© IWA Publishing 2010 Hydrology Research | 41.5 | 2010

Système Hydrologique Europeén (SHE): review and perspectives after 30 years development in distributed physically-based hydrological modelling

Jens Christian Refsgaard, Børge Storm and Thomas Clausen

ABSTRACT

The European Hydrological System (or *Système Hydrologique Europeén*, SHE) was initiated as a collaborative venture in 1976 between the Danish Hydraulic Institute (Denmark), Institute of Hydrology (UK) and SOGREAH (France). The present paper reviews the development history of the SHE and discusses the practical and scientific difficulties encountered during the different stages of the development. Comparison is made with eight other well-known model codes with respect to development stage and code dissemination among researchers and practitioners. Finally, the scientific developments and disputes on physically-based distributed modelling are discussed and the future perspectives outlined. The SHE venture has resulted in significant contributions to hydrological science, both in terms of model codes and new scientific insight. The fundamental scientific problems related to the inability to incorporate local scale spatial heterogeneity, scaling and uncertainty that were formulated are fundamentally still unresolved. Thus, in spite of the original visions, the hydrological community has not yet witnessed a model that in a universal sense (i.e. at all scales and for all internal variables) simulates accurate results for the right reasons. Instead, much of the scientific progress achieved during the recent years has dealt with how to live with these recognized problems.

Key words | distributed hydrological model, history, MIKE SHE, physically-based, SHE, SHETRAN

Jens Christian Refsgaard (corresponding author)
Geological Survey of Denmark and
Greenland (GEUS),
Øster Voldgade 10,
DK-1350 Copenhagen,
Denmark
E-mail: *jcr@geus.dk*

Børge Storm
Thomas Clausen
DHI,
Agern Alle 5,
DK-2970 Hørsholm,
Denmark

INTRODUCTION

Four decades ago Freeze & Harlan (1969) published their call for development of distributed physically-based catchment models. This was a paradigm shift as compared to the lumped conceptual models, e.g. the Stanford Watershed Model (Crawford & Linsley 1966), which was considered a pioneering work at that time and had been successfully used for some years. The first hydrological model code that was developed to fully meet the concept outlined by Freeze & Harlan (1969) was the Système Hydrologique Européen (SHE) (Abbott *et al.* 1986*a*,*b*). Today, two codes exist based on the original SHE concept: the MIKE SHE code further developed by DHI, Denmark (Refsgaard & Storm 1995; Graham & Butts 2005), and the SHETRAN code further developed by University of Newcastle, UK (Bathurst & O'Connell 1992; Bathurst *et al.* 1995).

The SHE provided a complete and integrated description of the land phase of the hydrological cycle by coupling the governing process equations of the major hydrological processes together. The processes included snowmelt, interception and evapotranspiration, subsurface flow in the unsaturated and saturated zones and surface flow on the ground surface and in rivers. All process descriptions were based on state-of-the-art knowledge. A framework called 'Chef d'orchestre' was designed to control how the individual numerical solutions were linked together, solved in parallel (in some cases iteratively) using appropriate time

doi: 10.2166/nh.2010.009

steps for the individual process descriptions. In both MIKE SHE and SHETRAN, the basic process descriptions have in various ways been modified and extended to include alternative or new process descriptions and numerical solution techniques.

Since the emergence of the SHE code, and subsequently the MIKE SHE and SHETRAN codes, several other codes of the distributed physically-based type have been developed. These codes can broadly be grouped in the following five categories:

1. Codes similar to SHE in concept and scope (Kutchment *et al.* 1996; Querner 1997);
2. Codes based on distributed groundwater models coupled with lumped or semi-distributed conceptual surface water modules (Refsgaard & Hansen 1982*a*,*b*; Miles & Rushton 1983; Christensen 1994; Wardlaw *et al.* 1994);
3. Codes primarily focusing on surface water or surface water near flow descriptions (e.g. Loague & Freeze 1985; Grayson *et al.* 1992*a*; Troch *et al.* 1993);
4. Codes based on coupling of existing groundwater codes and existing lumped conceptual or semi-distributed surface water codes (e.g. Perkins & Sophocleous 1999; Markstrom *et al.* 2008; Kim *et al.* 2008); and
5. Codes providing more detailed process descriptions, such as a full 3D description of the combined unsaturated-saturated zone and more sophisticated numerical solutions (e.g. VanderKwaak & Loague 2001; Panday & Hayakorn 2004; Therrien *et al.* 2006).

None of the above codes except from category (5) have capabilities that in principle are beyond those of the SHE family codes. The codes from the first three categories are, to our knowledge, not in operational use today. Codes from the fifth category, and in particular HydroGeoSphere (Therrien *et al.* 2006; Goderniaux *et al.* 2009), have the potential to become a new paradigm shift with capabilities significantly beyond those of the SHE type of codes.

During the years the SHE concept and its capabilities as outlined by Abbott *et al.* (1986*a*,*b*) have been subject to discussion and dispute in the scientific community (e.g. Beven 1989, 1996*a*,*b*; Bergström 1991; Grayson *et al.* 1992*b*; Refsgaard *et al.* 1996). In the practitioners' community, where distributed physically-based models have become more and more accepted and used, the debate has been more pragmatic focusing on functionality of the various model codes (Kaiser-Hill 2001).

The objectives of the present paper are: (a) to review and discuss the development history of the SHE and (b) to review and discuss the scientific and practical insights gained in connection with the development and application of distributed physically-based models during the last three decades.

DEVELOPMENT HISTORY OF SHE

Motivation and background

The need for a distributed physically-based hydrological modelling system was recognized at the beginning of the 1970s and resulted in the decision to develop the European Hydrological System (*Système Hydrologique Européen*) or SHE as a collaborative venture between three European organizations. Existing hydrological modelling systems of lumped conceptual type such as the Stanford (Crawford & Linsley 1966) were considered inappropriate for a range of analyses as they would not be able to adequately address problems arising from e.g. the adverse impacts of man's activities on the hydrological cycle (such as assessing impacts of land use change and simulating sediment and water quality processes at field and catchment scale, accounting for both surface water and groundwater processes).

The SHE development began in 1976 by establishing the 'Association pour le SHE' (ASHE) consisting of the British Institute of Hydrology, the Danish Hydraulic Institute (DHI) and the French consulting company SOGREAH. The ASHE was established partly to regulate the responsibilities and exploitation rights among the three rather different ASHE partners and partly to manage a repayable loan from the European Commission that funded some of the initial developments. In 1985 the British responsibility for the SHE was transferred from the Institute of Hydrology to the Water Resource Systems Research Unit at the University of Newcastle-upon-Tyne. In 1990 the French responsibility was transferred from SOGREAH to the Laboratoire d'Hydraulique de France. In both cases the key scientists were also transferred.

The three ASHE partners had complementary background and skills at the initiation of the SHE development. SOGREAH was one of the first organizations to apply computationally hydraulics techniques to support water management in practice. In particular, the model developed of the Mekong Delta in the mid 1960s was exceptionally detailed for its time. DHI moved into computational hydraulics around 1970 and developed two general modelling systems for handling one-dimensional and two-dimensional hydraulics flow systems for rivers and coastal areas, respectively. Both these organizations were pioneers in numerical techniques for hydrodynamic systems (computational hydraulics). The Institute of Hydrology, UK, an internationally leading research institute with state-of-the art knowledge on hydrological processes and some previous experience in hydrological modelling, provided complementary knowledge about hydrological processes into the partnership.

The authors have been involved in the development and application of the SHE and MIKE SHE at DHI since 1982. This paper is confined to a historical analysis based on our own experience through our work at DHI and, to a minor extent, the initiatives and work by DHI's ASHE partners.

The developments taken place during more than three decades can conveniently be divided into four stages:

- Stage 1: Initial development. This included the development of the concept and the first version of the model code, including its first test against field data on small research catchments.
- Stage 2: Consolidation. Here the code was applied in the very first commercial projects at DHI and in additional research projects. In this connection the code was improved with a series of enhancements in process descriptions, numerical efficiency and usability.
- Stage 3: Software packaging and dissemination. This was a stage where commercial applications and revenue from dissemination of the software became important to sustain further development and maintenance of the code.
- Stage 4: Integration. Here the code is not primarily being developed as a stand-alone code but instead being used as an element in a larger integrated modelling system.

Stage 1: initial development (1976–1986)

The first stage of the SHE development has been described in details by Abbott *et al.* (1986*a*) and only a brief summary is given here. The development of the concept and the first prototype ready for testing took approximately five years until 1981. During the following couple of years the coupled process equations were tested and improved before the code, in 1984, was able to produce its first real hydrographs. A schematic representation of the structure of the SHE from this stage is shown in Figure 1.

The first results published by Bathurst (1986) showed simulation results for five rainfall events with durations of 3–5 days for the 10.55 km^2 River Wye research catchment in Wales. The model had a spatial grid cell resolution of 250 m × 250 m (169 grid cells). The small number of grid cells and the short simulation periods were dictated by the limited capacity of computers at that time. Similar experiences were gained by some of the initial researchers outside the ASHE organization that tested the SHE code, e.g. the Hydrology Centre in New Zealand (Ibbitt 1985) and the University of Braunschweig, Germany (Rohdenburg *et al.* 1986).

This stage was characterized by many practical problems related to conceptualization, numerical algorithms and coding. The first scientific paper (Beven *et al.* 1980) appeared after four years, but it is worth noting that it took 10 years until the first model results were published in a peer-reviewed scientific journal.

Stage 2: consolidation (1986–1992)

After the initial tests the SHE code was applied in several studies (mainly externally funded research projects, but also consultancy projects). During this period the SHE applications were continuously pushed to the limit taking advantage of any progress in computer technology to create larger and more complicated models for larger areas. Some of the significant experiences gained and improvements to the Abbott *et al.* (1986*b*) version were as follows.

