

---

# Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment

Zhiyu Liu<sup>1</sup>, Mario L.V. Martina<sup>2</sup> and Ezio Todini<sup>2</sup>

<sup>1</sup>Bureau of Hydrology, Ministry of Water Resources, 2 Lane 2, Baiguang Road, Beijing 100053, P.R. China

<sup>2</sup>Department of Geo-Environmental Sciences, University of Bologna, Via Zamboni 67, 40126, Bologna, Italy

Email for corresponding author: liuzy@mwr.gov.cn

---

## Abstract

TOPKAPI is a physically-based, fully distributed hydrological model with a simple and parsimonious parameterisation. The original TOPKAPI is structured around five modules that represent evapotranspiration, snowmelt, soil water, surface water and channel water, respectively. Percolation to deep soil layers was ignored in the old version of the TOPKAPI model since it was not important in the basins to which the model was originally applied. Based on published literature, this study developed a new version of the TOPKAPI model, in which the new modules of interception, infiltration, percolation, groundwater flow and lake/reservoir routing are included. This paper presents an application study that makes a first attempt to derive information from public domains through the internet on the topography, soil and land use types for a case study Chinese catchment – the Upper Xixian catchment in Huaihe River with an area of about 10 000 km<sup>2</sup>, and apply a new version of TOPKAPI to the catchment for flood simulation. A model parameter value adjustment was performed using six months of the 1998 dataset. Calibration did not use a curve fitting process, but was chiefly based upon moderate variations of parameter values from those estimated on physical grounds, as is common in traditional calibration. The hydrometeorological dataset of 2002 was then used to validate the model, both against the outlet discharge as well as at an internal gauging station. Finally, to complete the model performance analysis, parameter uncertainty and its effects on predictive uncertainty were also assessed by estimating a posterior parameter probability density via Bayesian inference.

**Keywords:** flood forecasting, TOPKAPI model, upper Xixian catchment

## Introduction

An increasing demand for mathematical models which can predict the hydrological effects of man-made changes to land use and the impact of climate change on water resources, has stimulated interest in physically-based, catchment hydrological models. With the recent advances in remote sensing, geographic information systems and computer technology, together with natural evolution in hydrological modelling, physically-based distributed or semi-distributed hydrological models with simple and parsimonious parameterisations, such as DHSVM (Wigmostra *et al.*, 1994), SLURP (Kite, 1995), *FEWS NET Stream Flow Model* (Artan *et al.*, 2001), TOPKAPI model (Todini and Ciarapica, 2001; Liu, 2002; Liu and Todini, 2002) have been developed in recent years, which are attractive for flow simulation and prediction, both in applied research and in the operational field.

TOPKAPI is a physically based, fully-distributed rainfall-runoff model, which is based on the lumping of a kinematic wave assumption in the soil, on the surface and in the drainage network, and leads to transforming the rainfall-runoff and runoff routing processes into three non-linear reservoir differential equations. Initially, the TOPKAPI model (Todini and Ciarapica, 2001; Liu and Todini, 2002) was structured around five modules that represent the evapotranspiration, snowmelt, soil water, surface water and channel water respectively. Percolation to deep soil layers was ignored since it was not important in the basins to which the model was applied originally (Todini and Ciarapica, 2001).

This study developed a new version of the TOPKAPI model, in which the new modules of interception, infiltration, percolation and groundwater flow are included. Percolation to deep soil layers is simulated through the

introduction of a second soil layer with different characteristics from the upper layer, and involving water movement in a vertical direction feeding into the aquifer. In addition, an approach for incorporating lakes and reservoirs is also included.

This paper reports the application of the new version of TOPKAPI (Liu and Todini, 2002) to the upper Xixian catchment about 10 000 km<sup>2</sup> in area located in eastern China, using public domain data sets of topography, soil and land use types readily available over the internet. Model calibration used six months of the 1998 dataset. No calibration in the common sense of a curve fitting process was done. Using the hydrometeorological dataset of 2002, the model was validated against the discharge not only at the outlet but also at an internal gauging station. A description of the structure and methodology, parameters and data requirements of the new TOPKAPI, with the emphasis on illustrating the new components of the model, is also presented.

## The new version of TOPKAPI model

### STRUCTURE AND METHODOLOGY

The TOPKAPI model is a grid-based simulation approach. Spatial distribution of catchment parameters, precipitation input and hydrological response is described in the horizontal by a lattice (the grid cells of a Digital Elevation (DEM)) and in the vertical by a column of horizontal layers at each grid square.

In the Soil Vegetation Atmosphere Transfer (SVAT) simulation, a grid cell is represented by a maximum of five

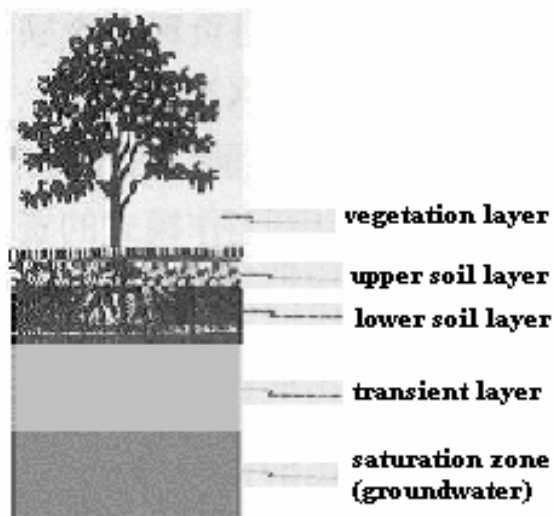


Fig.1. Sketch of five horizontal layers in a grid square in a SVAT simulation

layers (Fig. 1). The vegetation layer on the top is the major link between atmosphere and soil. The soil below has three horizons: Horizon 1 (upper soil layer, 0–30 cm depth) accounts for infiltration and evaporation processes; Horizon 2 (lower soil layer, 30–150 cm depth in general, up to 10m depth) mainly determines transpiration and percolation into deep soil layers; Horizon 3 is the transient transport zone in which percolation is transformed to groundwater recharge. The lowest layer of the column is the upper groundwater aquifer, if present (Framework project, 2000).

Taking account of the fact that the soil characteristics (depth, permeability) for the up to four soil layers are not always available, a two-layer soil model is proposed in the new TOPKAPI simulation, i.e. the soil column is divided into the upper soil layer and the lower soil layer (Fig. 2). The upper soil layer can be regarded as the ‘soil layer’ in the old TOPKAPI model (and the total of the upper and lower soil layers in a SVAT simulation as well); it is a soil water layer of limited thickness and with generally high hydraulic conductivity, which contributes to surface runoff if its soil moisture content exceeds its saturation level and loses water by drainage and percolation to the lower soil layer if its moisture content exceeds its field capacity. The lower soil layer can be thought of as the total of the transient transport layer and the groundwater-saturated zone in a SVAT simulation, it is a sub-soil layer below the upper soil layer and above the imperious bedrock layer, with lower hydraulic conductivity. The lower soil layer is introduced in the new TOPKAPI to account for percolation to deep soil layers and groundwater flow. Local saturated zones can exist as perched groundwater above the boundary between the upper soil layer and the lower soil layer.

The model is based on the idea of combining the kinematic approach with the topography of the basin described by a DEM. The flow paths and slopes are evaluated starting from the DEM, according to a neighbourhood relationship based on the principle of minimal energy cost (Band, 1986). Each grid cell of the DEM is assigned a value for each of the physical characteristics represented in the model. Precipitation is assumed constant over the single cell, by means of area-distribution techniques, such as Thiessen polygons and Block Kriging (de Marsily, 1986). The integration in space of the kinematic wave equations results in three ‘structurally-similar’ non-linear reservoir equations describing interflow, surface flow and channel flow (Liu, 2002).

The new TOPKAPI model is constructed around ten components: interception, evapotranspiration, snowmelt, infiltration, interflow and percolation, vertical recharge to groundwater table, groundwater flow, surface flow, channel flow as well as lake/reservoir routing. All the components may be activated on an individual grid cell of the DEM.

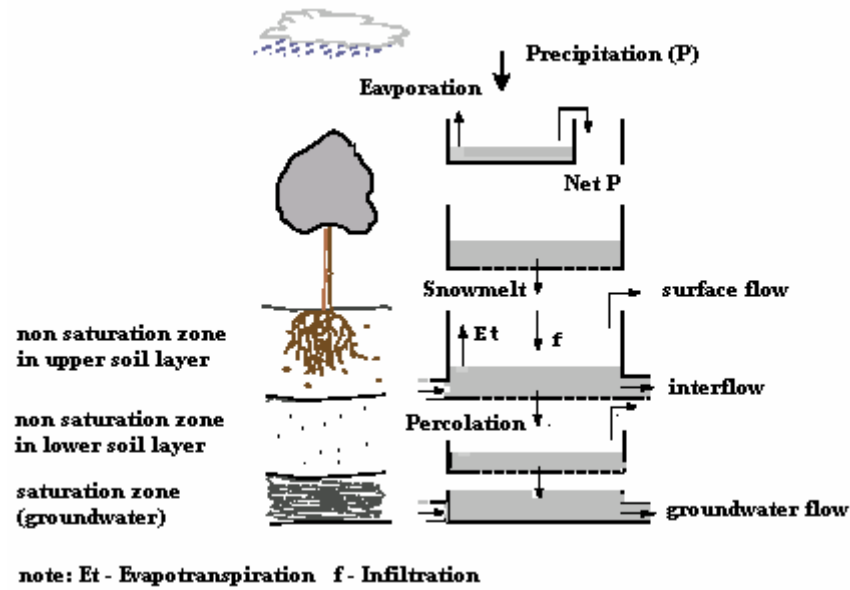


Fig. 2. Water balance simulation in the new TOPKAPI model

#### INTERCEPTION COMPONENT

Interception includes rainfall interception and snowfall interception. For snowfall interception, the new TOPKAPI uses equivalent water depth to represent the snowpack on the vegetation instead of actual snow depth.

A simple empirical equation such as Eqn. (1) can be used to determine the interception capacity of particular vegetation from the information about the vegetation's crop density and leaf-area-index (LAI) for each species and for each month of the year (Chen, 1996).

$$S_r(t) = S_{r_0} \cdot Cropd(t) \frac{LAI(t)}{LAI_0} \quad (1)$$

where  $S_r(t)$  = vegetation interception capacity at time  $t$ ;  $S_{r_0}$  = maximum vegetation interception capacity in a year;  $LAI(t)$  = ratio of total leaf area to area of ground covered by vegetation — the leaf-area-index of the vegetation at time  $t$ ;  $LAI_0$  = maximum LAI for the vegetation in a year;  $Cropd(t)$  = proportion of ground in planview hidden by vegetation – the crop density in the cell at time  $t$ .

