A spatially distributed flash flood forecasting model

Günter Blöschl*, Christian Reszler, Jürgen Komma

Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Karlsplatz 13/222, A-1040 Vienna, Austria

Received 16 March 2007; received in revised form 25 June 2007; accepted 26 June 2007
Available online 31 August 2007

Abstract

This paper presents a distributed model that is in operational use for forecasting flash floods in northern Austria. The main challenge in developing the model was parameter identification which was addressed by a modelling strategy that involved a model structure defined at the model element scale and multi-source model identification. The model represents runoff generation on a grid basis and lumped routing in the river reaches. Ensemble Kalman Filtering is used to update the model states (grid soil moisture) based on observed runoff. The forecast errors as a function of forecast lead time are evaluated for a number of major events in the 622 km² Kamp catchment and range from 10% to 30% for 4–24 h lead times, respectively.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Forecasting; Parameter identification; Kalman Filter; Floods; Distributed modelling; Stream routing; Model accuracy; Dominant processes concept

1. Introduction

Recent years have seen an explosion in the development and use of spatially distributed models in hydrology. For the particular case of flash flood forecasting their merits are obvious. Spatially distributed data on the landscape are widely available and are awaiting use in predictive analysis. Rainfall inputs are increasingly available in a spatially distributed fashion and one would expect that the location of rainfall relative to the runoff contributing areas is important for making accurate forecasts. The computational resources typically installed in forecasting centres make complex spatial computations feasible. The huge amount of information stored in the databases might suggest that the development of distributed hydrological models has been reduced to a software engineering task but it is argued in this paper that indeed it has not. It is a genuinely hydrological task that requires knowledge of the hydrological processes involved and the skill of parameterising them in suitable ways. This is in the spirit of the 10 iterative steps in development and evaluation of models proposed by Jakeman et al. (2006).

The aim of this paper is to discuss some of the challenges of distributed modelling in the context of developing a distributed flood forecasting system. The discussion will be illustrated by the example of the flood forecasting system of the Kamp catchment in Austria.

The paper is organised as follows. Section 2 discusses issues in distributed modelling and a strategy to model building. Section 3 gives a description of the Kamp catchment. Section 4 presents the model structure and the input data used. Section 5 gives the results of the parameter identification procedure and Section 6 reports on the operational use and real time updating.

2. Issues in distributed modelling and a strategy to model building

With the computational resources available today to most modellers, it has become feasible to build and apply highly complex distributed hydrological models that represent many different processes and consist of many model elements. Among the first to recognise, however, that, in hydrology,
“finer” is not necessarily “better” were Stephenson and Freeze (1974) and there is a long track record of studies demonstrating and discussing the difficulties in model identification and calibration once the model becomes too complex (e.g., Loague and Freeze, 1985; Beven, 1989, 2001; Blöschl, 2005). What is the reason for this counterintuitive fact, which is apparently at variance with experience in fluid dynamics and other geosciences? There now is a growing awareness that distributed hydrological models are different from models in sister disciplines in at least three important aspects. First, and probably most important, the media properties (both soil and vegetation) are highly heterogeneous and essentially always unknown or at least poorly known. There will always exist some variability within a grid element — no matter how fine the model resolution is — that cannot be resolved. Also, not only is the landscape heterogeneous but the heterogeneity is complex and an adequate statistical distribution of it is difficult to find. Second, there is no unique hydrological equation that can be derived from first principles, so most of the model equations are empirical in nature and tend to depend on the hydrological setting. Third, hydrological models are very much dependent on their boundary conditions, and these are often poorly defined. The “model dynamics” are relatively less important than, say, those in fluid dynamics. While it is possible to study the global dynamics of the atmosphere by spinning up a model and let it run for a period, this is not possible for a hydrological model.

These three aspects have two important implications for distributed modelling. The first is that there will always be some degree of calibration needed for any model to accurately represent the hydrological processes in a particular case. The second is that the appropriate choice of model complexity at the element scale depends on how much information is available on the natural variability. A model with very small elements and many process descriptions that, in principle can represent great detail, will unlikely have value over coarser models unless the data are available to define the variability of the model parameters (Gräson and Blöschl, 2000a). It is indeed a common situation for practical applications of distributed models that too complex a model with limited data are used which causes identifiability problems. In the context of this paper these issues are addressed by adopting a modelling strategy that is based on two principles: (a) model structure defined at the model element scale, and (b) multi-source model identification and verification.

(a) Model structure: the idea of avoiding excessive model complexity has a long tradition in science starting from the ideas of 14th century philosopher William of Ockham. An amazing range of modelling approaches exists in hydrology. On the one end of the spectrum of approaches are complex physically based models with the SHE Model (Abbott et al., 1986) probably being the classical example of models that are based on point (or laboratory) scale equations. Point scale equations can be straightforwardly extended to catchments, aquifers, reaches, etc. provided the boundary conditions are known and the media characteristics are known spatially (e.g. uniform) at the scale of the equations. However, hydrological systems are never completely uniform in terms of their parameters, fluxes and states, and are often not even approximately uniform and the variability is rarely known (Blöschl and Zehe, 2005; Blöschl, 2006). This is the rationale of using simpler models including models based on the systems approach or the related downward approach (Klemes, 1983; Sivapalan et al., 2003). For example, Jakeman and Hornberger (1993) and Littlewood et al. (2007), suggested that transfer function models involving four parameters may suffice to accurately represent the runoff dynamics from a catchment. In the context of distributed modelling, four parameters may not be enough to represent the complex interplay between rainfall patterns and the landscape (Moretti and Montanari, 2007; Krysanova et al., 2007). However, it may be prudent to formulate the model equations directly at the model element scale. This supports the choice of conceptual models that are based on solving ordinary differential equations rather than partial differential equations as is the case in physically based models. The idea is that this type of model allows some level of hydrological interpretation of the parameters defined at the model element scale rather than at the point scale. Interpretability of model parameters may be an advantage in the parameter identification step. Additionally, these models are usually numerically robust and efficient which is important in an operational context, particularly if ensemble methods are used, e.g., for updating the runoff model in a real time mode.