1. New process equations were introduced to accommodate specific needs or extend the capability of SHE. This included alternative evapotranspiration equations, alternative unsaturated zone Equations (simplifications

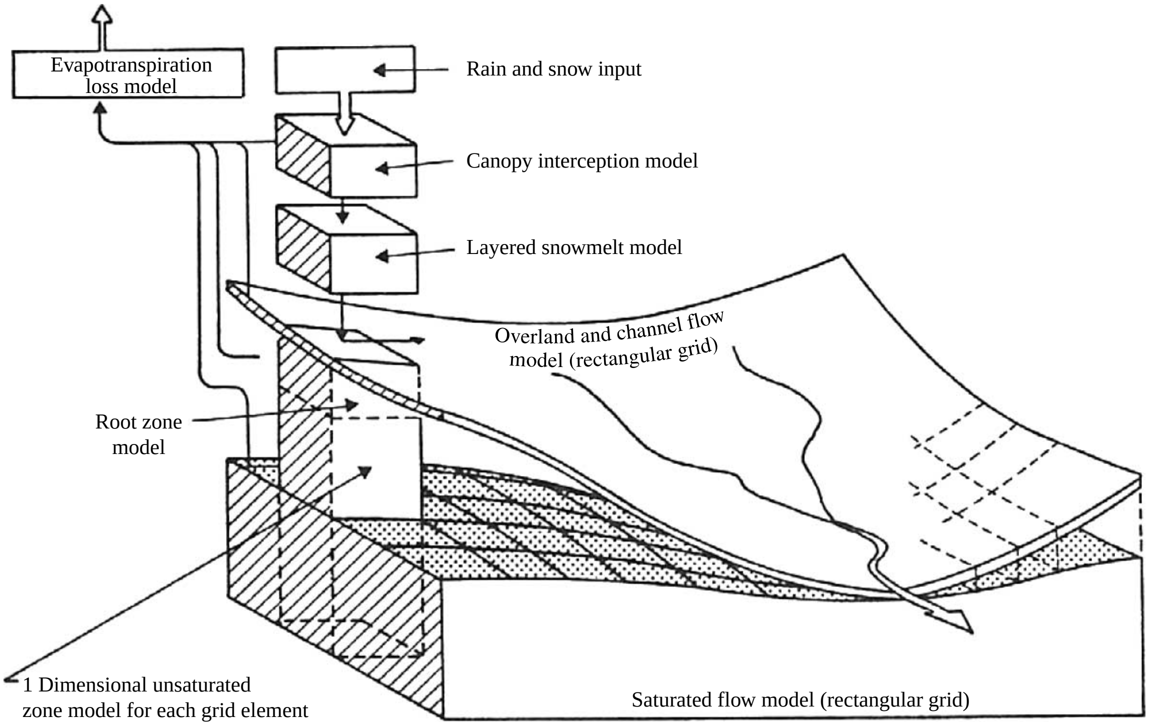

Figure 1 | The SHE structure illustrated during the initial stage. Figure from Abbott *et al.* (1986*b*) adapted from Beven *et al.* (1980).

of Richards' equation) and a new three-dimensional and multi-layered saturated zone solution including enhanced river–aquifer exchange.

2. The existing process equations did not allow appropriate simulations of some catchment processes, because it would require a spatial detail in the numerical models that was not feasible with the existing computer capabilities (this is true even today). For example, with a spatial grid cell resolution of e.g. 500 m in a regional model of several hundreds of squared kilometres, it would not be possible to simulate the 'interflow' type of response of the hydrograph. Several conceptual approaches were therefore introduced. A so-called tile drain option was introduced. This allowed the near-surface groundwater to be routed to the nearest downstream river point, when the water table rises to the tile drain level. Another important conceptual process was a so-called 'bypass flow' option in the soil column. It had been observed in the early applications that the flow through the soil matrix described by Richards' equation would not create sufficiently fast response on the groundwater table or in the drain flow. By allowing a fraction of the infiltration to be routed fast to the groundwater table under certain moisture conditions, it was possible to achieve much better results.
3. Add-on modules were developed for transport-dispersion, water quality, irrigation and soil erosion.
4. The numerical code was stabilized. Most of the problems, and definitely the most time-consuming problems to solve, were related to the couplings between the different components. Because the equations are solved in parallel (where one component provides the boundary condition to another) and because of vast differences in the temporal dynamics in the various components, a lot of special conditions had to be introduced. Examples include situations such as: (a) to which components should the water and the solute in the unsaturated zone from the previous time step go if the unsaturated zone disappears in the present time step; and (b) what should be the initial conditions of water and solute in the unsaturated zone if it was fully saturated (and hence part of the groundwater component) in the previous time step. The most difficult of all the couplings was that between unsaturated and saturated zones (Storm 1991).

Probably at least half of the work related to consolidating the code during this stage was linked to stabilizing the numerical algorithms of the various couplings.

5. The numerical algorithms of the code were optimized. This included two main aspects: (a) introduction of water balance controls and correction of errors that in some situations violated the water balance by more than 20% and (b) tuning of the computational efficiency, e.g. by speeding up the algorithms and introducing adaptive time steps and different time steps in the different components of the SHE, e.g. overland flow in the order of minutes and saturated zone in the order of days. This resulted in a code that was 5–10 times faster than the code from Stage 1 and that in almost all cases was free of water balance errors.
6. It was also realized that solving the Richards' equation for hundreds of cells in the horizontal grid was too time consuming. A method was developed to scan the model grid and identify groups of 'identical soil column conditions'. Columns with the same soil and vegetation characteristics, the same rainfall and potential evapotranspiration input and approximately the same depth to the groundwater table would give similar infiltration and groundwater recharge flow. Richards' equation was then solved in just one representative soil column within each category and the boundary conditions, infiltration and groundwater recharge, were transferred in each time step to the other cells within the same category. This is similar to the principles behind 'hydrological response units' commonly used in semi-distributed models (e.g. Knudsen *et al.* 1986). This would reduce the computational burden significantly (in some cases up to a factor of 10) and allow users to create a larger spatially distributed model.
7. A comprehensive pre- and post-processor program package was developed. This enabled users to work with spatially distributed maps, manipulate large datasets efficiently and visualize input data and simulation results.

Some of these developments took place as partner activities involving all three ASHE partners, but most of these were carried out by DHI in order to make it feasible to use SHE in commercial projects. This set strong limitations on the time available for delivering a final well-calibrated model. The single most important project during this period was a collaborative effort between the three ASHE partners and the National Institute of Hydrology in India (Jain *et al.* 1992; Refsgaard *et al.* 1992; Lohani *et al.* 1993).

Several research projects executed in parallel during this period resulted in the development of a handful of different project specific code versions, programmed by different scientists. Afterwards, it required a substantial effort to port all the important improvements back into one common code. This was done by DHI at the end of this second stage, and each of the three ASHE partners was given the new consolidated SHE code. This was the last common SHE code, before the development efforts of the three partners were divided into separate paths.

Stage 3: software packaging and dissemination (1992–2003)

General: ASHE partners

The experiences from the last years of Stage 2 gradually led to the recognition that substantial resources were required to continuously update the code with improvements from various research projects and, not least, with necessary adjustments dictated by continuous updates of computer software platforms. The code had become so complex that it required several man months of work every time a new version of the operational system was introduced. Such code maintenance was crucial for keeping the code alive, but scientifically it was not rewarding and not easy to finance through research project applications. Furthermore, usability improvements with more sophisticated graphical interface options were needed to make SHE attractive and possible to use, but this was not something that could attract research funding. Therefore, new strategies were necessary to secure funding for the further development. At the beginning of Stage 3 the legitimate interests of the three ASHE partners were so different that it led to three different strategies as described below.

- DHI decided on a commercial track. DHI invested in development of a software package with a graphical user interface. This became the MIKE SHE (Refsgaard & Storm 1995). The MIKE SHE was then used for three purposes: (1) to generate income through specialized consultancy jobs; (2) to act as a platform for further research and development; and later (3) to generate income through software sales of MIKE SHE.

- The University of Newcastle-upon-Tyne decided on a classical academic track. Newcastle made further developments of the code funded through research contracts. This became the SHETRAN (Bathurst & O'Connell 1992; Bathurst *et al.* 1995). The SHETRAN has since then been the platform for substantial research contributions from the Newcastle group.
- Laboratoire d'Hydraulique de France decided to associate with DHI on promoting and utilizing MIKE SHE in French-speaking countries.

These different strategies were a natural result of the different status of the three organizations. DHI is a self-owned technological service institute with a limited core funding. DHI mainly depends on research and consultancy contracts, primarily international, and had previously been successful in launching the MOUSE and MIKE11 codes as commercial software packages to generate revenues for sustaining and further developing these packages. The University of Newcastle-upon-Tyne, on the other hand, was a university with some core funding and was in addition able to continuously attract significant research funding.

The natural conclusion of these different tracks was a gradual dissolution of the ASHE partnership, although it officially did not take place until the loan to the EC had been repaid a few years later.

The MIKE SHE

The transformation of the SHE into the MIKE SHE began in the late 1980s and focused on developing a graphical user interface that could make the code easier to use. The first graphical user interface was developed for UNIX, but a few years later was also made available for desktop with Microsoft Windows using a windows emulator. The status of MIKE SHE by the end of this stage is summarized by Graham & Butts (2005) and the developments of additional options for process descriptions are illustrated in Figure 2.

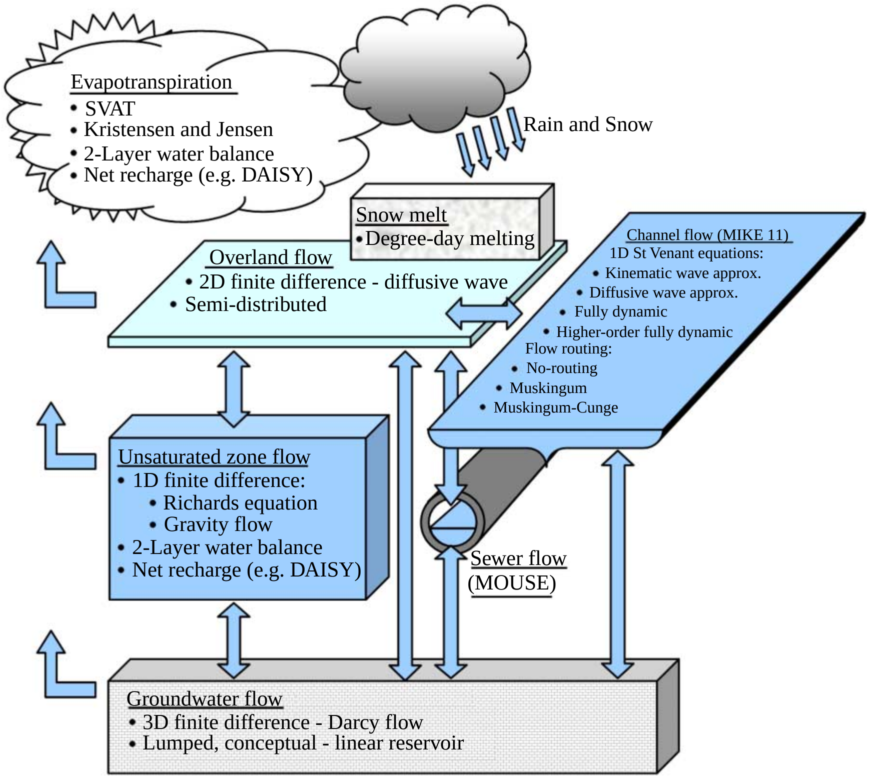

Figure 2 | Schematic view of the structure and processes in MIKE SHE, including the available numerical engines for each process (Graham & Butts 2005).

