Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Hydrology 326 (2006) 379–393
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Parameter estimation in stochastic rainfall-runoff models

Harpa Jonsdottir [a,*], Henrik Madsen [a], Olafur Petur Palsson [b]

[a] Department of Informatics and Mathematical Modelling, Bldg. 321 DTU, DK-2800 Lyngby, Denmark
[b] Department of Mechanical and Industrial Engineering, University of Iceland, Hjardarhaga 2-6, 107 Reykjavik, Iceland

Received 14 November 2003; revised 31 October 2005; accepted 9 November 2005

Abstract

A parameter estimation method for stochastic rainfall-runoff models is presented. The model considered in the paper is a conceptual stochastic model, formulated in continuous-discrete state space form. The model is small and a fully automatic optimization is, therefore, possible for estimating all the parameters, including the noise terms. The parameter estimation method is a maximum likelihood method (ML) where the likelihood function is evaluated using a Kalman filter technique. The ML method estimates the parameters in a prediction error settings, i.e. the sum of squared prediction error is minimized. For a comparison the parameters are also estimated by an output error method, where the sum of squared simulation error is minimized. The former methodology is optimal for short-term prediction whereas the latter is optimal for simulations. Hence, depending on the purpose it is possible to select whether the parameter values are optimal for simulation or prediction. The data originates from Iceland and the model is designed for Icelandic conditions, including a snow routine for mountainous areas. The model demands only two input data series, precipitation and temperature and one output data series, the discharge. In spite of being based on relatively limited input information, the model performs well and the parameter estimation method is promising for future model development.

Keywords: Conceptual stochastic model; Rainfall-runoff model; Parameter estimation; Maximum likelihood; Extended Kalman filter; Prediction and simulation

1. Introduction

All hydrological models are approximations of reality, and hence the output of a system can never be predicted exactly and the problem is how to achieve an acceptable and operational model.

The numerous hydrological models which already exist vary in their model construction, partly because the models serve somewhat different purposes. There are models for design of drainage systems, models for flood forecasting, models for water quality, etc. Singh and Woolhiser (2002) give a comprehensive overview of mathematical modelling of watershed hydrology, however, a brief overview will be given here. The HBV model, see Bergström (1975, 1995), is a standard model in the Scandinavian countries

* Corresponding author. Tel.: +354 5696051; fax: +354 5688896;
E-mail address: hj@os.is (H. Jonsdottir).

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2005.11.004

Nomenclature

a	Low pass filtering constant [1/*day*]
b	Constant in $\psi(\cdot)$, controlling the smoothness
b	Center of the threshold function $\phi(\cdot)$
b_1	Sharpness of the threshold function $\phi(\cdot)$
c:	Precipitation correction factor
f	Filtration from upper reservoir to lower reservoir [1/*day*]
K	Constant representing the base flow [*m*/*day*]
k	Constant in $\psi(\cdot)$, controlling the smoothness
k_1	Routing constant, from upper surface reservoir [1/*day*]
k_2	Routing constant, from lower surface reservoir [1/*day*]
M	Constant in $\psi(\cdot)$, controlling the upper limit value
N	Snow cover [*m*]
P	Measured precipitation [*mm*]
pdd	Positive degree day constant for melting [*m*/([°C]day)]
S_1	Upper surface water reservoir [*m*]
S_2	Lower surface water reservoir [*m*]
s_{11}	White noise process for the observations
$T(t)$	Measured temperature [°C]
$T_s(t)$	Low pass filtered temperature [°C]
Y	Measured discharge [*m*/*day*]
$d\omega_I$	One dimensional Wiener process
$\phi(x)$	Smooth threshold function (sigmoid function) $1/(1+\exp(b_0+b_1x))$
$\psi(x)$	Indicator function for the snow $\psi(x)=M\exp(-b\exp(-kx))$
σ_{ii}	Incremental covariance of the Wiener process

