure 9. The structure in the K_s plot is evident, with a logarithmic increase in Darcian velocity magnitude with K_s value. The structure of *UP* and H_s are less simple, with some high values of these factors leading to high velocity values. Looking at the *UP* in detail, the highest values of velocity are found at a factor range centred about ~0·17.

We must determine what these results mean for hypothesis testing. For both the hydraulic head and the Darcian velocity predicted by the model, we have seen that a wide variety of behavioural results exist. The model factors change in their sensitivity with different

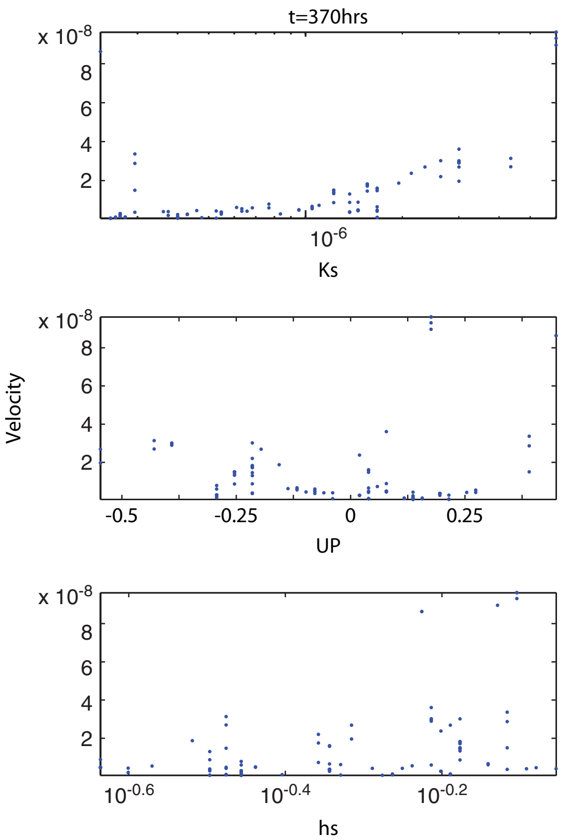

Figure 9. Scatter plots of the x-component of the Darcian velocity magnitude (at a point $x = 5$ m from piezometer S2) against K_s, UP and H_s factors

methods and through time. These sensitivities are sometimes complicated, with the example above of UP having a threshold-like behaviour predicting high velocity values in a particular range. We cannot reject the range of velocities predicted by our model (our hypothesis), but the key to confirming this is to take some more measurements, and continue the iterative scientific process (Hooper, 2001). However, we do now have an understanding of the major controls in our model on the flux of water (and solutes) into the floodplain during flood events.

UTILITY OF MULTI-METHOD GLOBAL SENSITIVITY ANALYSIS

The overall goal of any sensitivity analysis is to aid decision making in order to improve the model (Saltelli, 2000, p. 8–9). However, there is a danger that the results of any sensitivity analysis exercise does not directly help to improve the model but instead produces so much information that the modeller is overwhelmed and does not know what to do with such information. A great deal of time and computational resources have been invested in the MMGSA exercise presented in this paper, and a substantial amount of information has been generated. This is in stark contrast to the simple one-factor-at-a-time sensitivity analysis performed by Cloke (2003) and Claxton (2002). One needs to consider is the added understanding generated by the MMGSA worth the additional effort? Here, we would argue yes. The methodology we have demonstrated in this article can be used as a framework to help identify dominant factors. We use a simple ranking table (see Table IV) and a range of graphical techniques to help to summarize the information generated. We have shown that an appreciation of the complexities of the sensitivities of different factors and their dynamism in time can help to evaluate the next steps in fieldwork and the development of a conceptual model. Without this, the subtleties of the control of the boundary conditions on the pressure/velocity response would have been less clear. This then leads to an informed allocation of resources for future measurements at this site, which include nests of piezometers at the right location as well as experiments designed to constrain the K_s (and H_s) parameter values spatially.

CONCLUSIONS

Global sensitivity analysis is an essential tool in using numerical models to study hydrological processes. Here we have combined a fuzzy performance function with different methods of calculating global sensitivity, namely the Sobol, K–L Entropy and Morris methods, to perform a multi-method global sensitivity analysis (MMGSA). We have illustrated the necessity and suitability of this methodology for complex process-based models, with an application of a finite element subsurface flow model (ESTEL-2D) to a flood inundation event on a floodplain of the River Severn. The suite of results generated by the MMGSA highlight the complexities inherent in the sensitivities of the model factors, but demonstrated that for this particular flooding case, the upslope pressure boundary condition UP and the Brooks–Corey air entry value H_s were particularly important. These are different results to the simple one-factor-at-a-time sensitivity analysis originally performed for the case, where the Hydraulic Conductivity K_s was found to be the most important parameter. There are distinct variations between the different GSA measures, but one cannot be said to be more correct than another, they are just different ways of calculating sensitivity. However, having this wealth of information helps one to make an informed decision regarding the model behaviour.