Specialized consultancy studies. As indicated by studies with intercomparisons between several integrated model codes (e.g. Kaiser-Hill 2001) the strength of MIKE SHE compared to other existing codes is its ability to handle large datasets and its integration with respect to both surface water and groundwater processes of water flow, transport and reactive processes. This has made MIKE SHE useful for specialized consultancy studies, often including a transfer of the final model to an authority. Three examples of major studies are as follows.

- The Danubian Lowland (1991–1995). An integrated modelling system with MIKE SHE as its core was established to assess the impacts of alternative operations of the new Gabcikovo hydropower scheme on the Danube bordering Slovakia and Hungary (Refsgaard *et al.* 1998). The modelling systems included complex interactions with feedbacks between river flows, sedimentation in reservoirs, infiltration from river and reservoir to aquifer, redox conditions in aquifer, agricultural yield and nitrate leaching influenced by groundwater tables and hence by river and reservoir operations.
- The South Florida Water Management District, SFWMD (1996–present). MIKE SHE has been applied for a number of local studies throughout the Florida region complementing the SFWMD regional model (Yan *et al.* 1998). The objective of the model applications was to establish management strategies for flood protection, water supply and environmental quality in the Everglades.
- The national hydrological model for Denmark, DK-model (1996–present). MIKE SHE has been used to establish an integrated groundwater/surface water hydrological model for the entire 43,000 km^2 of Denmark (Henriksen *et al.* 2003). The model has been used for assessing the environmentally sustainable groundwater abstraction taking into account that the conditions in the aquatic ecosystems are not allowed to deteriorate (Henriksen *et al.* 2008).

Research achievements. MIKE SHE has been used as platform for many research projects during the years, initially at DHI but more gradually at a large number of universities and research organizations. This is illustrated by the search results of journal papers and their citations in Table 1 (see discussion below on the weaknesses of this search). The ISI results indicate that MIKE SHE has obtained a widespread use in the scientific community with more than 100 journal papers authored by scientists in 24 different countries and with 32% of the papers being authored or co-authored by a scientist from Denmark. For comparison, the SHETRAN code has also had widespread scientific application but mainly in the UK.

Software product business. In connection with the consultancy and research projects, MIKE SHE was transferred to a large number of external users. During the last decade MIKE SHE has been maintained and serviced as a commercial software product, as part of DHI software product portfolio. This has further expanded the use of MIKE SHE among practitioners and researchers. Compared to other DHI software products for rivers, sewer networks or marine and coastal applications, the MIKE SHE product business has been minor relative to consultancy applications, albeit sufficiently large to sustain continuous maintenance and improvements.

Figure 3 shows the temporal development in income generated from sales and service of MIKE SHE and MIKE 11 (DHI's commercial river hydrodynamic packages). The income from MIKE SHE sales has tripled during the past 10 years, but is only 10–15% of the income for MIKE 11. Table 2 shows the number of organizations that have a current version of MIKE SHE installed. It is distributed quite evenly between universities, public authorities and private companies, and also has quite a wide geographical distribution. The number of installations is probably small compared to e.g. some of the main groundwater codes.

Stage 4: integration (2003–present)

During the last six years the MIKE SHE code has been modernized both with respect to process descriptions and user interface. The user community has steadily increased, but MIKE SHE has by no means become a standard tool among water resources practitioners or researchers.

Due to the increasing need to manage more complex problems in an integrated manner in connection with environmental and climate change impact studies, it is increasingly recognized that there is a need for

Table 1 | Statistics from Web of Science on journal papers and from Google on websites on searching with code name as keywords. For Google search the model code name AND 'model' was used together with the specific search listed in the notes below the table. The search was made on 30 June 2009

Model code	**ISI Web of Science**					**Google**	**Ratio of Google hits to ISI citations**
	Number of ISI papers	**Number of ISI citations to papers**	**First year of publication**	**Number of countries involved in (co)authoring ISI papers**	**Country of developer and % of papers (co)authored from this country**	**Number of hits (websites)**	
SHE	20	1,206[*]	1986	8	UK + DK + F (75)	–[†]	–
MIKE SHE	108	1,079[‡]	1993	24	DK (32)	13,000[§]	12
SHETRAN	41	437[‖]	1992	16	UK (83)	1,000[¶]	2
HydroGeoSphere	7	12[**]	2006	5	CA (100)	300[††]	25
GSFLOW	0	–[**]	–	–	–	300[‡‡]	–
SWAT	808	5,534[**]	1990	53	US (55)	1,660,000[§§]	300
TOPMODEL	370	7,122[**]	1984	39	UK (21)	41,000[‖‖]	6
HBV	136	1,574[**]	1991	32	SE (38)	18,000[¶¶]	11
MIKE 11	76	353[***]	1991	29	DK (25)	16,000[†††]	45
MODFLOW	536	2,110[**]	1992	59	US (46)	41,000[‡‡‡]	19
FEFLOW	44	167[**]	1996	19	DE (27)	11,000[§§§]	66

[*] 'European Hydrological System' OR 'Systeme Hydrologique Europeen' for the period 1976 to 1993.

[†] It is not possible to search for 'SHE' in Google and get a meaningful result. Furthermore the SHE codes ceased to exist before the Internet appeared.

[‡] 'MIKE SHE' OR 'MIKESHE' OR 'MIKE-SHE'.

[§] {'MIKE SHE' AND 'water' AND ['model' OR 'software']} OR {'MIKESHE' AND 'water' AND ['model' OR 'software']}.

[‖] 'SHETRAN' OR 'SHESED'.

[¶] {'SHETRAN' AND 'water' AND ['model' OR 'software']} OR {'SHESED' AND 'water' AND ['model' OR 'software']}.

[**] Name of respective model code as in first column of table.

[††] 'HydroGeoSphere' AND 'water' AND ['model' OR 'software'].

[‡‡] 'GSFLOW' AND 'water' AND ['model' OR 'software'].

[§§] 'SWAT' AND 'water' AND ['model' OR 'software'].

[‖‖] 'TOPMODEL' AND 'water' AND ['model' OR 'software'].

[¶¶] 'HBV' AND 'water' AND ['model' OR 'software'] excluding pages with ['DNA' OR 'virus' OR 'infect' OR 'hepatitis' OR 'cardio' OR 'hemoglobin'].

[***] 'MIKE 11' OR 'MIKE11' OR 'MIKE-11'.

[†††] {'MIKE 11' AND 'water' AND ['model' OR 'software']} OR {'MIKE11' AND 'water' AND ['model' OR 'software']}.

[‡‡‡] 'MODFLOW' AND 'water' AND ['model' OR 'software'].

[§§§] 'FEFLOW' AND 'water' AND ['model' OR 'software'].

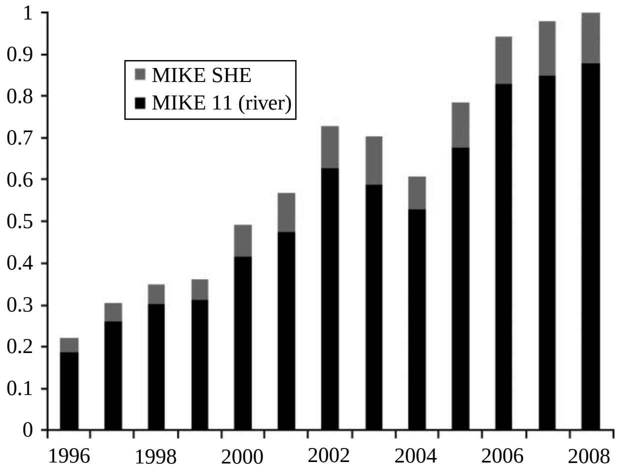

Figure 3 | The temporal development in income (normalized) from sales and service of MIKE SHE and MIKE 11.

integrated models. This typically includes coupling of model codes originating from different domains and different software developers. Coupling of model codes is, however, far from trivial. It requires suitable software techniques such as the MODFLOW specific approach (Banta *et al.* 2008) or the more general OpenMI (Gregersen *et al.* 2007). In addition, coupling of codes requires institutional and attitude changes.

MIKE SHE, as well as the other DHI software products, is provided to users in terms of executable code. These packages are internally compatible, and many couplings between the various packages have been developed during the past two decades. However, it is not easy to integrate such executable codes with other codes, and it is not possible for users to do so themselves. If, for instance, a user wants to replace one of the existing evapotranspiration modules in MIKE SHE with another description, it requires coding to be carried out by DHI. This reduces the flexibility of the codes. DHI has therefore engaged in the OpenMI development implying that, when DHI has made its codes OpenMI compatible, users can apply the OpenMI public domain software and themselves couple the DHI codes to other codes and in this way make adaptations. MIKE SHE is in the process of being made OpenMI compatible.

Table 2 | Number of organizations that have at least one installation of the current version of MIKE SHE (2008/2009)

Region	University	Public authority	Private company	Total
North America	63	23	24	110
South America	4	6	0	10
Europe	56	58	38	152
Asia	29	27	7	63
Africa	2	5	1	8
Oceania	12	9	8	29
Total	166	128	78	372

Key factors affecting the development

As compared to the development of new codes today, the speed of the SHE development was (particularly in Stage 1) very slow. The major difficulties foreseen at the outset were (Abbott *et al.* 1986*a*): (a) the resources (human expertise and computing capability) required to develop the SHE; (b) the necessity of keeping the computational burden associated with running the SHE within bounds of the current generation of mainframe computers; and (c) the potentially large amount of data which would be necessary for a practical application of the SHE.

The perceived limits of resources was probably the single most important factor in the decision for the three partners to make a joint development instead of moving alone on this unknown and difficult route. The major difficulty imposed by this decision turned out to be the logistics of integrating a code that was developed at three offices in three countries with very different programming traditions–one decade before the emergence of telefax and two decades before the Internet. There is no doubt that this obstacle reduced the implementation speed considerably. During Stage 1, the three organizations contributed with complementary capabilities and varying strengths over time. If the three organizations had started in parallel, we doubt that any of them would have been able to reach the goal alone. With today's resources we would definitely have decided for an alternative development strategy, but with the situation as it was in 1976 we believe that the decision was probably the best one.