(b) Multi-source model identification: this strategy builds on the notion that runoff data are a necessary, but not a sufficient, condition for identifying model parameters in a realistic way. Gräson and Blöschl (2000b) have argued that the development, calibration and testing of distributed models should ideally involve observed spatial patterns of catchment response, and that the use of runoff data alone can be greatly misleading. These patterns of catchment response can come from a number of sources. Recent years have seen an increase in the availability of ground-based pattern data in catchments and from remote sensing, up to the global scale. This has led to a number of examples of using patterns for developing and testing distributed models most of which demonstrated the value of observed patterns. The type of variable to be used clearly depends on the hydrological processes that are relevant in a particular hydro-climatologic setting. For example, in snow dominated regimes, snow cover patterns have been shown to be useful for testing distributed models (Blöschl et al., 1991). Other examples include inundation patterns, soil moisture patterns and the spatial distribution of the groundwater table (Gräson et al., 2002). In the context of the present study, a range of spatial data have been used that are complementary. These data include piezometric heads, spatial patterns of snow, both from satellite data and ground-based data, inundation patterns as well as soft information, e.g., on surface flow pathways during
past floods. Also, different types of events have been ana-
lysed (e.g. convective events and snowmelt events) to bet-
ter constrain the parameter space of the model.

3. The Kamp catchment

The Kamp catchment is located in northern Austria, ap-
proximately 120 km north-west of Vienna, between 48.37
and 48.76° N, and 14.78 and 15.78° E. The catchment area of
the most downstream forecast point is 1550 km². Elevations
range from 300 to 1000 m a.s.l. The geology of the catchment
is mainly granite and gneiss. Weathering has produced sandy
soils with a large storage capacity throughout the catchment.
Fifty per cent of the catchment is forested. The catchment is
heterogeneous in terms of its land use. As an illustration of
this, Fig. 1 shows a Landsat image of part of the Kamp area
in Austria with a 1 km grid indicated. There is hetero-
geneity including forest, different paddocks and urban area at
scales much finer than the grid shown. Mean annual precipita-
tion in the Kamp catchment is about 900 mm of which about
300 mm become runoff (Parajka et al., 2005). The western part
of the Kamp catchment drains into the Kamp reservoir scheme
of the EVN-AG hydropower operator. The scheme consists of
the Ottenstein, Dobra und Thurnberg reservoirs.

Flood generation in the Kamp catchment is characterised
by a number of particularities as compared to other catchments
in Austria. First, it is only a small proportion of rainfall that
contributes to direct runoff. As rainfall increases in magnitude,
the runoff response characteristics change fundamentally be-
cause of the soil moisture changes in the catchment. This
may result in floods that are much larger than average ones.
Fig. 2 shows the observed runoff at the Kamp at Zwettl stream
gauge from October 2001 to September 2002. During most of
the year, runoff is less than 10 m³/s while the flood in August
2002 peaked at a discharge of 460 m³/s. This type of runoff
regime makes runoff modelling challenging. The Kamp flood
forecasting system continuously simulates the catchment soil
moisture state in order to obtain antecedent soil moisture as
an initial condition for the forecasts. Second, the Kamp catch-
ment is very heterogeneous, both in terms of the catchment
characteristics and in terms of the rainfall input. The interplay
of the spatial distributions of rainfall with the runoff contrib-
uting areas controls the magnitude and the shape of flood
hydrograph. This interplay is represented in the flood forecast-
ing system by using a distributed hydrological model with spa-
tially distributed rainfall input.

4. Model structure

The Kamp catchment was divided into 13 subcatchments
(1–13 in Fig. 3) and in each of them runoff generation is esti-
ated on a 1 km square grid. The subcatchments are connected
by 10 routing modules (a–j in Fig. 3) as well as three modules
that represent the hydraulic characteristics of the Ottenstein,
Dobra and Thurnberg reservoirs (A, B, C in Fig. 3). The entire
forecast system runs on a time step of 15 min.

4.1. Pixel scale processes

The catchment is represented by a total of 1550 square grid
elements. For each grid element, snow processes, soil moisture

![Fig. 1. Landsat image of part of the Kamp catchment Austria around the township of Zwettl (grid element C38/R26). Forests are dark, paddocks are light grey. A 1 km grid is shown.](image-url)
processes and hillslope scale routing are simulated. These processes are formulated directly at the model element scale of 1 km².

Snow model: the snow routine represents snow accumulation and melt by a simple degree day concept. Precipitation routing reaches (\(a\)) processes are formulated directly at the model element scale of the precipitation gauges during snowfall is corrected by melting is represented in the model by increasing M

\[
S_{WE,i} = S_{WE,i-1} + (C_s P_s - M) \Delta t
\]

where \(\Delta t\) is the time step of 15 min.