We have highlighted the difficulties in computing sensitivities under the constraints of model non-convergence and numerical instabilities, and caution against the assumption that one can interpolate over holes in the response surface.

We demonstrate the utility of the MMGSA method for model understanding, in particular for the prediction of state variables such as Darcian velocity vectors directed towards the hillslope. We now have a better understanding of the Darcian velocity in our model, and the likely range of velocities possible. Understanding the controls on the numerical simulation of hydrological processes allows informed decision making: for example, where next to deploy field resources.

ACKNOWLEDGEMENTS

The Leighton Hydrology research site was initially sponsored by the Natural Environment Research Council and is now supported by Electricité de France. Hannah Cloke is supported by funding from the Nuffield Foundation and the University of London Central Research Fund. Florian Pappenberger is partially funded by the Flood Risk Management Research Consortium (FRMRC) which is supported by Grant GR/S76304 from the Engineering and Physical Sciences Research Council, in partnership with the Natural Environment Research Council, the DEFRA/EA Joint Research Programme on Flood and Coastal Defence, UKWIR, the Scottish Executive and the Rivers Agency (NI). The ESTEL-2D model is owned by Electricité de France. We would also like to gratefully acknowledge Amanda Mumford (*née* Claxton) for early ESTEL-2D modelling work on the River Severn flood events. Finally, we would like to thank Dr Bettina Schaefli and an anonymous reviewer for their valuable comments which greatly improved this manuscript.

REFERENCES

Bates PD, Stewart MD, Desitter A, Anderson MG, Renaud J-P, Smith JA. 2000. Numerical simulation of floodplain hydrology. *Water Resources Research* **36**: 2517–2530.

Beven K. 2006. A Manifesto for the equifinality thesis. *Journal of Hydrology* **320**: 18–36.

Beven K, Smith P, Freer J. 2007. Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Montovan and Ezio Todini. *Journal of Hydrology*. **338** (3–4): 315–318.

Brooks RH, Corey AT. 1964. Hydraulic properties of porous media. *Colorado State University Paper no 3*. Colorado State University, Fort Collins, Colorado.

Burdine NT. 1953. Relative permeability calculation size distribution data. *Transactions of American Institute of Mining, Metallurgical and Petroleum Engineering* **198**: 71–78.

Burt TP, Stewart MD, Bates PD, Claxton AJ, Anderson MG, Price DA. 2002. Water table fluctuations within the floodplain of the River Severn, England. *Journal of Hydrology* **262**: 1–20.

Campolongo F, Tarantola S, Saltelli A. 1999. Tackling quantitatively large dimensionality problems. *Computer Physics Communications* **117**: 75–85.

Celia MA, Bouloutas ET, Zarba RL. 1990. A General Mass-Conservative Numerical-Solution for the Unsaturated Flow Equation. *Water Resources Research* **26**(7): 1483–1496.

Claxton AJ. 2002. *Modelling subsurface water and chemical transport through floodplain systems*. PhD Thesis, University of Bristol, Bristol.

Claxton AJ, Bates PD, Cloke HL. 2003. Mixing of hillslope, river and alluvial groundwaters in lowland floodplains. *Ground Water* **41**(7): 926–936.

Cloke HL. 2003. *Modelling Riparian Hydrology and Streamflow Generation*. PhD Thesis, University of Bristol.

Cloke HL, Renaud J-P, Claxton AJ, McDonnell JJ, Anderson MG, Blake JR, Bates PD. 2003. The effect of model configuration on modelled hillslope-riparian interactions. *Journal of Hydrology* **279**: 167–181.

Cloke HL, Anderson MG, McDonnell JJ, Renaud J-P. 2006a. Using numerical modelling to evaluate the capillary fringe groundwater ridging hypothesis of streamflow generation. *Journal of Hydrology* **316**: 141–162.

Cloke HL, Anderson MG, Renaud J-P. 2006b. Development of a modelling methodology for the investigation or riparian hydrological processes. *Hydrological Processes* **20**(1): 85–107.

Coyle D, Buxton MJ, O'Brien BJ. 2003. Measures of importance for economic analysis based on decision modeling. *Journal of Clinical Epidemiology* **56**(10): 989–997.

Frey HC, Patil SR. 2002. Identification and review of sensitivity analysis methods. *Risk Analysis* **22**(3): 553–578.

Hall JW, Tarantola S, Bates PD, Horritt MS. 2005. Distributed Sensitivity Analysis of Flood Inundation Model Calibration. *Journal of Hydraulic Engineering* **131**(2): 117–126.

Helton JC, Davis FJ. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety* **81**(1): 23–69.

Hervouet JM, Bates P. 2000. The TELEMAC modelling system—Special issue. *Hydrological Processes* **14**(13): 2207–2208.