The computational burden of running the SHE on the mainframe computers of that period was immense. Overnight computer runs and excessive computer bills put a serious limit to the efficiency of the programming work. The limitation was illustrated by the size of the first test run

(Bathurst 1986) with 169 spatial grids and simulations for 3–5 days events. Any hydrologist knew that this did not make much sense in practice, as the hydrograph for such short events to a large extent is controlled by (calibrated) initial conditions. However, larger catchments or longer time periods were simply not possible. This was the single most important obstacle for efficiency in the code development. When it became possible to port the code to a 32-bit UNIX based in-house microcomputer around 1985, we could pay our computer bills and get much improved turnaround times for model runs. This led to a dramatic improved efficiency of the entire development work and soon made it feasible to apply the model code for practical consultancy work.

When we came to the first model applications outside small research catchments, we came to the third of the perceived difficulties: the large amount of data required. This is discussed in the article describing the experiences in setting up the SHE for six catchments in India with a total of 15,000 km^2 (Refsgaard *et al.* 1992). This led initially to development of comprehensive data processing utility software to the SHE and gradually to the development of GIS-related functionality in later versions of MIKE SHE. MIKE SHE has proven to be well suited for handling large data amounts in connection with complex modelling tasks (Refsgaard *et al.* 1992; Styczen & Storm 1993*a*,*b*; Jayatilaka *et al.* 1998; Henriksen *et al.* 2003). The main problem today is access to useful spatial data.

In addition to the foreseen difficulties, the different natural interests of the three organizations eventually led to the dissolution of the ASHE at a time where a well-functioning code made it possible for the individual partners to continue on their individual tracks.

The adaptation of the SHE concept for practical applications has taken much longer than originally foreseen by the SHE developers. Although we have seen an increasing use of MIKE SHE in the university world in the recent years, the dissemination of the technology is still proceeding today with a speed that is much lower than its potential, when taking into account that both the scientific community and the policy makers emphasize the need for integrated water management at catchment scale. One reason for this is that some of the expectations of the capabilities of the new concepts and technology were unrealistically high. In other words, the hypotheses behind these expectations were wishful thinking that could not be confirmed by scientific testing. But even for fields of applications which proved scientifically sound, the speed of dissemination has been relatively slow. What is the reason for this?

Based on our experience in promoting the MIKE SHE technology, we believe that the main constraints are institutional. To use MIKE SHE for integrated catchment modelling is a complex task requiring an experienced modelling team with skills in a variety of disciplines such as hydrogeology, soil science, agronomy and computational hydraulics. Such a multidisciplinary team does not often exist in a single research organization or public authority, where professionals are typically organized in mono-disciplinary departments. It is for instance seldom that hydrogeologists and surface water hydrologists/ecologists belong to the same department, and often they are even in different organizations. Furthermore, due to the multi-disciplinary complexity of integrated catchment modelling, the requirements of modelling team expertise are higher than for simpler types of modelling.

COMPARISON WITH OTHER CODES

Dissemination and impact

To illustrate the impacts of the SHE codes and compare with other well-known model codes, a search was conducted in the databases Web of Science and Google (Table 1). We selected 11 codes that are applied both by researchers and by practitioners.

(1) The SHE family of codes described in this paper, i.e. SHE (Abbott *et al.* 1986*b*; Refsgaard *et al.* 1992), (2) MIKE SHE (Refsgaard & Storm 1995; Graham & Butts 2005) and (3) SHETRAN (Bathurst *et al.* 1995).

(4) The HydroGeoSphere (Therrien *et al.* 2006; Goderniaux *et al.* 2009). This is a new code developed by a consortium of mainly universities. A unique feature of HydroGeoSphere is that it solves the partial differential equations governing surface and subsurface systems, including the three-dimensional form of Richards' equation, simultaneously.

(5) GSFLOW (Markstrom *et al.* 2008) developed by US Geological Survey by coupling of two well-known codes: PRSM and MODFLOW. PRSM (Leavesley *et al.* 1983, 2005) is a precipitation runoff modelling system with conceptual process descriptions and with discretization into hydrological response units. MODFLOW is a physically-based groundwater model (see below). GSFLOW itself is very new and has therefore limited dissemination, while the two underlying codes are well established.

(6) SWAT (Arnold *et al.* 1993, 1998) developed by US Department of Agriculture is widely used for simulation of runoff and water quality at catchment scale. It has a very detailed agriculturally based root zone description operating on hydrological response units. The groundwater component is basically a baseflow routing using linear reservoir techniques.

(7) TOPMODEL (Beven & Kirkby 1979; Beven *et al.* 1984; Beven *et al.* 1995) is developed by University of Lancaster. It is distributed and has conceptual process descriptions. TOPMODEL is one of the most widely used precipitation runoff model codes in the research community.

(8) HBV (Bergström & Forsman 1973; Bergström 1995) is developed by the Swedish Meteorological and Hydrological Institute. It is a traditional rainfall–runoff model which has a lumped conceptual process description at its core, but which has been extended with discretization into hydrological response units. HBV is extensively used both for research and for practical applications.

(9) MIKE 11 (Havnø *et al.* 1995) is developed by DHI. It is a river modelling system comprising a number of optional rainfall runoff modules, e.g. the lumped conceptual NAM (Nielsen & Hansen 1973), hydrodynamic river routing tools and a large suite of sediment and water quality add-on components. MIKE 11 is a commercial software package extensively used both for research and practical applications.

(10) MODFLOW (McDonald & Harbaugh 1988; Harbaugh 2005) is developed by US Geological Survey. It is a three-dimensional finite difference groundwater model code, with many add-on components. MODFLOW is the most widely used groundwater modelling package both for research and practical applications.

(11) FEFLOW (Trefry & Muffels 2007) is a three-dimensional finite-element groundwater modelling code developed by the private company WASY. It is widely used for complex professional applications.

Altogether, the 11 codes represent a broad range of process descriptions from distributed physically-based (SHE, MIKE SHE, SHETRAN, HydroGeoSphere, GSFLOW) over more surface-water-oriented (SWAT, TOPMODEL, HBV, MIKE11) to pure groundwater model codes (MODFLOW, FEFLOW). As another dimension, they range from typical university codes (SHETRAN, HydroGeoSphere, TOPMODEL) to proprietary commercial codes (MIKE SHE, MIKE 11, FEFLOW) with government agency codes (GSFLOW, SWAT, HBV, MODFLOW) in between. It includes many old, well-established codes as well as a few new programs (HydroGeoSphere, GSFLOW).

Web of Science holds information on the number of ISI journal papers and their citations, while Google can provide information about the number of websites that contain information related to a given model code name. The key words used in the search are listed below Table 1. Such a simple search in general databases has considerable methodological weaknesses. For example, in Web of Science we know of papers that are not included in the list with the given search words (probably the authors have not used these terms in their list of key words nor in their title or abstract). The first TOPMODEL publication (Beven & Kirkby 1979) that has more than 1000 ISI citations is such an example. We have not corrected the figures by adding these papers, because we assume that similar biases exist for other codes. As nobody is aware of all papers dealing with a specific code it is not possible to produce numbers without errors. We can assume that the figures related to Web of Science are biased towards giving too low figures for all model codes, but we do not know if the bias is similar for all codes.

The hits on Google are more difficult and less credible than those on Web of Science. With a first search simply using a model code name, a lot of irrelevant sites appeared. We therefore required that the website contained three words: (1) the name of the model code, (2) the term ‘water’ and (3) one of the two terms ‘model’ or ‘software’. With these search criteria, we checked a number of the hit websites to ensure that the number of irrelevant sites was below 10%. This was the case for all codes, except the HBV where we put additional requirements to the search criteria. This procedure has probably resulted in an underestimation of the true number of websites dealing with specific codes because e.g. you may have websites dealing with MIKE SHE without containing the term 'water'. Although all codes were treated equally, we cannot know whether the errors introduced are of the same relative magnitude for all codes.

Due to these methodological problems the numbers in Table 1 are uncertain, but we are confident that the orders of magnitude reflect realities. In addition to the uncertainty of the numbers themselves, it should be noted that the numbers do not directly show the dissemination and impact of the model codes. They are only indicators with room for alternative interpretations. Altogether, the numbers in Table 1 must be interpreted with great care. With these reservations, the following can be inferred from Table 1.

- The Web of Science figures may be seen as indicators of the dissemination and impact of the codes in the scientific community. HydroGeoSphere and GSFLOW have been launched so recently that they have not yet had a chance to prove themselves. It appears that all the other codes are widely used, both in terms of citations and in terms of actual research carried out by scientists outside the country where the respective codes were developed. This also applies to the SHE, MIKE SHE and SHETRAN, where it is particularly noted that the relatively few papers from the first period of the SHE development have been very widely cited. It can also be seen that the MIKE SHE is more widely used by scientists outside Denmark than the SHETRAN is used by researchers outside UK. This is most likely because MIKE SHE is commercially marketed globally and has a graphical user interface that makes it easier for a new user.
- The code with the most widespread scientific use in terms of authors spread over the largest number of countries is MODFLOW. This is probably because MODFLOW is the 'de facto' industry standard in groundwater modelling that other groundwater model codes need to benchmark themselves against. However, it should be noted that the two commercial codes MIKE 11 and FEFLOW have a very large geographical distribution in terms of scientists (co)authoring papers relative to their much lower number of journal papers.
- The Google hits include both websites from the scientific community and websites from the practical water management community (government, water authorities, water companies, consulting companies, etc.). Thus the number of Google hits can be interpreted as the impact sampled much beyond the scientific community and hence reflecting the dissemination towards practitioners. SWAT is the model code with by far the largest numbers of Google hits. This may indicate that SWAT has a widespread use among practitioners.
- When looking at the ratio of number of Google hits and number of ISI citations (last column in Table 1) it is seen that the university codes (SHETRAN and TOPMODEL) clearly have the smallest ratios, while two of the commercially marketed codes (MIKE 11 and FEFLOW) together with SWAT have large ratios. It is interesting to note that MIKE SHE, which is also a commercial code, behaves very differently from MIKE 11 and FEFLOW. Instead, MIKE SHE has a Google/ISI ratio comparable to the public domain codes HBV and MODFLOW. This is well in line with the observation described above that MIKE SHE has never become a real commercial success relative to e.g. MIKE 11 and that there continues to be considerable interest for using MIKE SHE as a tool in a broad spectrum of research projects.

Development stage

Different codes have different development histories based on funding opportunities, market demands, institutional context and personnel involved. Based on the hypothesis that the SHE development history reflects common features of code developments, we have compared the SHE family of codes to the eight other model codes. For this purpose we have made an attempt to categorize the 11 codes in four development stages corresponding to the four stages in which the SHE development was described above (Table 3).