Soil moisture accounting: the sum of rain and melt, \(P_t + M\), is split into a component \(dS\) that increases soil moisture of a top layer, \(S_s\), and a component \(Q_p\) that contributes to runoff. The components are split as a function of \(S_s\):

\[
Q_p = \left( \frac{S_s}{P_t} \right)^\beta (P_t + M)
\]

where \(L_s\) is the maximum soil moisture storage (Bergström, 1976); \(\beta\) controls the characteristics of runoff generation and is termed the non-linearity parameter. If the top soil layer is saturated, i.e., \(S_s > L_s\), all rainfall and snowmelt contributes to runoff and \(dS = 0\). If the top soil layer is not saturated, i.e., \(S_s < L_s\), rainfall and snowmelt contribute to runoff as well as to increasing \(S_s\) through \(dS > 0\):

\[
dS = P_t + M - Q_p - Q_{by}, \quad \text{if } P_t + M - Q_p - Q_{by} > 0
dS = 0, \quad \text{otherwise}
\]

while no bypass flow was assumed to occur for dry and very wet soils. Changes in the soil moisture of the top soil layer \(S_s\) from time step \(i - 1\) to \(i\) are accounted for by

\[
S_{by} = \alpha_{by}(P_t + M), \quad \text{if } \alpha_{by}(P_t + M) < L_{by}
Q_{by} = L_{by}, \quad \text{otherwise}
\]

where \(L_p\) is a parameter termed the limit for potential evaporation and leaves this reservoir through three paths: percolation to a snow correction factor, \(C_s\). Changes in the snow water equivalent from time step \(i - 1\) to \(i\) are accounted for by

\[
S_{WE,i} = S_{WE,i-1} + (C_s P_s - M) \Delta t
\]
the lower and groundwater zones with a given percolation rate $c_p$, outflow from the reservoir with a fast storage coefficient of $k_1$, and, if a threshold $L_1$ of the storage state is exceeded, through an additional outlet with a very fast storage coefficient of $k_0$. The percolation rate $c_p$ is split into two components by a fraction $a_0$ which flow into the lower zone reservoir and the groundwater reservoir. Bypass flow $Q_{by}$ is directly added to the lower zone reservoir. $Q_0$ represents fast runoff reflecting surface runoff or near surface runoff. $Q_1$ is a somewhat slower component reflecting interflow. $Q_2$ is a slower component from the lower soil zone. $Q_3$ is the slowest component and is attributed to groundwater flow. This conceptualisation is consistent with the observed runoff data in the catchments, the hydrogeological interpretation, as well as other data as used in the parameter identification step. The analysis of catchment response data also suggested that $k_1$ and $k_2$ should be related to $S_c$. A linear relationship was assumed:

\[
k_1 = k_1^* \left( 1 + \frac{\delta_1 \cdot S_c}{L_g} \right)
\]

where $k_1^*$ is a storage coefficient and $\delta_1$ is a free parameter. An analogous relationship for $k_2$ was used. Finally, the runoff data indicated that percolation $c_p$ changes with soil moisture. It was hence related to the storage of the top soil $S_c$ by

\[
c_p = \left( \frac{S_c}{L_g} \right)^\gamma L_g
\]

where $L_g$ is the maximum percolation rate. Both $L_g$ and $\gamma$ are free parameters. For those catchments where part of the discharge is in the deep subsurface and not captured by the stream gauge, the slowest groundwater component is reduced by a factor $f_3 < 1$ to account for deep percolation. Total runoff $Q_t$ from a pixel then consists of the following components:

\[
Q_t = Q_0 + Q_1 + Q_2 + Q_3 + f_3
\]

A total of 21 parameters exist for each pixel. Snow model parameters: $T_e$, $T_m$, $T_m$, $D$, $C_d$; soil moisture accounting parameters: $L_a$, $\beta$, $\alpha_b$, $L_b$, $L_p$; hillslope scale routing parameters: $k_0$, $k_1^*, \delta_1$, $k_2^*, \delta_2$, $k_3$, $L_1$, $L_{e^*}$, $\gamma$, $\alpha_p$, $f_3$.

### 4.2. Within-catchment routing

The outflow from the reservoirs, $Q_c$, is then convoluted by a transfer function which represents the runoff routing in the streams within each of the catchments. As a transfer function, a linear storage cascade with the parameters $n_c$ (number of reservoirs) and $k_c$ (time parameter of each reservoir) is used. For parsimony, it is assumed that these parameters do not vary spatially within each catchment but do vary with time to reflect non-linearities in the within-catchment routing. Specifically, it was assumed that $k_c$ decreases stepwise above dry weather flow $Q_{w}$ and a high flow $Q_{h}$ to represent the increase in stream connectivity as runoff increases:

\[
k_c = \begin{cases} 
3 & \text{if } Q_t < Q_{w}
\end{cases}
\]

$$k_c = \begin{cases} 
\frac{Q_t}{Q_{w}} & \text{if } Q_{w} \leq Q_t \leq Q_{h}
\end{cases}
\]

$$k_c = \begin{cases} 
\frac{Q_t}{Q_{h}} & \text{if } Q_t > Q_{h}
\end{cases}
\]

with $f_3 < f_2 < 1$. The convolution is performed in the state space notation in a similar way as stream routing (see Section 4.3). The sum of this convoluted runoff over each direct catchment is used as the lateral inflow to the stream routing model of each river reach. A total of six within-catchment routing parameters exist for each catchment: $n_c$, $k_{c1}^*$, $f_{c1}$, $f_{c2}$, $Q_{w}$, $Q_{h}$.