Actual vegetation interception during a simulation is determined by the amount of precipitation and the deficit in the vegetation water storage, which is shown in Eqn. (2).

$$S_{r_a}(t) = \begin{cases} \min[PE(t), SC_d(t)] & S_r(t) > SC_d(t), \quad S_r(t) > PE(t) \\ SC_d(t) & S_r(t) > SC_d(t), \quad S_r(t) \leq PE(t) \\ PE(t) & S_r(t) \leq SC_d(t), \quad S_r(t) > PE(t) \\ \min[PE(t), SC_d(t)] & S_r(t) \leq SC_d(t), \quad S_r(t) \geq PE(t) \end{cases} \quad (2)$$

with  $SC_d(t) = [SC_0 Cropd(t) LAI(t) / LAI_0 - SC(t-T)]$  and  $PE(t) = P(t) - Ep(t)$  where  $T$  = the computation time interval;  $S_{r_a}(t)$  = actual vegetation interception at time  $t$ ;  $SC_0$  = vegetation storage capacity in a year;  $SC(t-T)$  = vegetation water storage at time  $t-T$ ;  $SC_d(t)$  = deficit of the vegetation water storage at time  $t$ ;  $PE(t)$  = net precipitation on the vegetation;  $P(t)$  = precipitation;  $Ep(t)$  = potential evaporation from the vegetation.

#### EVAPOTRANSPIRATION COMPONENT

Evapotranspiration can either be introduced directly as an input to the model by computing it externally or estimated internally by a simplified equation derived from the radiation method (Doorenbos *et al.*, 1984), based on the air temperature and on topographic, geographic and climatic information, which was developed for the ARNO model (Todini, 1996). Different capacities of evapo-transpiration from varying land uses are basically affected by transpiration and evaporation from water intercepted by the given vegetation. For different types of land use, monthly crop coefficients (Doorembos, 1984; Maidment, 1993) are given, reflecting the state of the plants in an annual growth cycle. Thus, the potential evapotranspiration for actual land use is calculated by Eqn. (3).

$$Et_p = Et_0 Kc_{crop} \quad (3)$$

where  $Et_p$  is potential evapotranspiration for actual land use;  $Et_0$  is potential evapotranspiration for a reference land cover -lawn;  $Kc_{crop}$  is the ratio between the crop potential

evapotranspiration rate and the corresponding reference evapotranspiration rate.

Doorenbos and Pruitt (1992) give a comprehensive database of FAO crop coefficients ( $K_c$ ) for different climatic conditions and phenological stages (initial, mid-season and late-season stages). They stressed the need to collect local data on growing season and rate of crop development. Reduction of potential evaporation into actual evaporation is a linear process depending on the soil moisture of the upper soil layer.

#### SNOWMELT COMPONENT

For reasons of limited data availability, the snow accumulation and melting (snowmelt) component is driven by a radiation estimate based upon the air temperature measurements, which is also borrowed from the ARNO model (Todini, 1996). The following steps, similar to those adopted in SHE (Abbott *et al.*, 1986a,b), are followed:

- estimation of radiation at the DEM grid (in the distributed model) or at the snow line (in the lumped model);
- decision whether the precipitation is solid or liquid;
- estimation of the water mass and energy budgets based on the hypothesis of zero snowmelt;
- comparison of the total available energy with that sustained as ice by the total available mass at 273 °K;
- computation of the snowmelt produced by the excess energy; and updating the water mass and energy budgets.

#### INFILTRATION

The infiltration capacity depends on the land cover property and the soil moisture condition. By comparing infiltration capacity and the available surface water, whichever is smaller, the actual infiltration amount is calculated. In the new TOPKAPI, only the effect of land cover on the infiltration is accounted for. The amount of precipitation infiltrating into the soil ( $f_a$ ) is estimated using Eqn. (4):

$$f_a = P_n \cdot K_l \quad (4)$$

where  $P_n$  = net precipitation,  $K_l$  = coefficient depending on the land cover type.

#### INTERFLOW (SUB-SURFACE FLOW) AND PERCOLATION

The sub-surface flow at a point in the soil can be approximated by means of a kinematic wave model. The

point assumption is integrated up to a finite dimension in a generic cell, thus converting the original differential equation into a non-linear reservoir differential equation (Eqn. (5)) based upon physically meaningful parameters, which can be solved numerically or analytically from the known initial and boundary condition for the cell (Liu and Todini, 2002). The averaged interflow from the cell during the computation time interval can be calculated from the respective water balance calculation in the cell. Subsequently, at each time step the saturation excess volume (surface runoff) can be obtained by calculating the soil water balance (Liu and Todini, 2002).

$$\frac{dv_1}{dt} = \left[ (f_a X + q_o^u + q_s^u) - f_b X \right] - \frac{C_1}{X^{\alpha_s}} v_1^{\alpha_s} \quad (5)$$

$$\text{with } C_1 = \frac{L k_{sh_1} \tan(\beta)}{\mathcal{G}_e^{\alpha_s} L^{\alpha_s}} \text{ and } \mathcal{G}_e = \mathcal{G}_s - \mathcal{G}_r$$

where  $v_1$  is the upper soil water volume stored in the cell in  $m^3$ ,  $X$  is the grid cell size in  $m$ ,  $f_a$  is the infiltration into the cell in  $m s^{-1}$ ,  $f_b$  is the percolation rate to the lower soil layer in  $m s^{-1}$ ,  $q_o^u$  is the discharge entering the active cell as overland flow from the upstream contributing area ( $m^2 s^{-1}$ ), and  $q_s^u$  is the discharge entering the active cell as subsurface flow from the upstream contributing area in  $m^2 s^{-1}$ ,  $L$  is the upper soil depth in  $m$ ,  $k_{sh_1}$  is the horizontal soil saturated hydraulic conductivity in the upper soil layer in  $m s^{-1}$ ,  $b$  is the land surface slope,  $\mathcal{G}_s$  is the saturated soil moisture content,  $\mathcal{G}_r$  is the residual soil moisture content,  $\mathcal{G}_e$  is the effective soil moisture content and  $\alpha_s$  is the exponent of the transmissivity law for the upper soil.

It is assumed that percolation starts if the soil moisture content of the upper soil layer exceeds its field capacity ( $\mathcal{G}_f$ ). The percolation rate from the upper soil layer is assumed to increase as a function of the soil water content according to an experimentally determined power law (Clapp and Hornberger, 1978) but not to exceed the saturated soil hydraulic conductivity in the underlying deeper layer:

$$f_b = \min \left[ k_{sv_1} \left( \frac{v_1}{v_{1m}} \right)^{\alpha_p}, k_{sv_2} \right] \quad (6)$$

$$\text{with } v_{1m} = (\mathcal{G}_s - \mathcal{G}_r) L X$$

where,  $v_{1m}$  is the saturated water volume in the upper soil layer;  $k_{sv_1}$  is the vertical soil saturated hydraulic conductivity in the upper soil layer;  $k_{sv_2}$  is the vertical soil saturated hydraulic conductivity in the lower soil layer; the exponent  $\alpha_p$  depends on the type of the soil and can vary from  $\alpha_p \cong 11$  for sand to  $\alpha_p \cong 25$  for clay.

VERTICAL RECHARGE TO THE GROUNDWATER

TABLE

The unsaturated zone in the lower soil layer can be identified as an intermediate region, where moisture level remains constant at the field capacity of the soil and rock of the region. It is thought of as the transient transport zone, which transforms percolation to groundwater recharge if the soil moisture content of the upper soil layer exceeds its field capacity. In the new TOPKAPI, the vertical recharge to the groundwater table in a generic cell during a time step,  $q_r$ , is computed by Eqn. (7):

$$q_r = f_b X \left( \frac{v_2 + v_3}{v_{3m}} \right)^{\alpha_r} \quad (7)$$

with  $v_{3m} = (d - L) \rho X$

where  $v_2$  is the water volume per unit width stored in the unsaturated zone of the lower soil layer in the cell;  $v_3$  is the saturated groundwater volume per unit width stored in the cell;  $v_{3m}$  is the maximum groundwater volume per unit width stored in the lower soil layer,  $d$  is the impermeable bedrock depth from the surface,  $r$  is the effective porosity of the lower soil and  $\alpha_r$  is the soil parameter related with the lower soil property.

GROUNDWATER FLOW

The groundwater flow is generated from the groundwater-saturated zone, which is the major source of channel flow during the non-flood season. Assuming that the slope of the bedrock layer in the flow direction is  $s_b$ , Darcy's law states that the groundwater flow,  $q_g$ , can be obtained as:

$$q_g = k_{sh_2} \rho h \frac{\partial(h + z_f)}{\partial x} = k_{sh_2} \left( \frac{\partial h}{\partial x} + s_b \right) \left( \frac{v_3}{X} \right) \quad (8)$$

where  $x$  is the main direction of flow along a cell,  $h$  is the groundwater-table depth,  $z_f$  is the bedrock elevation relative to an absolute coordinate system,  $k_{sh_2}$  is the horizontal soil saturated hydraulic conductivity in the lower soil layer.

The variation through time of the groundwater volume in each grid cell is modelled by a linear reservoir equation. This combines Darcy's law and the mass conservation of one-dimensional laminar flow in an isotropic, homogeneous aquifer, to give:

$$\frac{dv_3}{dt} = (q_r + q_h) - k_{sh_2} \frac{\partial(h + z_f)}{\partial x} \left( \frac{v_3}{X} \right) \quad (9)$$

where  $q_h$  is the lateral groundwater flow into the cell.

Since  $\frac{\partial h}{\partial x}$  is much smaller than  $\frac{\partial z_f}{\partial x} = s_b$ , Eqn. (9) can be thus approximated into Eqn. (10) as:

$$\frac{dv_3}{dt} = (q_r + q_h) - \frac{k_{sh_2} s_b}{X} v_3 \quad (10)$$

In general, the term  $\frac{\partial h}{\partial x}$  has time variations that can be appreciated on the order of days. Since the integration time step in the model is kept in the order of hours, the variation in time of  $\frac{\partial h}{\partial x}$  is very small and, with a negligible error, the value estimated at the end of the previous time step can be used.