From Table 3, the following observations can be made.

- The entire SHE development history is exceptionally long (more than 30 years as independent code). Today nobody can probably afford a Stage 1 with 10 years

Table 3 | Comparison of development stage of hydrological model codes. The development stages correspond to the four phases described in the SHE development history

Model code	Type of model	Organization developing code	Development stage*	Comment	References
SHE	Distributed physically-based	ASHE	2	Abandoned as an operational code in 1992	Abbott *et al.* (1986*b*) and Refsgaard *et al.* (1992)
MIKE SHE	Distributed physically-based	Private	3–4		Refsgaard & Storm (1995) and Graham & Butts (2005)
SHETRAN	Distributed physically-based	University	2		Bathurst *et al.* (1995)
HydroGeoSphere	Distributed physically-based	University	1–2	Includes 3D Richards', fractured media	Therrien *et al.* (2006)
GSFLOW	Distributed Physically-based	Government Agency	1–2	Coupling of MODFLOW and PRMS	Markstrom *et al.* (2008)
SWAT	Semi-distributed Conceptual	Government Agency	3	Mainly surface water	Arnold *et al.* (1993)
TOPMODEL	Distributed conceptual	University	2	Mainly surface water	Beven & Kirkby (1979) and Beven *et al.* (1984, 1995)
HBV	Semi-distributed conceptual	Government agency	3–4	Mainly surface water	Bergström & Forsman (1973) and Bergström (1995)
MIKE 11	Lumped conceptual	Private	4	Mainly surface water	Nielsen & Hansen (1973) and Havnø *et al.* (1995)
MODFLOW	Distributed physically-based	Government agency	3	Groundwater, build into software packages by private companies	Mcdonald & Harbaugh (1988) and Harbaugh (2005)
FEFLOW	Distributed physically-based	Private	3	Groundwater	Trefry & Muffels (2007)

*Characteristics of development stages: *Stage 1: Initial development*: The code is under constant development and can only be used by the development team; *Stage 2: Consolidation*: The code is more stable and basically documented, but not necessarily very user friendly. It will sometimes require minor coding to adapt it to particular applications and it will typically require that the modellers develop additional software pieces to support pre- and post-processing; *Stage 3: Software packaging and dissemination*: The code is now converted to a professional, well-documented software package with graphical user interface and powerful pre- and post-processing supporting tools; *Stage 4: Integration*: The code is a professional software package integrated into a broader software platform.

before producing results. Most new codes today therefore do not start from scratch as the SHE did. Instead, they merge some existing codes and add something new. This is exactly what GSFLOW and, to a less extent, HydroGeoSphere have done. Because the available tools for code development are much more advanced than 30 years ago, the new codes can today move relatively quickly through Stage 1 and into Stage 2.

- There is a long route for new codes such as HydroGeoSphere until it reaches Stages 3–4. Stage 2 involves a consolidation requiring that test results and experience from a large number of different types of model applications are built into the code. This requires many years of work. New codes that are developed by coupling existing, well-tested codes can obviously do this faster than new codes developed from scratch. For instance, GSFLOW can be assumed well tested as far as the old modules are concerned, so that it is mainly the couplings that need to be consolidated. In this respect it is worth noting that the largest problems in stabilizing the SHE code in Stage 2 were related to the couplings. For HydroGeoSphere, which also includes concepts that are not yet well tested (such as fractured flow at different scales and three-dimensional Richards' equation at catchment scale), more work and hence more time can be expected.
- All the codes that have reached Stages 3–4 are developed by private organizations or government agencies. University codes tend to stop at Stage 2 and leave the commercialization aspects of software packaging to other organizations.

SCIENTIFIC DEVELOPMENTS AS SEEN IN RETROSPECT

The pioneering achievement

As evident from the number of citations to the first SHE publications (Table 1), the development of the first version of the SHE (Abbott *et al.* 1986*a*,*b*) and subsequently the first test on field data from a small research catchment (Bathurst 1986) were important achievements for the hydrological science. It represented a new approach and the first of its kind with a physically-based representation. Although the equations were all well known from various disciplines (soil science, computational hydraulics, groundwater modelling), the coupling of these equations constituted a quantum leap in complexity compared to contemporary simulation codes for catchment hydrology.

The next step to test the approach on ordinary-sized catchments with data that could be obtained from various sources was the study of six large catchments in India reported by the end of Stage 2 (Jain *et al.* 1992; Refsgaard *et al.* 1992) and the application to nutrient modelling for a 425 km^2 Danish catchment (Styczen & Storm 1993*a*,*b*). At that time lumped conceptual catchment model codes such as HBV (Bergström & Forsman 1973) and NAM (Nielsen & Hansen 1973) had been used operationally for two decades, typically for catchments ranging from a few squared kilometres to more than 10,000 km^2. At the same time, distributed physically-based models had mainly been tested on flood events on small (0.04–10 km^2) catchments that would typically have a good data coverage due to experimental instrumentation (Loague & Freeze 1985; Bathurst 1986; Grayson *et al.* 1992*a*,*b*; Troch *et al.* 1993). To our knowledge the only examples until then of distributed physically-based model studies including applications on catchments of several hundred squared kilometres and continuous simulation for periods of several years were coupled groundwater/surface water models (Miles & Rushton 1983; Christensen 1994; Wardlaw *et al.* 1994) that all had distributed physically-based groundwater components and lumped (or semi-distributed) conceptual surface water components.

During the following few years, a few additional catchment scale studies with continuous simulations of distributed physically-based models emerged. One example is Querner (1997) who applied the MOGROW to the 6.5 km^2 Hupselse Beek catchment simulating both discharge and groundwater head dynamics. Another example is Kutchment *et al.* (1996) who simulated surface water processes for the 3,315 km^2 Ouse catchment. The study of Kutchment *et al.* (1996) had many similarities to the India study (Jain *et al.* 1992; Refsgaard *et al.* 1992) with respect to model conceptualization and conclusions.

The main scientific contribution of the first SHE study (Abbott *et al.* 1986*a*,*b*; Bathurst 1986) was therefore the development of the first distributed physically-based model, while the study in India (Jain *et al.* 1992; Refsgaard *et al.* 1992) was the first to demonstrate that distributed physically-based models could be established for catchments of more than 1,000 km^2 and with ordinary data availability.

Disputes about the approach

From the very beginning the SHE concept was controversial in the scientific community. Initially, a key criticism was that the concept was unrealistic. The 10-year long initial development stage and the severe limitations dictated by the available computer facilities indicated that the SHE was not, as originally anticipated, realistic for practical applications in the 1970s or early 1980s. From a pure application point of view, it may be argued that the development started too early; perhaps that made it even more interesting as a research venture. When we were eventually–towards the late 1980s–able to produce simulations for ordinary catchments for periods of several years (during overnight computer runs), this debate ceased.

In the few years after the first SHE publications, Beven (1989) raised a more fundamental critique against the whole concept and the way physically-based models had been promoted by e.g. Abbott *et al.* (1986*a*) and Bathurst (1986). He did not agree with the expectations about capabilities and potential achievements that the physically-based models described in these early SHE papers. Beven pointed to, amongst others, the following key problems: (a) the process equations were simplifications leading to model structure uncertainty; (b) spatial heterogeneity at subgrid scale was not included in the physically-based models (the current generation of distributed physically-based models were in reality lumped conceptual models at a more sophisticated level) and (c) there was a great danger of overparameterization, if it was attempted to simulate all hydrological processes thought to be relevant and fit the related parameters against observed discharge data only.

Beven concluded that, for future applications, attempts should be made to obtain realistic estimates of the uncertainty associated with their predictions, particularly in the case of evaluating future scenarios of the effects of management strategies.

In retrospect, it is remarkable that the response from the SHE teams to the severe critique by Beven (1989) did not emerge immediately, but only developed gradually. In the subsequent SHE publications (Bathurst & O'Connell 1992; Refsgaard *et al.* 1992), a small part of Beven's critique was discussed by acknowledging that the process representation at large grid squares (1 or 2 km) was causing significant violations of some of the process descriptions, that "some degree of lumping and conceptualization had to take place at the grid scale", and that "scale problems were important". Both papers concluded that the SHE was indeed a suitable tool to support water management, while Beven (1989) had stated that the physically-based models "are not well suited to applications to real catchments", because the physical equations describing surface and unsaturated flow will not apply in domains with sub-grid scale heterogeneity. Furthermore, neither Refsgaard *et al.* (1992) nor Bathurst & O'Connell (1992) commented on Beven's main conclusion on the need for uncertainty assessment.

We believe that the main reason for this apparent disagreement lies in ambiguous terminology: it was not very clearly stated what was meant by "suitable" or "not well-suited". As far as our views are concerned, we did not clearly differentiate between universal applicability and site and purpose-specific applicability of a model until a few years later (Refsgaard & Knudsen 1996; Refsgaard 1997). Following this fundamental shift from the views of universal applicability expressed by Abbott *et al.* (1986*a*,*b*), several studies have been reported illustrating and discussing the limitations of the SHE code for specific applications (Refsgaard 1997; Refsgaard *et al.* 1999; Andersen *et al.* 2001, 2002; Hansen *et al.* 2007, 2008, 2009).

Numerous authors have contributed to the debate on the validity and usefulness of the SHE type of approach. As an example, when discussing the limitations of the process descriptions and the often quite poor accuracy of simulations internally in a catchment of physically-based models, Grayson *et al.* (1992*b*) noted that "there is a certain arrogance associated with physically-based models regarding their superiority over lumped parameter or empirical models" and that "the real issue is whether, in practice, there is any difference between such models except for the vastly increased time required to calibrate the numerous parameters associated with physically-based models". As another example, Seyfried & Wilcox (1995) discussed the critical importance of spatial variability in relation to process descriptions, grid scales and measurement scales. They noted the irony that while description of spatial variability is presumed to be a strength of physically-based models, it is only possible to include spatial variability if it can be explicitly described. They concluded that it is necessary to include stochastic approaches to account for spatial variability when using physically-based models.

Good modelling practice

Seen in the perspective of present protocols for good modelling practice (Refsgaard & Henriksen 2004; Refsgaard *et al.* 2005) the original SHE publications (Abbott *et al.* 1986*a*,*b*; Bathurst 1986; Bathurst & O'Connell 1992; Refsgaard *et al.* 1992) reveal some significant shortcomings, such as: (a) overparameterization, i.e. too many parameter values were assessed through calibration; (b) model capabilities were stated without proper model validation tests; and (c) uncertainty assessments were completely absent.