### 4.3. Stream routing

In a similar way as catchment processes are formulated directly at the grid scale, stream routing processes are formulated directly at the reach scale by making use of a lumped routing model. The main advantage of lumped routing models in a flood forecasting context is numerical stability and computational efficiency. Also, as there are usually only a small number of model parameters involved, these can be tested for plausibility to maximise the reliability and credibility of the forecasting procedure. A linear storage cascade in the state space notation of Szolgay (2004) is used here with discharge dependent parameters. If one assumes that the input vector $U$ to each reservoir is constant within a time interval $(i, i - 1)$ of duration $\Delta t$,

$$S_i = F_{i-1} S_{i-1} + G_{i-1} U_{i-1}
$$

$$Q_t = H_i \cdot S_i
$$

where $S$ and $Q$ are the $(n_r \times 1)$ state vectors of reservoir storages and outflow with $n_r$ being the number of reservoirs. $H$ is an $(n_r \times n_u)$ matrix that contains the inverse of the time parameter $k_c$ in the diagonal

$$H = \left( \frac{1}{k_c}, \frac{1}{k_c^*}, \ldots, \frac{1}{k_c^*} \right) \cdot I
$$
between the river reaches. A total of seven stream routing pas-

\[ F(t, \zeta) = e^{-\Delta t/\kappa} \frac{\Delta t^2}{(t - \zeta)} I_k \quad (16) \]

\[ G(t, \zeta) = k - e^{-\Delta t/\kappa} \sum_{i=0}^{t-\zeta} \frac{\Delta t^2}{v!} k^i \quad (17) \]

for \( \epsilon \) greater than or equal to \( \zeta \), and \( F = 0, G = 0 \) for \( \epsilon \) less than \( \zeta \), where \( \epsilon \) and \( \zeta \) relate to the rows and columns of the matrices, respectively. The duration \( \Delta t \) of the time interval is 15 min. Inflow \( U_{i,t-1}^{(1)} \) to each reach is the outflow from the upstream reach. Lateral inflow from the direct catchments is added to the downstream node. To account for non-linear routing effects, \( k \) is allowed to vary as a function of the inflow to the river reach based on the concept of Becker and Kundzewicz (1987). Varying \( k \) is straightforward in the state space notation as Eqs. (15)–(17) are evaluated for each time step as a function of the states \( S \) and \( Q \) of the previous time step and the value of \( k \) that is consistent with the inflow between the previous time step and the current time step. Typically, \( k \) is expected to increase with discharge due to the non-linearity in the flow—resistance relationship but beyond bank full discharge, \( k \) is expected to decrease because of inundation into the flood plain. Hence \( k \) was assumed to be a piecewise linear function of the inflow to the reach, \( U_{i,t-1}^{(1)} \):

\[ k_i = k_i' \quad \text{if} \quad U_{i,t-1}^{(1)} < Q_{i0} \]

\[ k_i = k_i' f_1 \quad \text{if} \quad Q_{i0} \leq U_{i,t-1}^{(1)} \leq Q_{i1} \]

\[ k_i = k_i' f_2 \quad \text{if} \quad U_{i,t-1}^{(1)} > Q_{i1} \quad (18) \]

and \( k_i \) is linearly interpolated between mean annual discharge \( Q_{i0} \) and \( Q_{i1} \). \( Q_{i0} \) is bank full discharge. The number of reservoirs \( n_r \) is not allowed to vary with the inflow but does vary between the river reaches. A total of seven stream routing parameters exist for each river reach: \( n_r, k_i', f_1, f_2, Q_{i0}, Q_{i1}, Q_{i2} \).

The outflow hydrographs from the hydropower scheme (reservoirs A, B, C in Fig. 3) can significantly differ from the inflow hydrographs. As a part of the Kamp flood forecasting system the operators communicate a planned release discharge over the next 48 h to the forecasting centre. These planned releases are routed to the downstream forecast points as one of the forecast variants in the forecasting system. As an alternative variant, the effect of the future operation of the scheme on the flood hydrograph is simulated by a reservoir simulation model which is, however, beyond the scope of this paper.

4.4. Input data

For the development of the distributed model, data from a total of 35 raingauges were used. Out of these, 19 raingauges recorded at a time interval of 15 min, the others were daily gauges. Sixteen of the recording raingauges are telemetered and are used for the operational forecasting. At each time step, the raingauge data are interpolated to the 1 km grid supported by climatologically scaled radar information.

Additionally, at each time step, deterministic precipitation forecasts are made by the Austrian Meteorological Office (ZAMG). The forecasts are at 15 min temporal resolution over a lead time of 48 h and are estimated as two components. The first component is an observation-based extrapolation or nowcast of the interpolated precipitation field using motion vectors determined from consecutive fields. The second component is a weighted mean of the forecast fields of the ALADIN and ECMWF numerical weather prediction (NWP) models (Wang et al., 2006). The weighting reduces areal precipitation forecast errors by about 20–30% (Haiden et al., 2006). Another weighting function is used for a smooth transition between the two components (nowcast and NWP forecast) (Golding, 1998). Analyses of the forecast performance indicated that, in most cases, over the first 2–6 h of the forecast the nowcast had smaller errors than the NWP forecast combination. The weighting function was hence chosen in a way to give full weight to the nowcast during the first 2 h, decreases linearly to zero at 6 h, and remains at zero for larger lead times.

Air temperatures observed at eight stations are interpolated to the 1 km grid. The temperature forecasts are based on a combination of the station data with the ALADIN forecasts.