and has also been used around the globe. The HBV model has been classified as a semi-distributed conceptual model and is based on the theory of linear reservoirs. The model has a number of free parameters, which are found by calibration. The model presented here is in the spirit of the HBV model, but additionally, the approach suggested in this paper includes a procedure for fully automatic parameter estimation. The NAM/MIKE11/MIKE21 models, see e.g. Nielsen and Hansen (1973); Gottlieb (1980); Havnø et al. (1995), are used for flood forecasting in Denmark and other European countries. The NAM model is based on similar principles as the HBV model and a further development of the model led to the MIKE11 software package which is a one dimensional modelling system for simulation of flow, sediment-transport and water quality. MIKE21 is a two dimensional version. The TOPMODEL, Beven and Kirkby (1979), has been used in Great Britain. The model is a set of conceptual tools that can be used to reproduce the hydrological behaviour of the catchment area in a distributed or semi-distributed way. The parameters are physically interpretable and the watershed is classified by using the so-called topographic index. The SHE model is a physically based, distributed watershed modelling system, developed jointly by the Danish Hydraulic Institute, the British Institute of Hydrology and SOGREAH in France. The SHE model is widely used, see Abbott et al. (1986); Bathurst (1986); Singh and Woolhiser (2002); Jain et al. (1992); Refsgaard et al. (1992). The MIKE SHE model is a further development of the SHE modelling concept Refsgaard and Storm (1995) and it has been used in many European counties. The ARNO model, Todini (1996), is a semi-distributed conceptual model, and it is well known in Italy. Like the HBV and NAM models the Tank model, Sugawara (1995), is a model based on linear reservoirs and it has been used in Japan. The Xinanjiang model, Zhao and Liu (1995); Zhao (2002), is a distributed, basin model for use in humid and semi-humid regions where the evaporation plays a major role. The model has been widely used in China since 1980. In Canada, the WATFLOOD model Singh and Woolhiser (2002), is being used. The WATFLOOD model is a distributed hydrological model based on the GRU (Group Response Unit) concept, i.e. all similarly vegetated areas within a sub-watershed are grouped as one response unit. The NWS River Forecast system, based on the Sacramento

Model, Burnash (1995), is a standard model in the United States for flood forecasting and in Australia the RORB model, Layrenson and Mein (1995), is commonly employed for flood forecasting and drainage design.

Increased computer power and data storage capabilities have opened the possibility for working with more detailed distributed models. Hence, many of the recently developed models are physically based distributed models, and they are occasionally used together with GIS (Geographic Information Systems). These models both utilize a large amount of information, but can also provide various information, however, if the input data (or information) are not available the model is of little use.

Black box models have also been used for flood forecasting, starting with linear transfer function models in the beginning of 1970s and since then various kinds of linear and nonlinear models. In recent years neural network models have been popular. Sajikumar and Thandaveswara (1999) used an artificial neural network as a nonlinear rainfall-runoff model for the river Lee in the UK and for the river Thuthapuzha in India. Shamseldin (1997) used neural networks for rainfall-runoff modelling which was tested on six different catchment areas.

The main advantage of black box models in hydrology is that they are not as data demanding as the physical models; this refers to all kinds of physical information about the watershed as well as long record of flow and precipitation.

Some of the conceptual models are not very data demanding and it is important to work with those kind of models as well, i.e. models with few input data, few parameters and limited prior information. The parameters in a lumped conceptual model can be interpreted as some kind of an average over a large area, but in general the most likely parameter values cannot be given, and the final parameter estimation must, therefore, be performed by calibration against observed data. Refsgaard et al. (1992) stated that in principle the parameters in a physically based model can be estimated by field measurements, but such an ideal situation requires comprehensive field data, which cover all the parameters. This situation rarely occurs and the problem of calibration will arise. Because of the large number of parameters in a physically based model the parameter estimation can not be done by free optimization for all parameters, however, an over all parameter estimation is possible for simpler models.

The state space formulation and the Kalman filter has been used in hydrology for years, representing both black box models and grey box models, i.e. conceptual physical models where parameter values are estimated using data. Szollosi-Nagy (1976) used a state space formulation for on-line parameter estimation in linear hydrography using a FIR model (Finite Impulse Response model). Todini (1978) presented a threshold ARMAX model, formulated in a state space form and the parameters estimated off line, i.e. in a batch form. Refsgaard et al. (1983) reformulated the NAM model in a state space form where two of the model parameters were time varying i.e. on-line estimated. Haltiner and Salas (1988) used ARMAX models, both with off-line (batch) parameter estimation and on-line parameter estimation method in the SRM model, see also Martinec (1960); Martinec and Rango (1986). All the above-mentioned models are formulated in a discrete time. In Georgakakos (1986a,b) rather large physical models are presented using a state space formulation. However, the model parameters are constants and not estimated. In Georgakakos et al. (1988) the Sakramento model (org. in Burnash et al. (1973)), is modified and formulated in state space form and some of the parameters are estimated. Rajaram and Georgakakos (1989) represent a model for acid decomposition in a lake watershed system formulated in a continuous-discrete state space form, and they estimated the parameters. Lee and V.P. Singh (1999) applied an on-line estimation to the Tank model (see e.g. Sugawara (1995)), for single storm at a time, calibrating the initial states manually. Lee and V.P. Singh (1999) also gave a short overview of application of the Kalman filter to hydrological problems upto 1999. In Ashan and O'Connor (1994) a general discussion about the use of Kalman filter in hydrology is found.