Hooper RP. 2001. Applying the scientific method to small catchment studies: A review of the Panola Mountain experience. *Hydrological Processes* **15**(10): 2039–2050.

Huang K, Mohanty BP, van Genuchten MT. 1996. A new convergence criterion for the modified picard iteration method to solve the variably saturated flow equation. *Journal of Hydrology* **178**: 1483–1496.

Krykacz-Hausmann B. 2001. Epistemic sensitivity analysis based on the concept of entropy. In *Proceedings of SAMO2001*, Madrid. Prado P, Bolado R (eds); 31–35.

Kullback S, Leibler RA. 1951. On Information and Sufficiency. *Annals of Mathematical Statistics* **22**(1): 79–86.

Liu H, Chen W, Sudjianto A. 2004. Relative entropy based method for global and regional sensitivity analysis in probabilistic design. In *Proceedings of DETC'04 ASME 2004 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Salt Lake City, Utah.

Meyer PD, Rockhold ML, Gee GW. 1997. *Uncertainty Analyses of Infiltration and Subsurface Flow and Transport for SDMP sites*. NUREG/CR-6565, US Nuclear Regulatory Commission.

Mokhtari A, Frey HC. 2005. Recommended practice regarding selection of sensitivity analysis methods applied to microbial food safety process risk models. *Human and Ecological Risk Assessment* **11**(3): 591–605.

Montanari A. 2005. Large sample behaviors of GLUE in assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research* **41**. DOI: 10·2/2004WR003826.

Morris MD. 1991. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics* **33**(2): 161–174.

O'Hagan A. 2006. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety* **91**: 1290–1300.

Pappenberger F, Beven K. 2004. Functional Classification and Evaluation of Hydrographs based on Multicomponent Mapping. *International Journal of River Basin Management* **2**(2): 89–100.

Pappenberger F, Beven KJ. 2006. Ignorance is bliss—or 7 reasons not to use uncertainty analysis. *Water Resources Research* **42**(5). DOI: 10·2/2005WR004820.

Pappenberger F, Beven K, Horritt M, Blazkova S. 2005. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *Journal of Hydrology* **302**: 46–69.

Pappenberger F, Iorgulescu I, Beven KJ. 2006a. Sensitivity Analysis based on Regional Splits (SARS—RT). *Environmental Modelling & Software* **21**(7): 976–990.

Pappenberger F, Matgen P, Beven K, Henry J-B, Pfister L, Fraipont de P. 2006b. Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources* **29**(10): 1430–1449.

Ratto M, Tarantola S, Saltelli A. 2001. Sensitivity analysis in model calibration: GSA-GLUE approach. *Computer Physics Communications* **136**(3): 212–224.

Renaud J-P, Cloke HL, Wang Y, Anderson MG, Wilkinson P, Lloyd D. 2003. Simulation numerique d'ecoulements en milieux poreux avec l'equation de Richards. *Revue Europeenne des Elements Finis* **12**(2–3): 203–220.

Rulon JJ, Rodway R, Freeze RA. 1985. The Development of Multiple Seepage Faces on Layered Slopes. *Water Resources Research* **21**(11): 1625–1636.

Saltelli A. 2000. What is Sensitivity Analysis. Chapter 1 in *Sensitivity Analysis*. Saltelli A, Chan K, Scott EM (eds). John Wiley & Sons, Ltd.: New York.

Saltelli A, Sobol IM. 1995. About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering & System Safety* **50**(3): 225–239.

Saltelli A, Tarantola S, Campolongo F, Ratto M. 2004. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons, Ltd.: New York.

Sobol IM. 1979. On the Systematic Search in a Hypercube. *Siam Journal on Numerical Analysis* **16**(5): 790–793.

Sobol IM. 1993. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiment* **1**: 407–414.

Stewart MD. 1998. *Severn fieldsite equipment installation report*. School of Geographical Sciences, University of Bristol.

Stewart MD, Bates PD, Anderson MG, Price DA, Burt TP. 1999. Modelling floods in hydrologically complex lowland river reaches. *Journal of Hydrology* **223**: 85–106.

van Genuchten MT. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* **44**: 892–898.

van Griensven A, Meixner T, Grunwald S, Bishop T, Diluzio M, Srinivasan R. 2006. A global sensitivity analysis tool for parameters of multi-variable catchment models. *Journal of Hydrology* **324**: 10–23.

Vogel T, van Genuchten MT, Cislerova M. 2001. Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions. *Advances in Water Resources* **24**: 133–144.

Young PC, Chotai A, Beven KJ. 2004. Data-based mechanistic modelling and the simplification of environmental systems. In *Environmental Modelling: Finding Simplicity in Complexity*. Wainwright J, Mulligan M (eds). John Wiley: Hoboken, New Jersey; 371–388.

DOI: 10.1002/hyp