The interpolated precipitation and air temperature fields are used to drive the runoff model to estimate the state variables such as soil moisture, soil and groundwater reservoir storage and snow water equivalent at each time step. These state variables are used as the initial conditions for the flood forecasts. The forecasts of precipitation and air temperature are then used for the actual flash flood forecasts over the lead time of 48 h.

For the development of the distributed model, runoff data between 1990 and 2005 from a total of 12 stream gauges were used. These range in catchment size from 77 to 1493 km². Out of these, nine stream gauges are telemetered and are used in the operational forecasting model for the updating procedures.

5. Parameter identification

5.1. Identifying parameters of pixel scale model and within-catchment routing

As runoff data are a necessary, but not a sufficient, condition for identifying model parameters in a realistic way, numerous other hydrological response data were used. Rather than optimising an objective function (e.g. Parajka et al., 2007), the identification of parameters builds on the ‘dominant processes concept’ of Grayson and Blöschl (2000b) which suggests that, at different locations and different points in time, a small number of processes will dominate over the rest. In a first step, Hydrological Response Units (HRUs) were defined manually rather than by over laying GIS maps, allowing some interpretation of the understanding of the hydrology of the area to be introduced. The HRUs are urban.
areas, steep slopes open, steep slopes forest, hills open, hills forest, tablelands, saturated areas, areas with aquifers, and lakes and reservoirs. Each 1 km pixel was assigned to one particular HRU. *A priori* parameter values were then assigned for each pixel (grid element) based on preliminary analyses of observed streamflow hydrographs and piezometric heads in the catchment as well as field surveys. The various pieces of information were then combined in an iterative way to construct a coherent picture of the functioning of the catchment system, on the basis of which plausible parameters for the HRUs were chosen. Different sources of information have been used:

The runoff simulations of the individual steps were compared with runoff data, stratified by time scale and hydrological situations which is in line with the ‘dominant processes concept’. A seasonal analysis allows one to infer the magnitude of the evaporation parameters, the percolation parameter and the parameters of the slow groundwater components. In the context of the flood forecasts, the seasonal dynamics are important to estimate well the initial conditions of the forecasts, in particular the catchment soil moisture state as well as the snow distribution. An analysis of the event hydrograph shapes allows one to infer the characteristics of fast catchment response as well as the associated model parameters. The event analysis was stratified by event magnitude and event types, again following the ‘dominant processes concept’. Synoptic (large scale) and convective (small scale) events, snowmelt events, and rain-on-snow events have characteristic runoff dynamics of their own (see, e.g., Merz and Blöschl, 2003). An example of a convective event is shown in Fig. 5.

Fig. 5. Convective event at the Kamp at Zwettl (622 km$^2$) in August 2003.

March 1998. The main difference from the convective event is the soil moisture status which was much higher during the snowmelt event. It is important to examine these event types separately for the model to be able to represent a spectrum of hydrological situations. This is essential in an operational forecasting context were situations that are not represented in the data used for model development are likely to occur. For example, after implementing the model in the forecasting centre in early 2006, unusually large snowfalls occurred — much larger than any in the past 16 year of runoff data used — and the following melt event was simulated well.

Discussions with locals were another source of information used. For example, these discussions provided information on flow pathways during past floods. This information was used to test the plausibility of the model simulations. Fig. 7 shows maps of simulated overland flow at the beginning of a large event (the August 2002 flood) and 18 h later during the same event. At the beginning of the event, it is mainly the sealed areas and the near stream areas that contribute to the flooding in the north-west of the catchment while there is almost no overland flow in the remainder of the catchment. Eighteen hours later, the spatial pattern has changed. The precipitation intensity is somewhat lower and hence overland flow on the sealed areas is lower. Additional areas contribute to overland flow, particularly the gullies and the rolling hills of the west of the catchment. The shift in the runoff patterns during the event is consistent with the understanding of runoff processes obtained during the reconnaissance trips and through the discussions with locals. Other spatial information that was used to test the plausibility of the model were snow data. Fig. 8 shows an example of this test, simulated snow water equivalent on the left and observed snow depth interpolated from snow depth readings on the right. The field recognisance trips also provided information on soil moisture and water logging which was used in a similar way. As part of the model identification, the model structure was adjusted to what has been considered is the hydrological functioning of this particular catchment. For example, a bypass flow component was added to represent the fast drainage of the top soil during events as a result of percolation into heavily weathered bedrock (Eq. (6)). Other examples are the functional dependences of some of the parameters on soil moisture (Eqs. (9) and (10)). For illustration, the parameters obtained for catchment 1 (Fig. 3) are given in Tables 1 and 2.

5.2. Identifying parameters of stream routing model

A lumped routing model is numerically more efficient and more robust than a hydrodynamic model which may be important in an operational flood forecasting context. However, calibration of the model parameters is more important, particularly if the interest resides in forecasting extreme floods where the routing characteristics of the flood plain play an important role. For small to medium sized catchments such as the Kamp examined here, lateral inflows to the river reaches can be very significant, so changes in the shape of the hydrograph cannot be uniquely traced back to channel routing. A double strategy of identifying the parameters of the routing model was hence adopted here. The first part of the strategy involved analysis of runoff data to which the routing parameters were calibrated for individual events. To minimise the effect of
lateral inflow, the events used for calibration were convective events where the rain storm was limited to a relatively small area upstream of the upstream gauge of the reach analysed with very little or no rainfall in the direct catchment of the reach. The second part of the strategy involved calibration of the lumped routing model to the results of hydrodynamic models. The idea of this calibration was to combine the merits of the two approaches, i.e., robustness, numerical efficiency and a small number of model parameters that can be checked for plausibility on the one hand, and reliable extrapolation to large events and estimation for ungauged river reaches on the other hand.