The solution of Eqn. (9) can be obtained analytically. The average groundwater flow from the cell during the computation time interval can be calculated on the respective water balance calculation in the cell.

Note that the groundwater may ex-filtrate into the upper soil layer as an input to the upper soil water reservoir, and it may also further ex-filtrate upward to feed overland flow at the ground surface; therefore, the water storages in the lower and the upper soil layers should be adjusted at the end of each computation of groundwater flow, which is done by means of the total water balance calculation in the soil column.

SURFACE FLOW (OVERLAND FLOW) COMPONENT

Overland flow routing is described similarly to the soil component, according to the kinematic approach (Wooding, 1965) in which the momentum equation is approximated by means of Manning's formula. By analogy, a similar process was effected for the soil; assuming that the surface water depth is constant over the cell and integrating the kinematic equation over the longitudinal dimension gives the non-linear reservoir model for the overland flow for a generic cell as:

$$\frac{dV_o}{dt} = r_o X^2 - \frac{C_o X}{X^{2\alpha_o}} V_o^{\alpha_o} \quad (11)$$

where  $V_o$  is the surface water volume in the cell in  $m^3$ ,  $r_o$  is the infiltration-excess and saturation-excess resulting from the solution of the soil water balance in  $[m s^{-1}]$ ,  $n_o$  is the Manning friction coefficient for the surface roughness in  $[m^{-1/3} s]$ ,  $C_o = (\tan \beta)^{1/2} / n_o$  is the coefficient relevant to the Manning formula for overland flow, and  $\alpha_o = 5/3$  is the exponent derived from using the Manning formula.

CHANNEL FLOW COMPONENT

Similar considerations also apply to the channel network, which is assumed to be tree-shaped with reaches having wide rectangular cross-sections. In this case the channel surface width is not constant but is assumed to increase towards the catchment outlet. On these assumptions, the



following non-linear reservoir model for the channel flow can be written for a generic reach:

$$\frac{dV_c}{dt} = (r_c XW + Q_c^u) - \frac{C_c W}{(XW)^{\alpha_c}} V_c^{\alpha_c} \quad (12)$$

where  $V_c$  is the volume of water stored in a generic channel reach in  $m^3$ ,  $W$  is the width of the rectangular channel reach in m, which is taken to increase as a function of the area drained by the cell on the basis of geomorphological considerations (Liu and Todini, 2002),  $Q_c^u$  is the inflow discharge from the upstream reaches in  $[m^3 s^{-1}]$ ,  $r_c$  is the lateral drainage input, including the overland runoff reaching the channel reach and the soil drainage reaching the channel reach in  $m s^{-1}$ ,  $C_c = s_0^{1/2}/n_c$  is the coefficient relevant to the Manning formula for channel flow,  $s_0$  is the river bed slope, assumed to be equal to the ground surface slope,  $n_c$  is the Manning friction coefficient for the channel roughness in  $[m^{-1/3} s]$  and  $\alpha_c = 5/3$  is the exponent derived from using the Manning formula.

#### LAKE/RESERVOIR ROUTING

In rainfall–runoff modelling, a lake or a reservoir can be thought of as a ‘multiple-input and single-output’ system. In the new TOPKAPI approach, a mechanism has been developed for identifying the receiving cells and the draining cells of the lake or the reservoir in a distributed model.

A reservoir in which the discharge is a function of water-surface elevation offers the simplest of all routing situations. Such a reservoir may have un-gated sluiceways and/or an uncontrolled spillway. Given the data of the elevation-storage curve and the elevation-discharge curve, the lake/reservoir routing can be implemented by using Eqn. (13):

$$\bar{Q}_i - Q_{o_1} + \left( \frac{V_1}{t} + \frac{Q_{o_1}}{2} \right) = \frac{V_2}{t} + \frac{Q_{o_2}}{2} \quad (13)$$

where  $\bar{Q}_i$  is the average inflow discharge into the reservoir during the time period  $t_1 \sim t_2$ ;  $Q_{o_1}, Q_{o_2}$  are the reservoir release discharges at time  $t_1, t_2$ , respectively;  $V_1, V_2$  are the reservoir storages at time  $t_1, t_2$ , respectively.

Solution of Eqn. (13) requires a routing curve showing  $V/t + Q_o/2$  versus  $Q_o$ . All terms on the left-hand side of the equation are known, and a value of  $V_2/t + Q_{o_2}/2$  can be computed. The corresponding value of  $Q_{o_2}$  can be determined from the routing curve. The computation is then repeated for succeeding routing periods.

#### DATA REQUIREMENTS AND PARAMETERS

The data required for the new TOPKAPI model include terrain data (e.g. DTM or DEM data, land survey data), soil survey data and vegetation or land-use data, LAI data,

geographical co-ordinates, precipitation data, evapo-transpiration or air temperature data.

Principal parameters required for each grid cell in the new TOPKAPI model are shown in Table 1. Because of their physical significance, in principle, the parameters can be measured directly through fieldwork. However, to do this in such a way as to provide coverage of the entire catchment would involve a prohibitive amount of work. Therefore, application of models such as TOPKAPI is likely to involve the transfer of parameter values measured at representative sites and the general use of parameter values available in the literature. Although it is physically-based, the model still needs calibration because of the uncertainty of the information on the topography, soil characteristics and land cover. Nonetheless, calibration of the TOPKAPI parameters is more an adjustment rather than a conventional calibration and is carried out by simple trial-and-error.

## The Upper Xixian catchment case study

#### CATCHMENT DESCRIPTION

The Upper Xixian catchment in China (Fig. 3), with an area of 10 100 km<sup>2</sup>, lies in the upper part of the Huaihe River which flows east to west to the Xixian station. It has three major tributaries: River Qingshui on the left side and River Shi and River Zhugan on the right side, on which there are two medium-sized reservoirs (Fig. 3). The watercourse of the Huaihe River within the catchment is 250 km long.

The local climate of the Eastern Asia monsoon results in the highest flooding risk in the period from June through August. The annual precipitation is around 1060 mm, and the yearly-averaged air temperature is around 15 °C in the catchment.

#### AVAILABLE GEOMORPHOLOGICAL AND LAND USE DATA

##### Topography

The first step in applying the new TOPKAPI model is to derive information on the topography of the catchment. For the Upper Xixian catchment, the United States Geological Survey (Gesch *et al.*, 1999) GTOPO30 public domain DEM from the Internet (<http://edcdaac.usgs.gov/gtopo30/gtopo30.html>, 2002) was used as the basic data source to define the catchment and its slopes and to derive the river network in the model by a GRASS-GIS (Todini *et al.*, 1999). Figure 4 shows the DEM on a grid scale of 1000 m with the derived river network by setting a drained-area threshold of 5 km<sup>2</sup> for the Upper Xixian catchment.

Table 1. Principal parameters required for each grid cell in the new TOPKAPI model

Model Component	Parameter
Interception	Vegetation interception capacity ( $S_{r0}$ ) Vegetation storage capacity ( $SC_0$ ) Maximum leaf-area-index ( $LAI_0$ )
Evapotranspiration	Daily maximum sunshine hours ( $n$ ) Monthly average air temperature ( $T_a$ ) Crop factors ( $K_c$ )
Snowmelt	Critical air temperature for determining the precipitation as snow or rain, $T_s$
Infiltration	Infiltration coefficient depending on the land cover type ( $K_I$ )
Interflow and percolation	Thickness of the upper soil layer ( $L$ ) Horizontal and vertical saturated hydraulic conductivity of the upper soil ( $k_{sh}$ and $k_{sv1}$ ) Effective soil moisture content of the upper soil ( $\theta_c = \theta_s - \theta_r$ ) Field capacity ( $\theta_p$ ) of the upper soil Exponent of the transmissivity law for the upper soil ( $a_s$ ) Exponent of the percolation law for the upper soil ( $a_p$ ) Vertical saturated hydraulic conductivity of the lower soil ( $k_{sv2}$ )
Groundwater recharge	Exponent of the vertical groundwater recharge equation for the lower soil, $a_r$
Groundwater flow	Horizontal saturated hydraulic conductivity of the lower soil ( $k_{sh2}$ ) Impermeable bedrock depth from the surface ( $d$ ) and slope ( $S_b$ ) Field capacity ( $\theta_p$ ) of the lower soil
Surface flow	Surface roughness ( $n_0$ ) Surface slope, tang ( $b$ )
Channel flow	Roughness for the channel according to the Strahler channel order ( $n_c$ ) Maximum and Minimum Channel widths ( $W_{max}$ , $W_{min}$ ) River bed slope ( $S_0$ )
Lake/reservoir routing	Routing curve