Calibration and overparameterization

It is today generally recognized that the number of free parameters to be adjusted through calibration should be limited to as few as possible to ensure that there is sufficient information content in the observed calibration data to provide robust estimates of the parameter values (Hill 1998). If too many parameter values are calibrated the result will often be a nice fitting of the discharge hydrograph or another calibration target during the calibration exercise, but a significantly poorer performance against other independent data. The first SHE test (Bathurst 1986) used a calibration approach allowing parameter values to vary as required to fit the observed data during the calibration phase. This approach was criticized by Beven (1989) as resulting in overparameterization. In the next major SHE study (Jain *et al.* 1992; Refsgaard *et al.* 1992), a more rigorous approach was developed where representative parameter values were associated with individual soil types, vegetation types, geological layers, etc. This reduced the number of free parameter coefficients to be adjusted in the subsequent calibration procedure. This procedure still resulted in 26 parameters to be calibrated for the 820 km^2 Kolar catchment in India, however. An attempt to reduce the number of free parameter values for the Senegal river basin, where the data availability is not better than the Indian catchment, resulted in a parameterization with only four free parameters (Andersen *et al.* 2001) implying that the parameterization approach adopted in the initial SHE studies were not yet finally developed. An attempt to assess all parameter values from field data and literature and provide simulations of discharge and nitrate leaching to groundwater without any model calibration was made by Refsgaard *et al.* (1999) with mixed success.

Model validation issues

It is today generally recognized that a model's predictive capability should be tested before specific statements on its field of applicability are made. It is also common wisdom today that model capabilities will vary from one catchment to another and from one type of application to another. Statements on model capabilities can therefore not be made universally but should be conditioned on specific sites and specific types of applications (Refsgaard & Henriksen 2004). In the initial SHE studies (e.g. Abbot *et al.* 1986*a*,*b*) there were many statements on potential applicability of the SHE code that easily could be interpreted as undocumented claims or over-selling.

During the course of the SHE development model validation, methodologies were applied and further developed (e.g. Ewen & Parkin 1996; Parkin *et al.* 1996; Refsgaard & Knudsen 1996; Refsgaard 1997). Much of this work was based on the principles proposed by Klemes (1986) that a model should be tested to show how well it can perform the kind of task for which it is specifically intended. Klemes' hierarchical test scheme include four types of tests corresponding to different situations with regard to whether data are available for calibration and whether the catchment conditions are stationary or the impact of some kind of human intervention has to be simulated: (a) the classical split-sample test; (b) the proxy-basin test, when there is not sufficient data for calibration; (c) the differential split-sample test, when a model is intended for simulation of impacts of human interventions such as land use change, climate change or groundwater abstraction; and (d) proxy-basin differential split-sample test, which is a combination of (b) and (c). One of the implications of this was a recognition that a distributed model that is validated for simulating catchment response often performs much poorer for internal sites and that it therefore can only be assumed valid with respect to the outputs that have been directly validated.

The principle that a model should never be considered universally validated, but can only be conditionally validated restricted by the availability of data and specifically performed validation tests, is well in line with Lane & Richards (2001) who argue that "evidence of a successful prediction in observed spaces and times (conventional validation) cannot provide a sufficient basis for use of a model beyond the set of situations for which the model has been empirically tested". The principles are also in accordance with the new coherent philosophy for modelling of the environment proposed by Beven (2002*a*) where he argues that it is required to be able to "define those areas of the model space where behavioural models occur".

This considerable development from the attitude of the early SHE studies to those of today reflects a general development in hydrological modelling over these decades. Other integrated physically-based models from the same early period (Miles & Rushton 1983; Christensen 1994; Wardlaw *et al.* 1994) had the same characteristics, i.e. only focus on calibration and model prediction but no model validation tests against independent data. The first SHE studies (Bathurst 1986; Refsgaard *et al.* 1992) applied the classical split-sample testing, which was also the standard practice in the community of lumped conceptual rainfall–runoff modelling (WMO 1975, 1988). No studies of those days went beyond the split-sample tests, although the studies often used or discussed the use of physically-based models for assessing the impacts of changing catchment conditions such as changes in land use or groundwater abstraction.

Uncertainty assessment

Uncertainty assessment is today a mandatory element of a good modelling practice (Refsgaard & Henriksen 2004; Refsgaard *et al.* 2005; Rosbjerg & Madsen 2005; Jakeman *et al.* 2006; Pappenberger & Beven 2006). This was not the case in the initial SHE studies, where the uncertainty issues were absent. Since then uncertainty has been addressed in several studies (Thorsen *et al.* 2001; Christiaens & Feyen 2002). The most difficult uncertainty problem to handle today is probably the model structure uncertainty. A common approach to deal with this is using multiple conceptual models (IPCC 2001; Beven 2002*b*; Neuman & Wierenga 2003; Refsgaard *et al.* 2006). MIKE SHE is due to its modular structure with several alternative conceptualizations for each of the components well suited for such purposes (Butts *et al.* 2004).

Although examples of model prediction uncertainty assessments had been reported previously from different modelling disciplines (e.g. Beck 1987), the first to emphasize the need to systematically perform uncertainty assessments related to catchment model predictions was probably Beven (1989). This was followed by Binley *et al.* (1991) who used Monte Carlo analysis to assess the predictive uncertainty for the Institute of Hydrology Distributed Model and by the introduction of the Generalized Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley 1992) after which uncertainty in catchment modelling was high on the agenda in the scientific community.

Fields of application

Initial expectations

According to Abbott *et al.* (1986*a*) it was envisaged that the SHE would be applicable to almost any kind of hydrological problem, although extensive further development and refinement would still be needed. They acknowledged that there were many problems for which the necessary solutions can be obtained using cheaper, conventional rainfall–runoff models. They therefore envisaged that the main types of SHE applications would be for the more complicated problems such as those listed below.

- *Catchment changes*. The strength of the SHE for assessing impacts of land use changes was believed to be that the parameters have direct physical interpretation and therefore can be assessed for the new state of the catchment before the new state actually occurs.
- *Ungauged catchments*. Here the expectation again was that the direct physical meaning of the model parameters would allow parameter values to be assessed directly from field data without calibration or at least using a much shorter hydrometeorological record for calibration.

- *Spatial variability in catchment inputs and outputs*. It was expected that distributed models would be better able to make use of distributed input data and simulate processes within a catchment, e.g. the effects on flood flows due to different directions of storm propagation.
- *Movement of pollutants and sediments*. As most water quality and sediment problems occur at much smaller scale than catchment scale, it was expected that distributed flow models could provide improved possibilities for water quality and sediment simulations.

Documented fields of application: status today

A number of intercomparison studies have been conducted to assess the capability of different model types, including physically-based models (Loague & Freeze 1985; Michaud & Sorooshian 1994; Refsgaard & Knudsen 1996; Reed *et al.* 2004). Based on these studies it can be concluded that if the purpose of modelling is limited to simulation of runoff under stationary catchment conditions and if data exist for calibration purpose, there is no scientifically documented reason to go beyond lumped conceptual models. We believe that the only factor that may change that conclusion is the introduction of new spatial data from new airborne or satellite sensors. Whereas these new data types have proven to have great value for many hydrological purposes and for special conditions (e.g. snow cover), they have in general not yet documented that they can provide distributed models with comparative advantages in simulation of catchment runoff.

The study by Refsgaard & Knudsen (1996) also suggested that, except for some indications of more narrow prediction intervals, physically-based models are not significantly better than lumped conceptual models for simulation of runoff from ungauged catchments.

The key problem for ungauged catchment simulation is related to assessing model parameters directly from field data. This has turned out to be possible in some cases and for some types of parameters, while inadequate process descriptions and scale problems put severe limitations to this approach in other situations (Beven 1995; Refsgaard 1997).

There have been many successful applications with water quality simulations using physically-based models, particularly for subsurface processes (e.g. Jensen *et al.* 1993). However, heterogeneity of geology and subsurface parameters at scales smaller than the computational unit in a catchment model is posing a fundamentally difficult upscaling challenge that need to be addressed before successful water quality simulations can become standard applications at catchment scale (Hansen *et al.* 2007, 2008).

CONCLUSIONS

The development of the Système Hydrologique Europeén (SHE) initiated more than 30 years ago has had a significant impact on the hydrological science. First of all, the development itself and some of the first research applications were major achievements and scientifically novel. Thus, for three decades the SHE and its daughter products MIKE SHE and SHETRAN were used as references and benchmarks when new codes were developed. Many codes exist that are superior to MIKE SHE and SHETRAN on single domains. However, it was not until the emergence of HydroGeoSphere during the last five years that we have seen a code for integrated physically-based catchment modelling that, with respect to concepts, versatility and numerical engines, goes beyond and has the potential to become superior to the well-proven MIKE SHE.

Secondly, the scientific challenges encountered (e.g. related to scaling problems and prediction uncertainty) when the new SHE concept was tested generated scientific discussions that contributed to the development of new insight and new methodologies for hydrological modelling.

The uptake of integrated physically-based modelling technology among practitioners has been much slower that expected by the SHE developers. A severe obstacle in this regard is that such modelling is very complex with very high demands to modelling teams in terms of scientific training, modelling experience and multi-disciplinary knowledge covering hydrogeology, soil science and surface water hydraulics. The last requirement, which is dependent on local institutional settings, is often not fulfilled.

If we revisit the key problems for distributed physically-based models formulated by Beven (1989), we have to conclude that the fundamental scientific problems are unresolved. First of all, model structures and process equations are still far from perfect. This is in many cases the dominant source of uncertainty when assessing prediction uncertainty (Refsgaard *et al.* 2006). Secondly, spatial heterogeneity at subgrid scale is generally still not included in the models. This is recognized as a major weakness of distributed models (Hansen *et al.* 2007, 2008). In spite of a few research studies to include spatial heterogeneity into stochastic partial differential process Equations (e.g. Jensen & Mantouglou 1992) these efforts have been limited to single domains and have not been permanently implemented in operational model codes. Thirdly, assessment of the physically-based parameters through direct field measurements and/or through calibration is confined by limited data and scale issues.

The scientific progress made during the past two decades on these fundamental issues has been very limited. Instead, the efforts have been devoted to finding solutions on how to live with these recognized problems. The most important contribution in this respect has so far been increased awareness and a broad suite of tools to analyze model prediction uncertainties. Thus, while Beven's call for scientific solutions to some fundamental problems is still unresolved, his call for modellers to provide realistic estimates of prediction uncertainties has resulted in substantial progress. Approaches to deal with model structure uncertainty are still not well developed and used, however.