In the Kamp catchment, the results of various one-dimensional and two-dimensional hydrodynamic models are available such...
as HEC RAS and the model of Nujic (1998). Scenarios include events that have actually occurred plus floods associated with return periods from 30 to 1000 years, where a prescribed hydrograph shape was used. These different event magnitudes allowed an assessment of the non-linearity in the routing processes which translates into a change of routing parameters with discharge in the model used here. Examples of the calibration of the lumped routing model to the results of a 2D hydrodynamic model for the reaches between Rosenburg and Stiefern, and Rosenburg and Zöbing are shown in Fig. 9. Within the accuracy to be expected from a flood forecasting model, the model fits are excellent. Very little would be gained from directly using the hydrodynamic model in the operational forecasts, provided the lumped model is calibrated well for the full range of flows to be expected at the sites.

The parameter estimates from the two sources were then combined to obtain a functional relationship between the routing parameters and discharge for each reach of the model. The typical pattern of the functional relationship between the routing time parameter with discharge is as follows: below mean annual flow $Q_0$, the routing parameter $f_i$ is approximately constant and decreases with increasing discharges because of increasing flow velocities. Beyond bank full discharge, $Q_2$, inundation of the flood plain occurs and travel times are larger which is represented as a step increase in the routing parameter $f_2$ beyond the value of $f_1$. For reaches g, i and j there is very little inundation as the banks are steep, so $f_2$ and $f_0$ are identical. Examples of the routing parameters for reaches a and i are given in Table 3.

6. Operational use and real time updating

The complete spatially distributed model was comprehensively tested in a simulation mode for all the available stream gauges. Observed historical rainfall input was used for these tests which involved both analyses at the event scale and the seasonal scale. Examples of the simulation mode tests at the seasonal scale are shown in Fig. 10 for three gauges in the Upper Kamp area (Zwettl at Zwettl, Kamp at Zwettl and Purzelkamp at Rastenberg). The model represents the streamflow dynamics very well. However, the performance of the model hinges on the accuracy of the rainfall data. Minor biases in rainfall may accumulate over weeks and months and produce streamflow estimates that are not as accurate as those shown in Fig. 10. In particular, in the real time mode of the forecasting system, only 16 raingauges are available while the model development was based on 35 raingauges. The result of small biases in rainfall is biases in the simulated soil moisture which in turn will affect the forecasts as runoff generation very strongly depends on antecedent soil moisture in the Kamp catchment. Other error sources are possible biases in the estimation of evaporation which will affect estimated soil moisture in a similar way. To minimise the forecast uncertainties, two updating algorithms have been used.

The first updating algorithm adjusts the catchment soil moisture state by making use of runoff data in a real time mode. The rationale of this is that runoff is an excellent indicator of the catchment soil moisture state. An updating method widely used is the Kalman Filter which consists of weighting measurements and simulation, the weight
Table 1
Parameter values for the rainfall-runoff model for catchment 1 (Fig. 3)

<table>
<thead>
<tr>
<th>HRU</th>
<th>( T_i ) (s)</th>
<th>( C_i )</th>
<th>( P_i ) (mm)</th>
<th>( \beta )</th>
<th>( \alpha_i ) (mm)</th>
<th>( \delta_i ) (h)</th>
<th>( \delta_i' ) (h)</th>
<th>( \delta_i'' ) (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban areas</td>
<td>1.0</td>
<td>20</td>
<td>1.0</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.0</td>
<td>4.8</td>
</tr>
<tr>
<td>Steep slopes open</td>
<td>1.5</td>
<td>2.0</td>
<td>1.0</td>
<td>1.8</td>
<td>0.5</td>
<td>1.5</td>
<td>0.1</td>
<td>4.8</td>
</tr>
<tr>
<td>Hills forest</td>
<td>1.5</td>
<td>8.0</td>
<td>5.0</td>
<td>10</td>
<td>0.5</td>
<td>1.5</td>
<td>0.2</td>
<td>4.8</td>
</tr>
<tr>
<td>Tablelands</td>
<td>3.0</td>
<td>5.0</td>
<td>3.0</td>
<td>1.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Table 2
Within-catchment routing parameters for catchment 1 (Fig. 3)

<table>
<thead>
<tr>
<th>( n_i )</th>
<th>( k_i ) (h)</th>
<th>( f_i )</th>
<th>( f''_i )</th>
<th>( Q_s ) (m³/s)</th>
<th>( Q_{s,0} ) (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.5</td>
<td>0.67</td>
<td>0.2</td>
<td>8</td>
<td>50</td>
</tr>
</tbody>
</table>