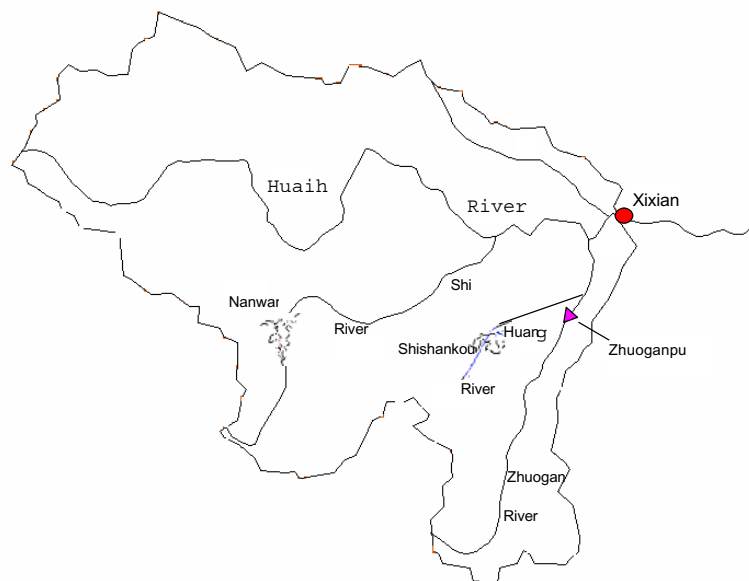


Fig. 3. The Upper Xixian catchment in Huaihe River basin

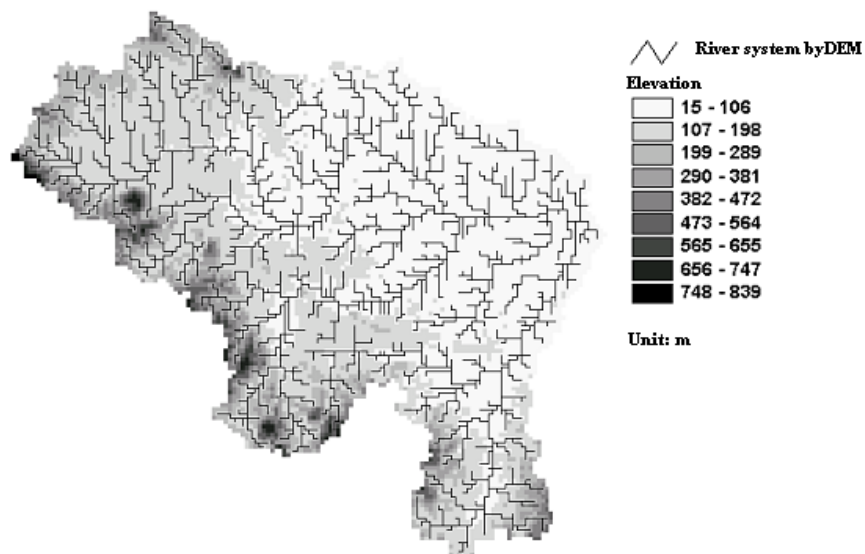


Fig. 4. DEM map and derived river network of the Upper Xixian catchment

*Soils*

Soil texture data were derived from the global soils dataset of Reynolds *et al.* (1999), which is available on line at <http://www.ngdc.noaa.gov/seg/eco/cdroms/reynolds/reynolds/reynolds.htm>, 2002. The dataset includes percentages of sand, silt and clay, and porosity, and it is based on the FAO Soil Map of the World linked to a global dataset of over 1300 soil pedons (FAO, 1996). The spatial resolution of the Reynolds map is 5 minutes (around 10 km) and there are two layers: 0–30 cm and 30–100 cm. For the Upper Xixian catchment, the soil map was classified into the four soil texture classes shown in Fig. 5.

*Land use*

The TOPKAPI model needs information on the distribution of the various land covers in the basin. For the Upper Xixian catchment, the UMD 1KM digital land cover map of the world (Hansen *et al.*, 2000) was used from the Internet (<http://www.geog.umd.edu/landcover/1km-map.html>, 2002). This was classified to the ten land classes shown in Fig. 6.

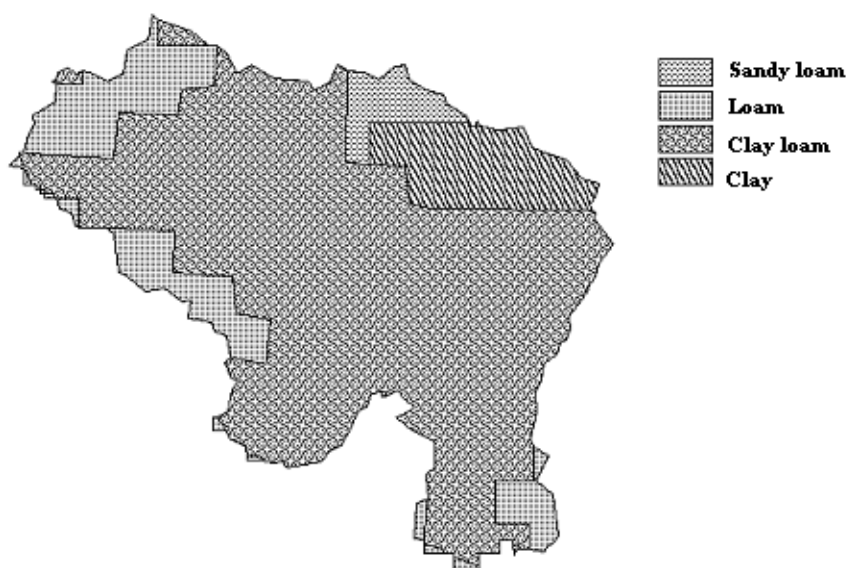


Fig. 5. Soil type map of the Upper Xixian catchment



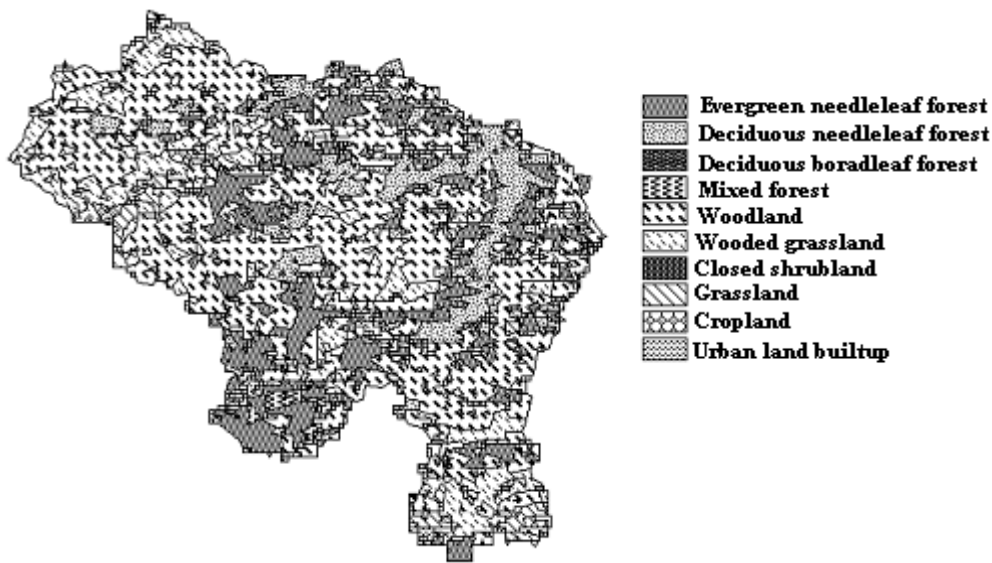


Fig. 6. Land cover map of the Upper Xixian catchment

## Calibration and validation of the new TOPKAPI

### HYDROMETEOROLOGICAL DATA AVAILABILITY

Floods have been experienced in the Upper Xixian catchment in recent years, 1991, 1998 and 2002. Six-hourly measurements from 24 raingages and 1 evaporation station are available while discharges are computed by means of a rating curve at Xixian. The hydrometeorological dataset for 1998 was selected for model calibration, since a major flood event with a peak discharge of  $4780 \text{ m}^3 \text{ s}^{-1}$  occurred in early August of that year. For model validation, the dataset of 2002 was chosen.

### TRADITIONAL CALIBRATION AND VALIDATION

The model calibration was performed at a 6-hour time-step using the hydrometeorological dataset of May 1 to Oct. 31, 1998. Due to the shortage of LAI data for the catchment, it was decided not to consider the interception component in the flood simulation. According to Li (2001), the effects of the two middle-size reservoirs on floods are sufficiently small to be neglected in flood simulation.

The initial soil saturation percentage was set as the same value of 0.9 for all cells at the beginning of the calibration period and it was assumed that there is no snow or surface water over the slopes and the water depth in a generic channel cell increases linearly with the channel width. The minimum and maximum width for the river network was set at 1 m and 400 m respectively. Since the elevations of impervious bedrock are unknown, it was assumed to be 30 m

from the surface for all the cells in the catchment. The distribution of the initial groundwater depth was estimated from the topographic index (Beven and Kirkby, 1979; Franchini *et al.*, 1996), given the catchment average groundwater depth of 5 m.

The initial model parameter values were estimated from the literature. Certain soil parameters (e.g. residual soil moisture content, saturated soil moisture content, field capacity, saturated soil hydraulic conductivity) were obtained from the USDA soil texture class index by referring to the USDA parameters table for the infiltration model of Green-Ampt (Maidment, 1993), and a soil parameter table used in the VIC (Variable Infiltration Capacity) model (Liang, *et al.*, 1996 a,b), which is available online at (<http://www.hydro.washington.edu/lettenmaier/models/vic/operation/info/soiltext.htm>, 2002). The Manning's overland flow roughness for different land cover types was estimated from Bedient and Huber (1992) and HEC-1 (1998), and the roughness of channel flow was estimated by referring to Chow (1959), Barnes (1967) and Maidment (1993) according to the Strahler channel order (Strahler, 1957). The final parameter values shown in Tables 2 and 3 were obtained by 'trial and error' on the basis of curve fitting. Figure 7 shows the comparison of the observed discharges and those calculated by the new TOPKAPI model at Xixian.

The new TOPKAPI is described as a physically-based, spatially distributed model. The model should be validated against internal data as well as the catchment outlet response, e.g. discharge at internal gauging stations or well levels (Ambroise *et al.*, 1995; Feyen *et al.*, 2000; Vázquez *et al.*, 2002).

Table 2. Principal calibrated soil parameters in the new TOPKAPI model for the Upper Xixian catchment

Soil type	$J_s - J_r$	$J_f - J_r$	$a_s$	$a_p$	$k_{sh}$ (m s <sup>-1</sup> )	$k_{sv1}$ (m s <sup>-1</sup> )	$L$ (m)
Sandy loam	0.412	0.172	2.5	-	1.46E-04	7.28E-07	0.80
Loam	0.433	0.264	2.5	-	5.47E-05	2.74E-07	0.50
Clay loam	0.432	0.312	2.5	3.0	4.92E-05	2.46E-07	0.30
Clay	0.385	0.275	2.5	-	8.83E-05	4.42E-07	0.20

Table 3. Principal calibrated soil parameters in the new TOPKAPI model for the Upper Xixian catchment

Land cover	$n_o$ (m <sup>-1/3</sup> s)	Strahler channel order	$n_c$ (m <sup>-1/3</sup> s)
Evergreen needleleaf forest	0.40	I	0.050
Decidulous needleleaf forest	0.30	II	0.040
Decidulous broadleleaf forest	0.30	III	0.035
Mixed forest	0.25	IV	0.030
Woodland	0.20	V	0.030
Wooded grassland	0.15	VI	0.025
Closed shrubland	0.15	0	0
Grassland	0.095	0	0
Cropland	0.085	0	0
Urban land builtup	0.070		