What does the future hold for likely developments and challenges? First of all, we do not believe that the fundamental problems related to imperfect model structures and spatial heterogeneity at subgrid scale can be solved in a foreseeable future. The main reason for this is that we will probably never get enough data to adequately describe all details of the geological and ecological systems in our catchments. New measuring techniques are gradually improving the data situation, but we still need major breakthroughs with quantum leaps in terms of more and better quality field data before we can start hoping for a solution to the fundamental scientific problems. The logical consequence of this line of argument is that one distributed physically-based model code, such as SHE or any other program, will never be able to address all problems adequately. Consequently, the answer to the fundamental problems does not lie in further increasing the model structure complexity.

While we are waiting for the breakthroughs in new measurement techniques–and it may be a long waiting period–we need to use tools based on our present knowledge: the existing modelling codes and new future codes which we can expect to be marginally better than the existing ones.

More importantly, how do we then use the models? First of all, we need to ensure a good practice of the entire modelling process including sound scientific practices, stakeholder involvement and quality assurance (Refsgaard & Henriksen 2004; Refsgaard *et al.* 2005; Jakeman *et al.* 2006). Secondly, we need to broaden the uncertainty assessment methodologies so that model structure uncertainty is always considered. The alternative blueprint for physically-based distributed modelling proposed by Beven (2002*b*) is a pragmatic realistic way forward. He suggests that any study should *a priori* include a number of different contending model structures and within each model structure a variety of parameter sets. These combinations of model structures and parameter sets should then be confined by testing against field data, where the combinations that do not reach some agreed level of acceptability are discarded as non-behavioural. This strategy of using multiple model structures is also strongly advocated by Neuman & Wierenga (2003), Poeter & Anderson (2005) and Refsgaard *et al.* (2006). The implications of such a strategy would be a need among practitioners for model codes with alternative process formulations (e.g. Butts *et al.* 2004) and software platforms, including several model codes and uncertainty assessments tools, so that a modeller can easily construct a suite of models with alternative model structures and use these to assess prediction uncertainties.

ACKNOWLEDGEMENTS

Thanks to numerous present and former colleagues, research partners and opponents in the scientific disputes and discussions during the past three decades. The authors greatly appreciate the constructive review comments by K. Beven and H. Gupta.

REFERENCES

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connel, P. E. & Rasmussen, J. 1986*a* An introduction to the European hydrological system–système Hydrologique Européen "SHE",1: history and philosophy of a physically-based distributed modelling system. *J. Hydrol.* **87**, 45–59.

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connel, P. E. & Rasmussen, J. 1986*b* An introduction to the European hydrological system–système Hydrologique Européen "SHE", 2: structure of a physically-based distributed modelling system. *J. Hydrol.* **87**, 61–77.

Andersen, J., Refsgaard, J. C. & Jensen, K. H. 2001 Distributed hydrological modelling of the Senegal river basin–model construction and validation. *J. Hydrol.* **247**, 200–214.

Andersen, J., Sandholt, I., Jensen, K. H., Refsgaard, J. C. & Gupta, H. 2002 Perspectives in using a remotely sensed dryness index in distributed hydrological models at the river-basin scale. *Hydrol. Processes* **16**, 2973–2987.

Arnold, J. G., Allen, P. M. & Bernhardt, G. 1993 A comprehensive surface-groundwater flow model. *J. Hydrol.* **142**, 47–69.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large area hydrologic modelling and assessment: part I. Model development. *J. Am. Water Resour. Assoc.* **34**(1), 73–89.

Banta, E. R., Hill, M. C., Poeter, E., Doherty, J. E. & Babendreier, J. 2008 Building model analysis applications with the Joint Universal Parameter Identification and Evaluation of Reliability (JUPITER) API. *Comput. Geosci.* **34**, 310–319.

Bathurst, J. C. 1986 Physically-based distributed modelling of an upland catchment using the système Hydrologique Européen. *J. Hydrol.* **87**, 79–102.

Bathurst, J. C. & O'Connell, P. E. 1992 The future of distributed modelling–the système Hydrologique Européen. *Hydrol. Processes* **6**(3), 265–277.

Bathurst, J. C., Wicks, J. M. & O'Connell, P. E. 1995 The SHE/SHESED basin scale water flow and sediment transport modelling system. In: Singh, V. P. (ed.) *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado, pp. 563–594.

Beck, M. B. 1987 Water quality modelling: a review of the analysis of uncertainty. *Water Resour. Res.* **23**(8), 1393–1442.

Bergström, S. 1991 Principles and confidence in hydrological modelling. *Nord. Hydrology* **22**, 123–136.

Bergström, S. 1995 The HBV model. In: Singh, V. P. (ed.) *Computer Models of Watershed Hydrology*. Water Resources Publications. Highlands Ranch, Colorado, pp. 443–476.

Bergström, S. & Forsman, A. 1973 Development of a conceptual deterministic rainfall-runoff model. *Nord. Hydrol.* **4**, 147–170.

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

Beven, K. 1995 Linking parameters across scales: Subgrid parameterization and scale dependent hydrological models. *Hydrol. Processes* **9**, 507–525.

Beven, K. 1996*a* A discussion of distributed hydrological modelling. In: Abbott, M. B. & Refsgaard, J. C. (eds) *Distributed Hydrological Modelling*. Kluwer Academic Publishers, Dordrecht, pp. 255–278.

Beven, K. 1996*b* Response to comments on 'a discussion of distributed hydrological modelling'. In: Abbott, M. B. & Refsgaard, J. C. (eds) *Distributed Hydrological Modelling*. Kluwer Academic Publishers, Dordrecht, pp. 289–295.

Beven, K. 2002*a* Towards a coherent philosophy for modelling the environment. *Proc. Roy. Soc. Lond. A* **458**(2026), 2465–2484.

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

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

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

Beven, K., Warren, R. & Zaoui, J. 1980 SHE: towards a methodology for physically-based distributed forecasting in hydrology. *IAHS Publication* **129**, 133–137.

Beven, K. J., Kirkby, M. J., Schoffield, N. & Tagg, A. 1984 Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *J. Hydrol.* **69**, 119–143.

Beven, K. J., Lamb, R., Quinn, P., Romanowicz, R. & Freer, J. 1995 TOPMODEL. In: Singh, V. P. (ed.) *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado, pp. 627–668.

Binley, A. M., Beven, K. J., Calver, A. & Watts, L. G. 1991 Changing responses in hydrology: assessing the uncertainty in physically based model predictions. *Water Resour. Res.* **27**(6), 1253–1261.

Butts, M. B., Payne, J. T., Kristensen, M. & Madsen, H. 2004 An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow prediction. *J. Hydrol.* **298**, 242–266.

Christensen, S. 1994 Hydrological model for the TudeÅ catchment. *Nord. Hydrol.* **25**, 145–166.

Christiaens, K. & Feyen, J. 2002 Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrol. Processes* **16**(2), 373–391.

Crawford, N. H. & Linsley, R. K. 1966 *Digital simulation in hydrology, Stanford Watershed Model IV*, Department of Civil Engineering, Stanford University, Technical Report 39.

Ewen, J. & Parkin, G. 1996 Validation of catchment models for predicting land-use and climate change impacts. 1. Method. *J. Hydrol.* **175**, 583–594.

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

Goderniaux, P., Brouyère, S., Fowler, H. J., Blenkinsop, S., Therrien, R., Orban, P. & Dassargues, A. 2009 Large scale surface-subsurface hydrological model to assess climate change impacts on groundwater reserves. *J. Hydrol.* **373**, 122–138.

Graham, D. N. & Butts, M. B. 2005 Flexible integrated watershed modelling with MIKE SHE. In: Singh, V. P. & Frevert, D. K. (eds) *Watershed Models*. CRC Press, Boca Raton, Florida, pp. 245–272.

Grayson, R. B., Moore, I. D. & McHahon, T. A. 1992*a* Physically based hydrologic modelling, 1. A terrain-based model for investigative purposes. *Water Resour. Res.* **28**(10), 2639–2658.

Grayson, R. B., Moore, I. D. & McHahon, T. A. 1992*b* Physically based hydrologic modelling, 2. Is the concept realistic? *Water Resour. Res.* **28**(10), 2639–2658.

Gregersen, J. B., Gijsbers, P. J. A. & Westen, S. J. P. 2007 OpenMI: opean modelling interface. *J. Hydroinformatics* **9**(3), 175–191.

Hansen, J. R., Refsgaard, J. C., Hansen, S. & Ernstsen, V. 2007 Problems with heterogeneity in physically based agricultural catchment models. *J. Hydrol.* **342**(1–2), 1–16.

Hansen, J. R., Ernstsen, V., Refsgaard, J. C. & Hansen, S. 2008 Field scale heterogeneity of redox conditions in till–upscaling to a catchment nitrate model. *Hydrogeol. J.* **16**, 1251–1266.

Hansen, J. R., Refsgaard, J. C., Ernstsen, V., Hansen, S., Styczen, M. & Poulsen, R. N. 2009 An integrated and physically based nitrogen cycle catchment model. *Hydrol. Res.* **40**(4), 347–363.

Harbaugh, A. W. 2005 MODFLOW-2005, the US Geological Survey modular ground-water model–the Ground Water Flow Process. US Geological Survey Techniques and Methods 6-A16.

Havnø, K., Madsen, M. N. & Dørge, J. 1995 MIKE 11–a generalized river modelling package. In: Singh, V. P. (ed.) *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado, pp. 733–782.

Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C. & Madsen, B. 2003 Methodology for construction, calibration and validation of a national hydrological model for Denmark. *J. Hydrol.* **280**, 52–71.

Henriksen, H. J., Troldborg, L., Højberg, A. L. & Refsgaard, J. C. 2008 Assessment of exploitable groundwater resources of Denmark by use of ensemble ressource indicators and numerical groundwater-surface water model. *J. Hydrol.* **348**(1–2), 224–240.

Hill, M. C. 1998 *Methods and guidelines for effective model calibration*. U.S. Geological Survey, Water-Resources Investigations Report 98-4005. Denver CO.

Ibbitt, R. 1985 SHE–Model of the 80s. STREAMLAND 39, Ministry of Works and Developments, New Zealand.

IPCC 2001 In: Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K. & Johnson, C. A. (eds) *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel of Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA, p. 881.

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.

Jakeman, A. J., Letcher, R. A. & Norton, J. P. 2006 Ten iterative steps in development and evaluation of environmental models. *Environ. Modell. Softw.* **21**, 602–614.

Jayatilaka, C. J., Storm, B. & Mudgeway, L. B. 1998 Simulation of water flow on irrigation bay scale with MIKE-SHE. *J. Hydrol.* **208**, 108–130.