(or Kalman gain) being a function of the measurement error and the model error. Originally, the Kalman Filter has been designed for linear models. Extended variants such as the extended Kalman Filter and the Ensemble Kalman Filter (EnKF) are applicable to non-linear models. In the EnKF, a number of realisations or ensemble members are used to estimate the Kalman gain (Evensen, 1994; Madsen and Cañizares, 1999). The advantages of the Ensemble Kalman Filter are that no linearisation is needed and it tends to be numerically robust. For the flood forecasts at the Kamp, the Ensemble Kalman Filter has hence been used. Runoff has been chosen as the state vector and, for consistency with the usual notation, is denoted by \( x \) here. The measurement error is the error in runoff measurements, the model error is attributed to the error in rainfall and evaporation input as these errors were assumed to be larger than the structural errors of the runoff model. The model \( \Phi(x) \) is now applied to each of the \( M \) members of the ensemble to estimate the state \( x \):

\[
x_{m,j} = \Phi(x_{m,j-1}, u_i + \epsilon_{m,i}), \quad m = 1, 2, \ldots, M
\]

where \( x_{m,j} \) is the runoff of ensemble member \( m \) at time step \( i \), \( x_{m,j-1} \) is the runoff at the previous time step, superscript \( f \) stands for forecast, superscript \( a \) stands for analysed, \( u_i \) is the model input (precipitation, evaporation) and \( \epsilon_{m,i} \) is the model error which is randomly drawn from a normal distribution with zero mean and model error covariance \( V_i \). As an \( a \ priori \) forecast, the mean value of the ensemble forecasts is adopted:

\[
x_f = \bar{x} = \frac{1}{M} \sum_{m=1}^{M} x_{m,i}
\]

The error covariance matrix \( P_f \) of the forecast is estimated from the ensemble forecasts as:

\[
P_f = S_f (S_f)^T
\]

with

\[
s_{f,m} = \frac{1}{\sqrt{M-1}}(x_{m,i} - \bar{x})
\]

where \( s_{f,m} \) is the \( m \)th column of \( S_f \). In a next step, the measurements \( z_i \) of runoff are contaminated by a measurement error \( \eta_{m,i} \) to generate an ensemble of \( M \) possible measurements:

\[
z_{m,i} = z_i + \eta_{m,i}, \quad m = 1, 2, \ldots, M
\]

where \( \eta_{m,i} \) is randomly drawn from a normal distribution with zero mean and covariance \( W_i \). Each ensemble member \( x_{m,i} \) is then updated according to
Routing parameters for reaches a and i (Fig. 3)

<table>
<thead>
<tr>
<th>Reach</th>
<th>n_s</th>
<th>k_s (h^-1)</th>
<th>f_1</th>
<th>f_2</th>
<th>Q_s (m^3/s)</th>
<th>Q_1 (m^3/s)</th>
<th>Q_2 (m^3/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>6</td>
<td>0.67</td>
<td>0.6</td>
<td>0.8</td>
<td>40</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>i</td>
<td>6</td>
<td>0.67</td>
<td>0.5</td>
<td>0.5</td>
<td>80</td>
<td>110</td>
<td>110</td>
</tr>
</tbody>
</table>

While updating methods commonly used in real time flood forecasting (e.g. Gutknecht, 1991) update runoff generation during events, the procedure presented here updates the entire model. A Kalman Filter (termed MOS or model output statistics) that updates runoff directly. Error analyses were performed from which the autocorrelation lag of events was found as 4 h. The entire model has then been set up in a forecasting mode that involves precipitation forecasts as well as both real time updating algorithms. An example of the output of the model is shown in Fig. 12. For clarity of presentation, the forecasts are shown at three hourly intervals while the operational forecasts are produced at quarter hourly intervals. During August 15, some precipitation has been forecasted but as soil moisture is relatively low, runoff hardly increases. On August 16 at 0:00, 20 mm of rainfall has been predicted over the following 12 h while 45 mm has actually been observed. The rising limb is hence underestimated. At 3:00 there is a similar underestimation but at 6:00, 20 mm of rainfall has been predicted and 19 mm observed. The rising limb is hence predicted very accurately. At the following time steps rainfall is smaller, so the dynamics of the hydrograph are mainly controlled by routing, and the predictions are accurate.

To examine the contributions of the individual model components to the forecast accuracy, comprehensive tests were performed. As this study is concerned with a flood forecasting system, forecast errors during floods were of interest rather than errors during medium and low flows. The forecast errors e_j were estimated for five major flood events on record as:

\[ e_j = \frac{1}{i_2 - i_1} \sum_{i=i_1}^{i_2} \left| \bar{Q}_j - Q_i \right| / Q_i \]  

where e_j is the mean normalised absolute error in per cent for lead time j, Q_1 is runoff at time step i that is forecasted with a lead time of j, Q_i is the observed runoff at time step i, and i_1 and i_2 are the time steps of the beginning and the end of the analysis interval, respectively. Fig. 13 shows the averages of the forecast errors over the five events, separately for entire events (Fig. 13, left) and the rising parts of the hydrographs only (Fig. 13, right). Four variants were analysed. In the first variant (thick solid lines in Fig. 13) future precipitation was assumed to be known from interpolating station data and the initial conditions of soil moisture were not updated. The average errors of the entire flood events are about 15% and do not depend on the lead time. This is because this is a simulation problem (rather than a forecast problem). In the second variant (dashed lines in Fig. 13) the initial conditions of soil moisture were updated by the Ensemble Kalman Filter, again based on the assumption that future precipitations were known. The updating reduces the errors, particularly for the short lead times that are closer to the time the forecasts are calculated. The thin
solid lines in Fig. 13 show the variant in which, additionally, the forecasts were updated by the error model (MOS) which further decreases the forecast errors for lead times of less than 6 h. Six hours is the same order of magnitude as the average autocorrelation length of the forecast errors during floods which is 4 h. In the fourth variant, the two updating procedures were used in the same way but future precipitation was not assumed to be known; rather the precipitation forecasts were used (dashed dotted lines in Fig. 13). For the first hours, the errors are identical with the previous variant as the forecasts are controlled by routing and fallen precipitation. For larger forecast lead times, the errors increase significantly. This is related to the uncertainty of the precipitation forecasts as well as to the combined forecast approach. For short lead times, the errors are small as runoff routing is the most accurate model component and the updating procedures

Fig. 10. Test of the combined rainfall runoff and routing model in a simulation mode for three gauges in the Upper Kamp area: Purzelkamp at Rastenberg (95 km², catchment 5), Zwettl at Zwettl (269 km², catchments 2 and 4 and reach b), Kamp at Zwettl (622 km², catchments 1, 2, 3, 4 and reaches a and b in Fig. 3), October 2000—September 2001. Thick lines, observations; thin lines, simulations. Top panel shows catchment precipitation of the Kamp at Zwettl catchment.