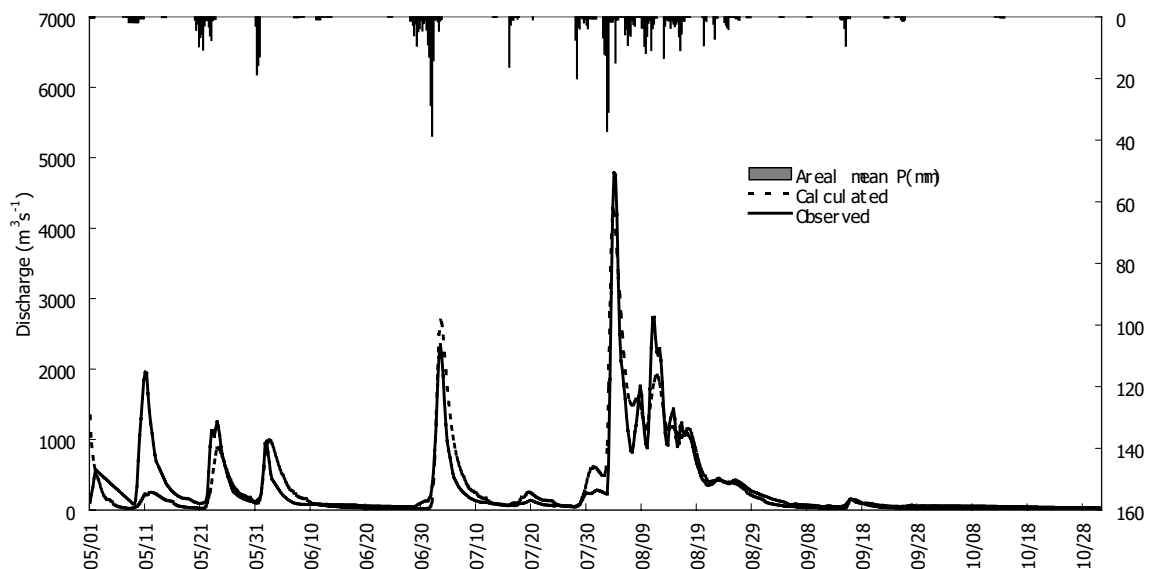


Fig. 7. Comparison of observed and calculated discharge for Xixian (From 1 May to 31 Oct., 1998)

Using the calibrated model parameter values, the new TOPKAPI was validated using the dataset of May 1 to Oct. 31, 2002. Comparison of observed and simulated hydrographs for the catchment outlet (Xixian) and for an

internal gauging station Zhuoganpu (see Fig.3) is shown in Figs. 8 and 9, respectively. The model calibration and validation performance statistics are shown in Table 4.

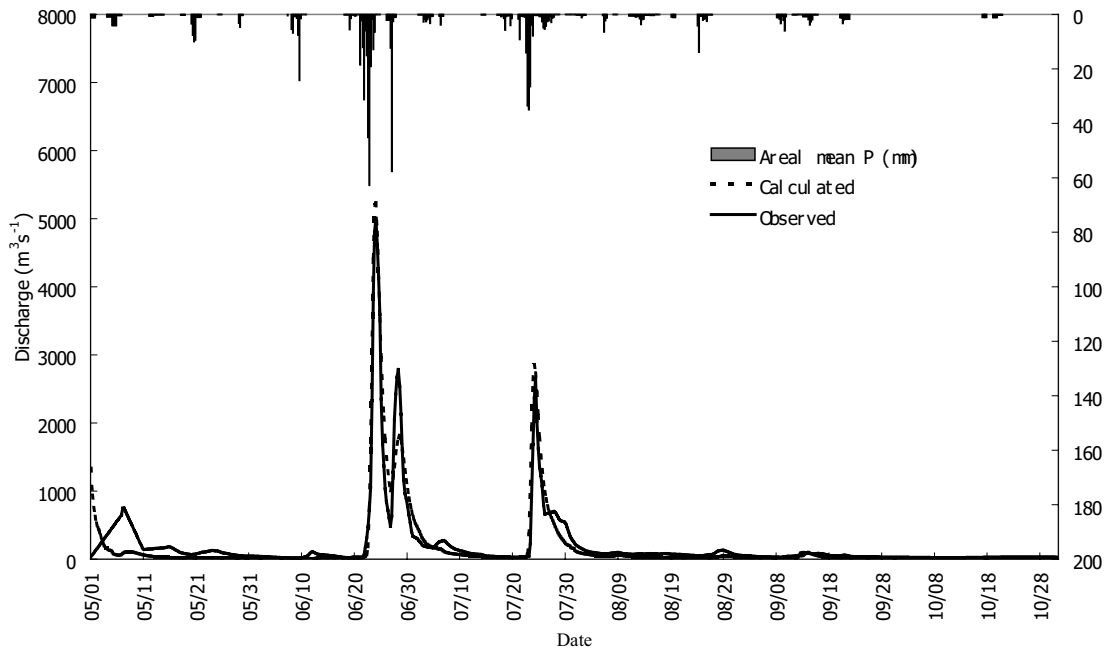


Fig. 8. Comparison of observed and calculated discharge for Xixian (from 1 May to 31 October, 2002).

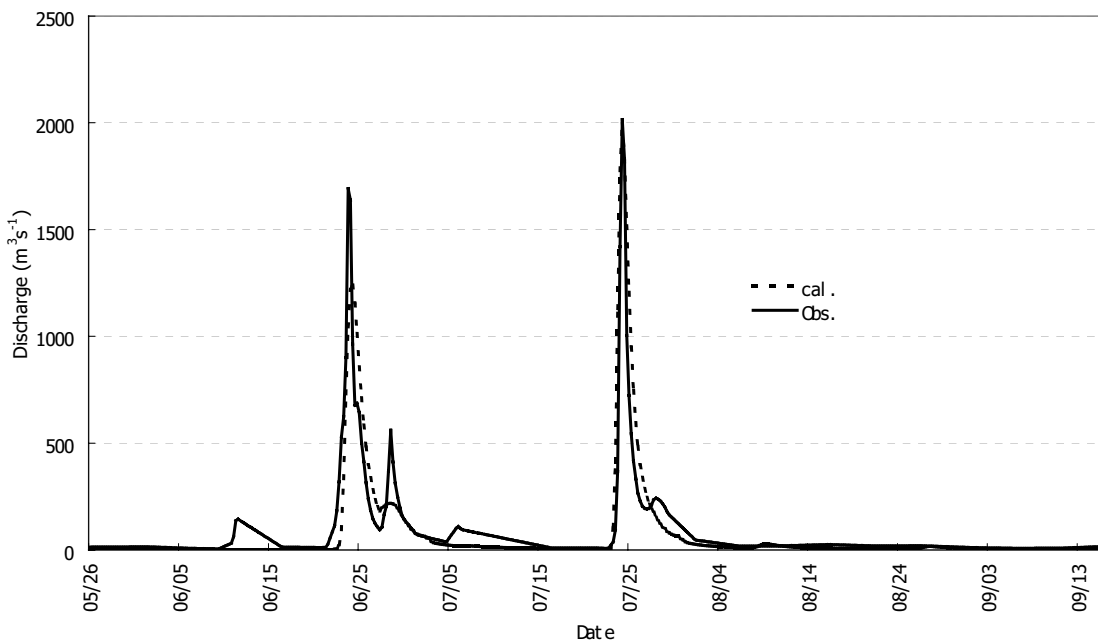


Fig. 9. Comparison of observed and calculated discharge for Zhuoganpu (from May to Sept., 2002).

Table 4. New TOPKAPI model calibration and validation performance statistics for Xixian

Period	Coefficient of Determination $r^2$	Observed peak discharge ( $m^3 s^{-1}$ )	Calculated peak discharge ( $m^3 s^{-1}$ )	Relative error of peak discharge (%)
Calibration : May–Oct, 1998	0.894	4780	4280	-10.5
Verification: May–Oct, 2002	0.844	5010	5230	4.4

DISCUSSION OF RESULTS OF TRADITIONAL CALIBRATION

As can be seen from the split sample testing, the new TOPKAPI shows relatively constant efficiencies both in the calibration and validation periods (Table 4), since the coefficients of determination ranges from 0.894 for the calibration period to 0.844 in the validation period. The relative errors (the ratio of the absolute error to the observed) of estimated maximum peak discharges are -10.5% for the calibration period, and 4.4% for the validation period, respectively.

Figures 7 and 8 demonstrate that the new TOPKAPI model (based on a DEM with a grid size of 1000 m) also performed well in simulating the floods and the low flows for the Xixian station.

The model did not simulate well the initial flood events in both the calibration and the validation periods. This is clearly due to the assumption made for the initial conditions of soil saturation, surface water depth, and channel water depth and so on, while the large difference between the calculated and the observed discharge in the event of 11 May, 1998 is caused largely by the underestimation of precipitation total in the catchment.

CALIBRATION BASED ON PARAMETER UNCERTAINTY AND PREDICTIVE UNCERTAINTY

It is now generally accepted that the uncertainty associated with model parameters and its implications for model output, namely the predictive uncertainty, should be explicitly recognised in the modelling procedure (Beven and Binley, 1992; Ewen and Parkin, 1996; Quinton, 1997; Christiaens and Feyen, 2002). Beven and Binley (1992) developed a generalised likelihood uncertainty estimation (GLUE) framework for representing model parameter and prediction uncertainty within the context of a Monte Carlo analysis. Unfortunately, as demonstrated by Mantovan and Todini (in press), the technique is incompatible with the Bayesian inference approach, which results in an incorrect estimation of both parameter and predictive uncertainties. Consequently, it was decided to use a more formal Bayesian inference approach.

To initiate the Bayesian inference process one has to specify  $f_0(\theta)$ , the probability density function of model parameters, which expresses an *a priori* subjective belief in the model parameter values. Successively, a likelihood function of the parameter values conditional to the observations needs to be found.

If  $\mathbf{y} = (y_1, y_2, \dots, y_n)^T \in R^n$  denotes the vector of the  $n$  observations of interest and  $\hat{\mathbf{y}}_n(\theta, \mathbf{X}_n) = (\hat{y}_1(\theta, \mathbf{x}_1), \hat{y}_2(\theta,$

$\mathbf{x}_2), \dots, \hat{y}_n(\theta, \mathbf{x}_n))^T$  the vector of a given model  $\mathbf{M}(\theta, \mathbf{X}_n)$  output, this will be a function of the  $p$ -dimensional vector of parameters  $\theta \in \Theta$  and of the observations input matrix  $\mathbf{X}_n = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$ . On the assumption that the probability density function of the observations conditional upon the model parameters and the observation input matrix (Eqn. 14) is known:

$$f(\mathbf{y}_n | \theta, \mathbf{X}_n) = f(y_1, y_2, \dots, y_n | \theta, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \quad (14)$$

once the observations have been acquired in statistical inference, the likelihood function  $L_n(\theta; \mathbf{y}_n, \mathbf{X}_n)$  of the parameters is defined as follows:

$$L_n(\theta; \mathbf{y}_n, \mathbf{X}_n) \propto f(\mathbf{y}_n | \theta, \mathbf{X}_n) \quad (15)$$

Knowing the parameters prior probability density function  $f_0(\theta)$  and the likelihood function  $L_n(\theta; \mathbf{y}_n, \mathbf{X}_n)$ , it is then possible to derive the posterior probability density of parameters conditional to the observations, using the Bayes theorem:

$$f_n(\theta | \mathbf{y}_n, \mathbf{X}_n) = \frac{L_n(\theta; \mathbf{y}_n, \mathbf{X}_n) f_0(\theta)}{\int L_n(\theta; \mathbf{y}_n, \mathbf{X}_n) f_0(\theta) d\theta} \quad (16)$$

The probability density function of Eqn. 16 is the probability density function that can be used for estimating the value of the parameters and for assessing the parameter uncertainty.