Jensen, K. H. & Mantoglou, A. 1992 Application of stochastic unsaturated flow theory, numerical simulations, and comparisons to field observations. *Water Resour. Res.* **28**, 269–284.

Jensen, K. H., Bitch, K. & Bjerg, P. L. 1993 Large scale dispersion experiments in a sandy aquifer in Denmark–observed tracer movements and numerical analyses. *Water Resour. Res.* **29**(3), 673–696.

Kaiser-Hill 2001 Model code and scenario selection report site: Wide water balance rocky flats environmental technology site. Report 01-RF-00337. Kaiser-Hill Company LLC.

Kim, N. W., Chung, I. M., Won, Y. S. & Arnold, J. G. 2008 Development and application of the integrated SWAT-MODFLOW model. *J. Hydrol.* **356**, 1–16.

Klemes, V. 1986 Operational testing of hydrological simulation models. *Hydrological Sci. J.* **31**, 13–24.

Knudsen, J., Thomsen, A. & Refsgaard, J. C. 1986 WATBAL: a semi-distributed, physically based hydrological modelling system. *Nord. Hydrol.* **17**, 347–362.

Kutchment, L. S., Demidov, V. N., Naden, P. S., Cooper, D. M. & Broadhurst, P. 1996 Rainfall-runoff modelling of the Ouse basin, North Yorkshire: an application of a physically based distributed model. *J. Hydrol.* **181**, 323–342.

Lane, S. A. & Richards, K. S. 2001 The ‘validation’ of hydrodynamic models: some critical perspectives. In: Anderson, M. G. & Bates, P. D. (eds) *Model Validation Perspectives in Hydrological Science*. John Wiley & Sons, Ltd, pp. 413–438.

Leavesley, G. H., Lichty, R. W., Troutman, B. M. & Saindon, L. G. 1983 *Precipitation-runoff modelling system*. User’s manual: US Geological Survey Water Resources Investigations Report 83-4238, p. 207.

Leavesley, G. H., Markstrom, S. L., Viger, R. J. & Hay, L. E. 2005 USGS modular modelling system (MMS)–precipitation-runoff modelling system (PRMS) MMS-PRMS. In: Singh, V. P. & Frevert, D. K. (eds) *Watershed Models*. CRC Press, pp. 159–177.

Loague, K. M. & Freeze, R. A. 1985 A comparison of rainfall-runoff modelling techniques on small upland catchments. *Water Resour. Res.* **21**(2), 229–248.

Lohani, V. K., Refsgaard, J. C., Clausen, T., Erlich, M. & Storm, B. 1993 Application of the SHE for irrigation command area studies in India. *J. Irrigation Drainage Eng.* **119**(1), 34–49.

Markstrom, S. L., Niswonger, R. G., Regan, R. S., Prudic, D. E. & Barlow, P. M. 2008 GSFLOW–Coupled ground-water and surface-water flow model based on the integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005). *US Geological Survey Techniques and Methods* 6-D1, p. 240.

McDonald, M. G. & Harbaugh, A. W. 1988 A modular three-dimensional finite difference ground-water flow model. *US Geological Survey Techniques of Water Resources Investigations*, Book 6, Chapter A-1, p. 586.

Michaud, J. & Sorooshian, S. 1994 Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resour. Res.* **30**(3), 593–605.

Miles, J. C. & Rushton, K. R. 1983 A coupled surface water and groundwater catchment model. *J. Hydrol.* **62**, 159–177.

Neuman, S. P. & Wierenga, P. J. 2003 *A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites*. University of Arizona, Report NUREG/CR-6805.

Nielsen, S. A. & Hansen, E. 1973 Numerical simulation of the rainfall-runoff process on a daily basis. *Nord. Hydrol.* **4**, 171–190.

Panday, S. & Hayakorn, P. S. 2004 A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Adv. Water Resour.* **27**(4), 361–382.

Pappenberger, F. & Beven, K. J. 2006 Ignorance is bliss: or seven reasons not to use uncertainty analysis. *Water Resour. Res.* **42**(5), W05302.

Parkin, G., O'Donnell, G., Ewen, J., Bathurst, J. C., O'Connell, P. E. & Lalabre, J. 1996 Validation of catchment models for predicting land-use and climate change impacts. 2. Case study for a Mediterranean catchment. *J. Hydrol.* **175**, 595–613.

Perkins, S. P. & Sophocleous, M. 1999 Development of a comprehensive watershed model applied to study stream yield under drought conditions. *Ground Water* **37**(3), 418–426.

Poeter, E. & Anderson, D. 2005 Multiple ranking and inference in ground water modeling. *Ground Water* **43**(4), 597–605.

Querner, E. P. 1997 Description and application of the combined surface and groundwater flow model MOGROW. *J. Hydrol.* **192**, 158–188.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F. & Seo, D.-J. 2004 Overall distributed model intercomparison project results. *J. Hydrol.* **298**, 27–60.

Refsgaard, J. C. 1997 Parametrisation, calibration and validation of distributed hydrological models. *J. Hydrol.* **198**, 69–97.

Refsgaard, J. C. & Hansen, E. 1982*a* A distributed groundwater/surface water model for the Suså catchment. Part 1: model description. *Nord. Hydrol.* **13**, 299–310.

Refsgaard, J. C. & Hansen, E. 1982*b* A distributed groundwater/surface water model for the Suså catchment. Part 2: simulations of streamflow depletions due to groundwater abstraction. *Nord. Hydrol.* **13**, 311–322.

Refsgaard, J. C. & Henriksen, H. J. 2004 Modelling guidelines–terminology and guiding principles. *Adv. Water Resour.* **27**, 71–82.

Refsgaard, J. C. & Knudsen, J. 1996 Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **32**(7), 2189–2202.

Refsgaard, J. C. & Storm, B. 1995 MIKE SHE. In: Singh, V. P. (ed.) *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado, pp. 809–846.

Refsgaard, J. C., Seth, S. M., Bathurst, J. C., Erlich, M., Storm, B., Jørgensen, G. H. & Chandra, S. 1992 Application of the SHE to catchments in India–part 1: general results. *J. Hydrol.* **140**, 1–23.

Refsgaard, J. C., Storm, B. & Abbott, M. B. 1996 Comments on 'a discussion of distributed hydrological modelling'. In: Abbott, M. B. & Refsgaard, J. C. (eds) *Distributed Hydrological Modelling*. Kluwer Academic Publishers, pp. 279–287.

Refsgaard, J. C., Sørensen, H. R., Mucha, I., Rodak, D., Hlavaty, Z., Bansky, L., Klucovska, J., Topolska, J., Takac, J., Kosc, V., Enggrob, H. G., Engesgaard, P., Jensen, J. K., Fiselier, J., Griffioen, J. & Hansen, S. 1998 An integrated model for the Danubian Lowland–methodology and applications. *Water Resour. Manage.* **12**, 433–465.

Refsgaard, J. C., Thorsen, M., Jensen, J. B., Kleeschulte, S. & Hansen, S. 1999 Large scale modelling of groundwater contamination from nitrogen leaching. *J. Hydrol.* **221**(3–4), 117–140.

Refsgaard, J. C., Henriksen, H. J., Harrar, W. G., Scholten, H. & Kassahun, A. 2005 Quality assurance in model based water management–review of existing practice and outline of new approaches. *Environ. Modell. Softw.* **20**, 1201–1215.

Refsgaard, J. C., van der Sluijs, J. P., Brown, J. & van der Keur, P. 2006 A framework for dealing with uncertainty due to model structure error. *Adv. Water Resour.* **29**, 1586–1597.

Rohdenburg, H. B., Diekkrüger, B. & Bork, H. R. 1986 Deterministic hydrological site and catchment models for the analysis of agro-ecosystems. *Catena* **13**, 119–137.

Rosbjerg, D. & Madsen, H. 2005 Concepts of hydrological modelling. In: Anderson, M. G. (ed.) *Encyclopedia of Hydrological Sciences*. John Wiley & Sons, Ltd.

Seyfried, M. S. & Wilcox, B. P. 1995 Scale and the nature of spatial variability: field examples having implications for hydrologic modelling. *Water Resour. Res.* **31**(1), 173–184.

Storm, B. 1991 Modelling of saturated flow and the coupling of the surface and subsurface flow. In: Bowles, D. S. & O'Connel, P. E. (eds) *Recent Advances in the Modelling of Hydrologic Systems*. Kluwer, Dordrecht, pp. 185–203.

Styczen, M. & Storm, B. 1993*a* Modelling of N-movements on catchment scale–a tool for analysis and decision making, 1. model description. *Fertilizer Res.* **36**, 1–6.

Styczen, M. & Storm, B. 1993*b* Modelling of N-movements on catchment scale–a tool for analysis and decision making, 2. a case study. *Fertilizer Res.* **36**, 7–17.

Therrien, R., McLaren, R., Sudicky, E. A. & Panday, S. 2006 *HydroGeo-Sphere, a Three-Dimensional Numerical Model Describing Fully- Integrated Subsurface and Surface Flow and Solute Transport*. Groundwater Simulation Group, Waterloo, Ont., Canada.

Thorsen, M., Refsgaard, J. C., Hansen, S., Pebesma, E., Jensen, J. B. & Kleeschulte, S. 2001 Assessment of uncertainty in simulation of nitrate leaching to aquifers at catchment scale. *J. Hydrol.* **242**, 210–227.

Trefry, M. G. & Muffels, C. 2007 FEFLOW: a finite-element ground water flow and transport modeling tool. *Groundwater J.* **45**(5), 525–528.

Troch, P. A., Mancini, M., Paniconi, C. & Wood, E. F. 1993 Evaluation of a distributed catchment scale water balance model. *Water Resour. Res.* **29**(6), 1805–1817.

VanderKwaak, J. E. & Loague, K. 2001 Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. *Water Resour. Res.* **37**(4), 999–1013.

Wardlaw, R. B., Wyness, A. & Rippon, P. 1994 Integrated catchment modelling. *Surv. Geophys.* **15**, 311–330.

WMO 1975 *Intercomparison of conceptual models used in operational hydrological forecasting*. WMO Operational Hydrology Report No 7, WMO No 429, World Meteorological Organisation, Geneva.

WMO 1988 *Intercomparison of models for snowmelt runoff*. WMO Operational Hydrology Report No 23, WMO No 646, World Meteorological Organisation, Geneva.

Yan, J., Hopkins, E., Sørensen, H. R. & Kjelds, J. T. 1998 Integrated hydrological modelling in South Florida Water Management District. *Water Resour. Eng.* **98**(1), 532–537.

First received 23 September 2009; accepted in revised form 1 December 2009. Available online June 2010