Available online at www.sciencedirect.com
ScienceDirect
Environmental Modelling & Software 23 (2008) 464–478
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

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 km^2 Kamp catchment and range from 10% to 30% for 4–24 h lead times, respectively.

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,

* Corresponding author. Tel.: +43 1 58801 22315; fax: +43 1 58801 22399.
E-mail address: bloeschl@hydro.tuwien.ac.at (G. Blöschl).

1364-8152/$ - see front matter
doi:10.1016/j.envsoft.2007.06.010

"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 (Grayson 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 (Klemeš, 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. Grayson 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 (Grayson 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