In hydrological applications, the prior probability density of the parameters is generally taken as a non-informative multi-uniform distribution, which has the advantage of cancelling out in Eqn. (16). The use of a non-informative prior in this case is justified since, when dealing with samples of the order of hundreds of observations, the effect of a prior density tends to vanish with respect to that of the likelihood. Therefore, the main problem is the specification of an appropriate likelihood.

To do so, in general, an additive error model is formulated:

$$\mathbf{y}_n = \mathbf{M}(\theta, \mathbf{X}_n) + \boldsymbol{\varepsilon}_n \quad (17)$$

with  $\boldsymbol{\varepsilon}_n = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$  the error vector.

Consequently, the probability density of the observations conditional on  $\mathbf{M}(\theta, \mathbf{X}_n)$ , namely on (a) the model, (b) the parameters and (c) the input observations, can be derived as a function of the stochastic properties of the error vector. Different hypotheses of distribution laws for the error vector have been proposed, for instance the multivariate  $t$ -distribution law (Johnson and Kotz, 1972; Mardia *et al.*, 1979; Ripley, 1997). If the model with additive errors is stochastically independent and if errors are equally

distributed, it can be assumed that:

$$f(\boldsymbol{\varepsilon}_n | \boldsymbol{\beta}, N) = \frac{N\boldsymbol{\beta}^{1/N}}{2\Gamma(1/N)} \exp(-\boldsymbol{\beta}|\boldsymbol{\varepsilon}_n|^N) \quad (18)$$

with  $N \geq 1$  (Box and Tiao, 1983; Bishop, 1996); alternatively, a multivariate distribution that accounts for the structural dependency among the errors must be defined.

Several batch or recursive algorithms for parameter estimation and uncertainty assessment have recently been developed following this assumption, such as the Parameter Identification Method based on the Location of Information (PIMLI) (Vrugt, 2004) or the Bayesian Recursive Estimation (BaRE) (Thiemann *et al.*, 2001).

A relatively robust approach is used in this paper by transforming the original data into a Gaussian space via the Normal Quantile Transform (Van der Waerten, 1952, 1953; Kelly and Krzysztofowicz, 1997) by probability matching. Namely, after ranking the original variables in increasing order, they are converted into the standard normal variable corresponding to probability  $i/(n_i + 1)$ , where  $i$  is the ranking order and  $n_i$  the sample size. In this way Eqn. 17 is transformed into:

$$\eta_n = \hat{\eta}_n(\hat{\mathbf{y}}_n(\mathbf{M}, \boldsymbol{\vartheta}, \mathbf{X}_n)) + \xi_n \quad (19)$$

with  $\eta_n(\mathbf{y}_n)$  and  $\hat{\eta}_n(\hat{\mathbf{y}}_n(\mathbf{M}, \boldsymbol{\vartheta}, \mathbf{X}_n))$  (where  $\hat{\mathbf{y}}_n(\mathbf{M}, \boldsymbol{\vartheta}, \mathbf{X}_n)$  represents the model output) normally distributed with mean zero and variance 1; consequently,  $\xi_n$  results normally distributed with mean zero and variance  $1 - \rho^2$ , with  $\rho$  the correlation coefficient between  $\eta_n$  and  $\hat{\eta}_n(\hat{\mathbf{y}}_n(\mathbf{M}, \boldsymbol{\vartheta}, \mathbf{X}_n))$ . This property allows the probability density function of observations conditional to the parameters to be defined as:

$$f(\boldsymbol{\eta}_n | \boldsymbol{\vartheta}, \hat{\mathbf{H}}_n) = f(\eta_1, \eta_2, \dots, \eta_n | \boldsymbol{\vartheta}, \hat{\eta}_1, \hat{\eta}_2, \dots, \hat{\eta}_n) \quad (20)$$

which is proportional to the likelihood of parameters according to Eqn. 15. Following the Bayes theorem, the posterior density of the parameters given the observations is then defined as:

$$f_n(\boldsymbol{\vartheta} | \boldsymbol{\eta}_n, \hat{\mathbf{H}}_n) = \frac{L_n(\boldsymbol{\vartheta}; \boldsymbol{\eta}_n, \hat{\mathbf{H}}_n) f_0(\boldsymbol{\vartheta})}{\int_{\boldsymbol{\vartheta}} L_n(\boldsymbol{\vartheta}; \boldsymbol{\eta}_n, \hat{\mathbf{H}}_n) f_0(\boldsymbol{\vartheta}) d\boldsymbol{\vartheta}} \quad (21)$$

Note that although the time correlation of  $\xi_n$  was not accounted for here, it can easily be incorporated by using a multivariate Normal distribution.

The posterior density of the parameters describes the parameter uncertainty directly and allows one to estimate the predictive uncertainty.

The Bayesian approach was applied to the case of the Upper Xixian basin using 650 observations relevant to the

calibration period; a number of scarcely significant observations towards the end of the last month were disregarded in order to reduce the computational effort.

The four most significant TOPKAPI model parameters were chosen for this analysis, namely the logarithm of the hydraulic conductivity at saturation, the soil thickness, the Manning coefficient for the land surface and the Manning coefficient for the river. Moreover, to reduce the number of parameters, instead of independently calibrating all the individual parameters for the different soil types, land uses and river reaches, it was decided to apply the Bayesian approach to a set of four multiplicative factors. These factors are then applied identically to all the different classes of soils, different land uses and river reaches. A number of 50,000 sets of four parameters was generated  $\boldsymbol{\vartheta}_j = (\vartheta_1, \vartheta_2, \dots, \vartheta_p)_j \quad \forall p = 1, \dots, 4; j = 1, \dots, 50,000$  using a multi-uniform distribution function each of which was run with the TOPKAPI model. This operation produced 50 000 sets of 650 time steps model outputs  $\hat{y}_{ij} \quad \forall i = 1, \dots, 650; j = 1, \dots, 50,000$  that, together with the observed discharges  $y_i$ , were converted, via the Normal Quantile Transform, into their images in the normal space  $\eta_i, \hat{\eta}_{ij} \quad \forall i = 1, \dots, 650; j = 1, \dots, 50,000$  from which it was possible to compute  $\xi_{nj} = (\xi_{1j}, \xi_{2j}, \dots, \xi_{4j})_j \quad \forall j = 1, \dots, 50,000$  and derive the likelihood of each generated parameter set

$$L_n(\boldsymbol{\vartheta}_j; \boldsymbol{\eta}_n, \hat{\mathbf{H}}_n) \propto f(\boldsymbol{\xi}_{nj} | \boldsymbol{\vartheta}_j) = \frac{\exp\left(-\frac{1}{2} \sum_{i=1}^n \xi_{ij}^2\right)}{(2\pi)^{n/2}} \quad (22)$$

$\forall j = 1, \dots, 50,000$

which coincides with  $L_n(\boldsymbol{\vartheta}_j; \mathbf{y}_n, \mathbf{X}_n)$  as well as with the posterior parameter density, given the peculiar choice of the non-informative multi-uniform prior.

Figure 10 displays the different transformations from the original data required to produce the normally distributed deviates used in the Bayesian approach.

A sensitivity analysis of the parameters was then performed, the results of which are given in Fig. 11 in terms of two parameters at a time (one parameter versus another). As opposed to what happens when GLUE is applied (Fig. 12), peak concentrations of parameter likelihood (see for instance the factors relevant to the land surface versus the river Manning friction coefficient) appear evident.

Once the posterior parameter density function has been constructed it is possible to derive predictive uncertainty of the model. To do so, it is necessary to define the probability density of the predictand conditional on the model output, which is a function of model parameters:

$$f(y_k | \mathbf{M}(\boldsymbol{\vartheta}, \mathbf{x}_k)) \quad (23)$$



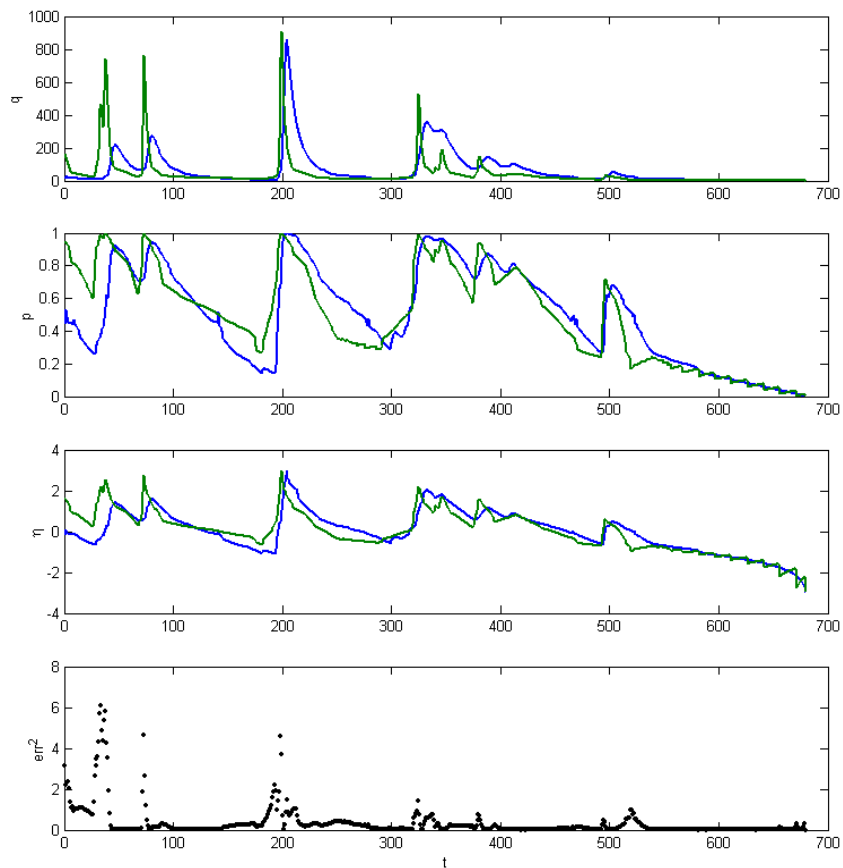


Fig. 10. Procedure to estimate the normally distributed errors associated with the  $j$ -th model: (1) observed and simulated discharges; (2) order probability associated with the observed and simulated discharges; (3) standard normal quantiles of the observed and simulated discharges; (4) squared errors. The normal black solid line represents the observed values and their transformations.