Fig. 11. Top: runoff simulations with EnKF updating (thin solid line) and without updating (dashed line). Bottom: cumulative errors of the two simulations. Kamp at Zwettl (622 km², catchments 1, 2, 3 and 4, and reaches a and b in Fig. 3) during November 2005—April 2006.
increase the accuracy additionally. Quantitatively, the average forecast errors range from 10% to 30% for 4–24 h lead times, respectively. It should be noted that this is an analysis of flood events only. The error statistics of forecasts over the entire year are much smaller.

Fig. 13 right shows similar analyses but for the rising limbs of the flood events only. Overall, the error patterns are similar although in absolute terms they are larger than those for the entire events, particularly for the variant of the precipitation forecasts (dashed dotted lines) which is the real time case. The larger errors are related to the fact that during the rising limb the uncertainty in precipitation affects flood flows more strongly than during the remainder of the hydrograph.

7. Conclusions

This paper has presented a strategy to model building where the model structure is defined at the model element scale and that is based on multi-source model identification and verification. Formulating the structure directly at the model element scale of 1 km² attempts to mimic the catchment functioning (Sivapalan, 2005) instead of using point scale relationships that may not be relevant in the presence of heterogeneity. Also, the approach allows parameters to be interpreted at the element scale, and it is numerically efficient. The multi-source model identification and verification strategy uses a multitude of hydrological response data, as available, in addition to runoff data. The strategy builds on the ‘dominant processes concept’ of Grayson and Blöschl (2000b) which suggests that, at different locations and different points in time, a small number of processes will dominate over the rest, such as fast runoff as a result of convective storms and subsurface driven runoff associated with snowmelt. It is important to examine these event types separately for the model to be able to represent a spectrum of hydrological situations. This is essential in an operational forecasting context where situations that are not represented in the data used for model development are likely to occur. The approach adopted here is fundamentally different from the common procedure of minimising objective functions that does not allow such a stratification and hence is less likely to capture different hydrological situations. In a similar way as catchment processes are formulated directly at the grid scale, stream routing processes are formulated directly at the reach scale by making use of a lumped routing model. A double strategy of identifying the parameters of the routing model was adopted here. The first part of the strategy involved analysis of runoff data to which the routing parameters were calibrated for individual

Fig. 12. Top panel: 12 h forecasts of cumulative precipitation. Bottom: 12 h runoff forecasts. Kamp at Zwettl (622 km², catchments 1, 2, 3 and 4, and reaches a and b in Fig. 3) during August 2005. Solid lines, observations; dashed lines, forecasts.

Fig. 13. Average forecast errors at the Kamp at Zwettl (622 km²) for various model components. Five large flood events during 2002–2005. Left: errors over the entire flood events. Right: errors over the rising limbs only.
events. The second part of the strategy involved calibration of the lumped routing model to the results of hydrodynamic models. The idea of this calibration is to combine the merits of the two approaches, i.e., robustness, numerical efficiency and a small number of model parameters that can be checked for plausibility on the one hand, and reliable extrapolation to large events and estimation for ungauged river reaches on the other hand.

The approach proposed here is in line with some of the steps in development and evaluation of models suggested by Jakeman et al. (2006). These steps include model choice with a focus on the nature of the data used to construct and test the model, providing a strong rationale for the choice of model family, and serious analysis, testing and discussion of model performance. However, the strategy proposed here differs from the one of Jakeman et al. (2006) in that, here, less emphasis is on formal calibration and more emphasis is on understanding the behaviour of the model vis-à-vis perceived catchment behaviour.

Comprehensive model tests indicate that not only is the methodology feasible but also are the simulation results accurate for a range of hydrological situations and a range of temporal scales. However, the performance of the model hinges on the accuracy of the rainfall data, and biases in rainfall may translate into biases in soil moisture and hence diminished forecast accuracies. Two updating algorithms have hence been implemented that use runoff in a real time mode. The first updating algorithm adjusts the catchment soil moisture state by the Ensemble Kalman Filter. The second updating algorithm exploits the autocorrelation of the forecast error and consists of an additive error model that updates runoff directly. The error analyses of five major events indicate that both algorithms improve the forecasts, the first over longer time scales and the second over shorter time scales. If the forecast lead times are longer than the catchment response time, the uncertainty of the precipitation forecasts dominates over the uncertainty of the hydrological model. For five major flood events, the average forecast errors range from 10% to 30% for 4–24 h lead times, respectively, but the errors are much smaller if the error statistics are evaluated over the entire year. The forecasting system has been implemented in the Kamp forecasting centre in early 2006 and has been in operational use since then.

Acknowledgements

Development of the forecasting model was funded by the State Government of Lower Austria and the EVN Hydropower Company, Austria. Financial support of the EC (project nr 037024, HYDRATE) is acknowledged. The authors would like to thank Dieter Gutknecht for numerous suggestions on this research as well as two anonymous reviewers for their valuable comments on the manuscript.

References