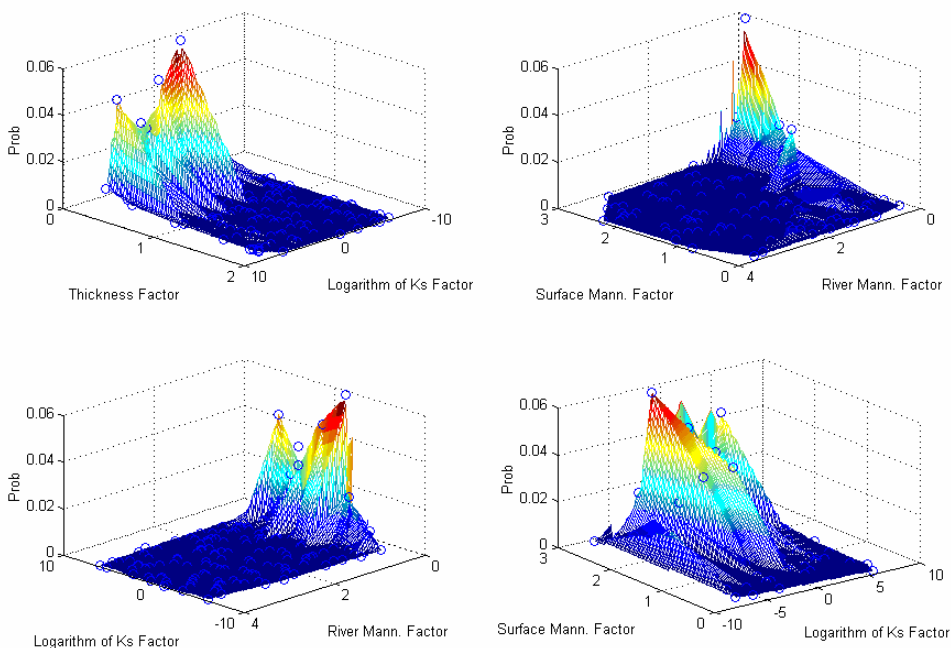


Fig. 11. Two by two parameter factors sensitivity analysis.

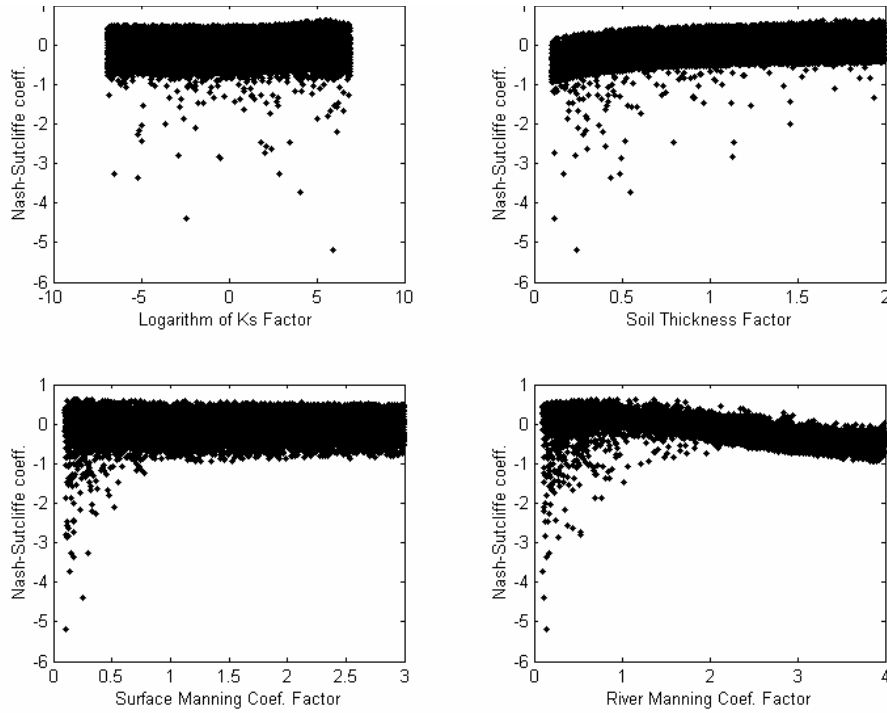


Fig. 12. *GLUE* dotty plots for the four parameter factors. The predictive uncertainty here discussed relates to the uncertainty of a predictand (in this case discharge or water stage or more in general a measurable quantity of interest) induced by the parameter uncertainty, conditional upon a model and its inputs, which are taken as given.

Once this density is known, the predictive uncertainty can be defined as follows:

$$\int_{\mathcal{D}} f(y_k | \mathbf{M}(\mathcal{G}, \mathbf{x}_k)) f(\mathcal{G} | \mathbf{y}_n, \mathbf{X}_n) d\mathcal{G} \quad (24)$$

As was done to derive the posterior parameter density, a predictive density relevant to each single parameter set  $\mathcal{G}_j$  can be obtained via the NQT approach initially by forming the joint probability distribution  $f(\eta_n, \hat{\eta}_n | \mathbf{M}, \mathcal{G}_j, \mathbf{x}_n)$  in the Normal space and by deriving the conditional density  $f(\eta_k | \hat{\eta}_k(\mathbf{M}, \mathcal{G}_j, \mathbf{x}_k))$  as shown in Fig. 13. The latter density can then be transformed back in the original space through the inverse of the NQT, to give  $f(y_k | \hat{y}_k(\mathbf{M}, \mathcal{G}_j, \mathbf{x}_k))$ .

Finally, the predictive density can be found by marginalising all these densities with respect to the parameters, which can be approximated by summing up the products of all the realisations times of their relevant probability density of occurrence, to give:

$$f(y_k | \mathbf{M}, \mathbf{x}_k, \mathbf{y}_n, \mathbf{X}_n) = \sum_{j=1}^m f(y_k | \hat{y}_k(\mathbf{M}, \mathcal{G}_j, \mathbf{x}_k)) f(\mathcal{G}_j | \mathbf{y}_n, \mathbf{X}_n) \quad (25)$$

where  $m = 50,000$  in this case.

In the upper part of Fig. 13, the observed discharge (solid line) is compared with the expected value conditional on the model, input and initial conditions (dotted line) plus and

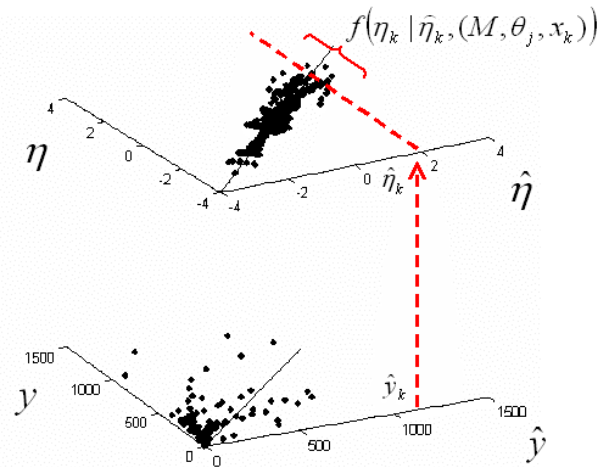


Fig. 13. The joint (dots cloud) and the conditional (section) probability densities in the Normal space

minus two sigma limits, while a number of estimated predictive probability distribution functions for selected cases are shown in the lower part. Figure 13 shows that the estimated predictive probability distribution (and consequently the uncertainty band) are non-Gaussian, highly variable in time, sometimes positively and sometimes negatively skewed, as would be expected.

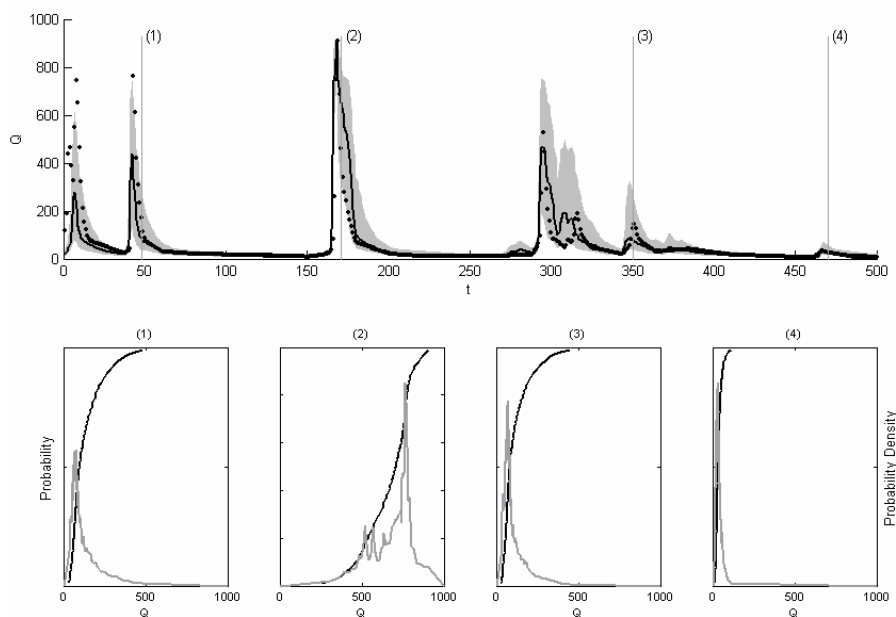


Fig. 14. Predictive uncertainty. Upper figure: observed discharge (solid), expected value (dotted) and the 95% confidence boundaries. (grey bands). Lower figure: predictive probability distribution function (black) and probability density function (grey) at selected times

## Conclusion

A new version of a physically-based, fully distributed TOPKAPI model was developed. The structure and methodology, parameters and data requirements of the new TOPKAPI, with the emphasis of illustrating the new components of the model is described. The application to the upper Xixian catchment (*c.* 10 000 km<sup>2</sup>) using public domain data sets readily available over the internet and applying the new TOPKAPI is described. For the two-year period, the discharge was simulated well against the response at the catchment outlet as well as at an internal gauging station.

The application of the new TOPKAPI model to the Upper Xixian catchment demonstrated that the model performed well in simulating floods and low flows. Since the model-required basin's surface data (topography, soils and land use/land cover, etc.) can be derived from the public domains through the Internet, it is expected that the TOPKAPI model will be applied in more catchments for flood forecasting. Once a parameter database for the TOPKAPI is prepared, the application of the model can be broadened to other applications, including extreme flood analysis in ungauged catchments.

The new TOPKAPI is described as a physically-based, spatially distributed model. Nonetheless, a formal Bayesian inference technique was applied to demonstrate how it is possible on the one hand to perform parameter sensitivity analysis and, on the other hand, to provide an estimate of

the predictive uncertainty.

The full Bayesian approach described in this paper can be used to assess the model performances using historical data, but can seldom be applied in real time due to the large amount of computational effort required. Ongoing research aims at demonstrating a possible alternative by using in real time applications the predictive probability density conditional to the Maximum Likelihood or the Maximum Posterior Density parameter estimates instead of the marginalised density of Eqn. (25). Preliminary results show that the two distributions are quite similar because the Posterior Parameter Density results are highly dense around its maximum.

## References

- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., 1986a. 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'Connell, P.E. and Rasmussen, J., 1986b. An introduction to the European Hydrological System – Système Hydrologique Européen, SHE, 2: Structure of a physically based, distributed modelling. *J. Hydrol.*, **87**, 61–77.
- Ambroise, B., Perrin, J.L. and Reutenauer, D., 1995. Multicriterion validation of a semidistributed conceptual model of the water cycle in the Fecht Catchment (Vosges Massif, France). *Water Resour. Res.*, **31**, 1467–1481.
- Anderton, S., Latron, J. and Gallart, F., 2002. Sensitivity analysis and multi-response, multi-criteria evaluation of a physically based distributed model. *Hydrol. Process.*, **16**, 333–353.

- Artan, G., Verdin, J. and Asante, K., 2001. A wide-area flood risk monitoring model. *Proc. 5th Int. Workshop on Application of Remote Sensing in Hydrology*, Montpellier, France.
- Band, L.E., 1986. Topographic partition of watersheds with digital elevation models. *Water Resour. Res.*, **22**, 15–24.
- Barnes, H.H., 1967. *Roughness characteristics of natural channels*. United States Geological Survey Water Supply Paper 1899. US Government Printing Office, Washington.
- Bedient, P.B. and Huber, W.C., 1992. *Hydrology and floodplain analysis*, 2<sup>nd</sup> edn., Addison-Wesley, Reading, MA, USA.
- Beven, K. and Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.*, **6**, 279–298.
- Beven, K.J. and Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrolog. Sci. Bull.*, **24**, 43–69.
- Bishop, C.M., 1996. *Neural networks for pattern recognition*. Clarendon Press, Oxford, UK.
- Box, G.E.P. and Tiao G.C., 1973. *Bayesian inference statistical analysis*. Addison-Wesley-Longman, Reading, MA, USA.
- Chen, H., 1996. *Object watershed link simulation (OWLS)*. Ph.D dissertation, Oregon State University, USA.
- Chow, V.T., 1959. *Open-channel hydraulics*. International student edition. McGraw-Hill Book Company, Inc., USA.
- Christiaens, K. and Feyen, J., 2002. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. *Hydrol. Process.*, **16**, 373–391.
- Clapp, R.B. and Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. *Water Res. Res.*, **14**, 601–604.
- de Marsily, Ghislain, 1986. Geostatistic and stochastic approach in hydrogeology (Chapter 11). In: *Quantitative hydrology: Groundwater Hydrology for Engineers*. Academic Press, Inc., California, USA. 286–329.
- Doorenbos, J. and Pruitt, W.O., 1992. *Crop water requirements*. FAO Irrigation and Drainage Paper No. 24. FAO, Rome, Italy.
- Doorenbos, J., Pruitt, W.O., Aboukhaled, A., Damagnez, J., Dastane, N.G., van den Berg, C., Rijtema, P.E., Ashford, O.M. and Frere, M., 1984. *Guidelines for predicting crop water requirements*. FAO Irrig. Drainage Pap., 24.
- Ewen, J. and Parkin, G., 1996. Validation of catchment models for predicting land-use and climate change impacts. 1. Method. *J. Hydrol.*, **175**, 583–594.
- FAO, 1996. *The digitized soil map of the world Including derived soil properties*. CD-ROM, Food and Agriculture Organization, Rome, Italy.
- Feyen, L., Vázquez, R., Christiaens, K., Sels, O. and Feyen, J., 2000. Application of a distributed physically-based hydrological model to a medium size catchment. *Hydrol. Earth Syst. Sci.*, **4**, 47–63.
- Framework project, 2000. *An integrated hydrological water balance model for the assessment of man's influence on severe floods*. Deliverable no. D.6.1 of FRAME (Flash-flood Risk Assessment under the impacts of land use changes and river Engineering works) project under EC contract ENV4-CT97-0529.
- Franchini, M., Wendling, J., Obled, C. and Todini, E., 1996. Physical interpretation and sensitivity analysis of the TOPMODEL. *J. Hydrol.*, **175**, 293–338.
- Gesch, D.B., Verdin, K.L. and Greenlee, S.K., 1999. New land surface digital elevation model covers the earth. *EOS, Trans. Amer. Geophys. Un.*, **80**, 69–70.
- HEC-1, 1998. *Flood hydrograph package user's manual*. Hydrologic Engineering Center, US Army Corps of Engineers.
- Hansen, M.C., Defries, R.S., Townshend, J.R.G. and Sohlberg, R., 2000. Global land cover classification at 1km spatial resolution using a classification tree approach. *Int. J. Remote Sens.*, **21**, 1331–1364.
- Johnson, N.L. and Kotz, S., 1970. *Continuous univariate distributions-2*. Houghton Mifflin Company, Boston, USA.
- Kelly, K.S. and Krzysztofowicz, R., 1997. A bivariate meta-Gaussian density for use in hydrology. *Stochast. Hydrol. Hydraul.*, **11**, 17–31.
- Kite, G.W., 1995. The SLURP model (Chapter 15). In: *Computer models of watershed hydrology*, V.P. Singh (Ed.). Water Resources Publications, Colorado, USA. 521–562.
- Li, Y., 2001. *Effects of the choice of data for calibration on the performance of the Xinanjiang model, applied to the Huai River [D]*. MSc. Thesis. IHE, Delft, The Netherlands.
- Liang, X., Lettenmaier, D.P. and Wood, E.F., 1996a. One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. *J. Geophys. Res.*, **101** (D16) 21,403–21,422.
- Liang, X., Wood, E.F. and Lettenmaier, D.P., 1996b. Surface soil moisture parameterization of the VIC-2L model: evaluation and modifications. *Global Planet. Change*, **13**, 195–206.
- Liu, Z., 2002. *Toward a comprehensive distributed/lumped rainfall-runoff model: analysis of available physically-based models and proposal of a new TOPKAPI model*. PhD dissertation, The University of Bologna, Italy.
- Liu, Z. and Todini, E., 2002. Towards a comprehensive physically-based rainfall-runoff model. *Hydrol. Earth Syst. Sci.*, **6**, 859–881.
- Maidment, D.R. (Ed.), 1993. *Handbook of Hydrology*, McGraw-Hill, New York, USA.
- Mantovan, P. and Todini, E., Hydrological forecasting: Incoherence of the GLUE methodology. *J. Hydrol.*, in press.
- Mardia, K.V., Kent, J.T. and Bibby, J.M., 1979. *Multivariate analysis*. Academic Press, London, UK.
- Quinton, J.N., 1997. Reducing predictive uncertainty in model simulations : a comparison of two methods using the European Soil Erosion Model (EUROSEM). *Catena*, **30**, 101–117.
- Ripley, B.D., 1987. *Stochastic simulation*. Wiley, Chichester, UK.
- Reynolds, C.A., Jackson, T.J. and Rawls, W.J., 1999. Estimated available water content from the FAO Soil Map of the World, Global Soil Profile Databases, and Pedo-transfer Functions, available online at <http://www.ngdc.noaa.gov/seg/eco/cdroms/reynolds/reynolds/reynolds.htm>.
- Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. *Trans. Amer. Geophys. Un.*, **38**, 913–920.
- Thiemann, M., Trosset, M., Gupta, H. and Sorooshian, S., 2001. Bayesian recursive parameter estimation for hydrologic models. *Water Resour. Res.*, **37**, 2521–2535.
- Todini, E., 1996. The ARNO rainfall-runoff model. *J. Hydrol.*, **175**, 339–382.
- Todini, E. and Ciarapica, L., 2001. The TOPKAPI model. (Chapter 12). In: *Mathematical models of large watershed hydrology*, V.P. Singh et al. (Eds.). Water Resources Publications, Littleton, Colorado, USA.
- Todini, E., Marsigli, M. and Zamboni, L., 1999. The role of GIS in the decision support system ODESSEI (Open architecture DEcision Support System for Environmental Impact assessment). 2<sup>nd</sup> Inter-Regional Conference on Environment-Water 99, Lausanne, Switzerland.
- Van der Waerden, B.L., 1952. Order tests for two-sample problem and their power. *Indagationes Mathematicae*, **14**, 253–458.
- Van der Waerden, B.L., 1953. Order tests for two-sample problem and their power. *Indagationes Mathematicae*, **15**, 303–316.

- Vázquez, R.F., Feyen, L., Feyen, J. and Refsgaard, J.C., 2002. Effect of grid size on effective parameters and model performance of the MIKE-SHE code. *Hydrol. Process.*, **16**, 355–372.
- Vrugt, J.A., 2004. *Towards improved treatment of parameter uncertainty in hydrologic modeling*. Ph.D. Thesis of University of Amsterdam, The Netherlands. ISBN:90-76894-46-9.
- Wigmosta, M.S., Vail, L.W. and Lettenmier, D.P., 1994. A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1165–1679.
- Wooding, R.A., 1965. A hydraulic modeling of the catchment-stream problem. 1. Kinematic wave theory. *J. Hydrol.*, **3**, 254–267.