In the following a stochastic lumped, conceptual rainfall-runoff model is developed. The model is formulated as a continuous-discrete time stochastic state space model. The dynamics are described by stochastic differential equations and the observations are described by equations relating the discrete time observations to the state variables at time points where observations are available. The main advantage of this model formulation is that the stochastic part permits a description of both the model and the measurement uncertainty, and hence more rigorous statistical methods can be used for parameter optimization. Furthermore, the stochastic modelling approach allows a much simpler model structure than a deterministic modelling approach since some of the variations observed in data are described by the stochastic part of the model. The model presented is a watershed model designed for discharge forecasting. It is a simple lumped reservoir model with two input variables, precipitation and temperature and one output variable, the discharge. Because of the simplicity of the model and few parameters it is possible to estimate all the parameters including threshold parameter in the snow routine and the system noise, which often has been difficult to identify in hydrology. The method suggested for parameter estimation (in batch form) is a maximum likelihood method, where the one step ahead prediction errors required for evaluating the likelihood function are evaluated using the Kalman filter technique. Moreover, the state space formulation allows the model to be used for simulation as well, however, good simulation results require different parameter values, Kristensen et al. (2004). Parameter values, which are suitable for simulation can be achieved by fixing the system noise to a small value and then estimating the remaining parameters. Conversely good prediction results are obtained by using parameter values where all the parameters have been optimized, including the parameters describing the system noise.

The paper is organized as follows. In Section 2 the availability of data is discussed. The model is described in Section 3 and the method for parameter estimation is described in Section 4. Section 5 includes discussion of some estimation principles. In Section 6 results are demonstrated and in Section 7 conclusions are drawn.

2. The data

The goal is to develop a model, which can be used in mountainous areas with snow accumulation. Such areas are often thinly populated and the meteorological observatories are often rather spread. However, there is a need for flood forecasting for various reasons, such as warning related to the spring floods or for operational planning of hydropower plants. The data used in this project originates from Iceland which has in general only mountainous catchment areas and it certainly is thinly populated with only 2.8 inhabitants per km^2, and only a few meteorological observatories exists.

Precipitation is the main input for hydrological models as the precipitation and the evaporation control the water balance. In Iceland, the evaporation plays only a minor role, but the precipitation is important. However, there is a shortage of good precipitation data, which indeed has an effect on the prediction performance of the model. The poor quality of precipitation data arises both from a rain gauges bias towards too small values, and a limited number of rain gauge measurement stations. Due to the influence of wind the amount of precipitation measured is an underestimate of the ‘ground true’ precipitation. Unfortunately, no experiments have been made in Iceland in order to develop models to adjust for this bias. Experiments, like for instance the Nordic project in Jokioinen in Finland during the years 1987–1993 (Førland et al. (1996)), had the purpose of developing models to describe the underestimate of the different rain gauges depending on weather condition. This experiment is of little use here since the wind speed in Iceland in general is much higher than in Finland. Furthermore, most of the meteorological stations in Iceland are located along the coastal line in the inhabited areas and most rivers, especially the larger ones, stretch far into the country and have thus watershed in high mountainous areas. Occasionally, there are no meteorological stations in the whole watershed and if any they are typically located near the coast. Precipitation lapse rate is also difficult to track since in practice the lapse rate depends highly on wind speed and direction in the mountainous areas.

The discharge data are calculated from water level data using the $Q-h$ formula. The errors of the discharge data are caused both by uncertainty of the water level data and uncertainty of the parameters in $Q-h$ formula. The errors of the water level measurements more or less only occur during the winter because of the icing, which causes the water level to rise even though the flow of water is not increasing. This has to be corrected manually.

In Iceland few discharge measurements with very large discharge exist. This is due to the fact that even though annual spring floods occur, it is difficult to predict peaks several days ahead, and the number of rivers that can be measured simultaneously is limited.

The discharge used in this project originates from the river Fnjóská in Northern Iceland, see Fig. 1. The river is a direct runoff river with no glacier in the watershed, whereas many larger rivers in Iceland have a glacier factor. The watershed is about 1132 km^2, the altitude range is between 44 and 1084 m, and 54% of the catchment area is above 800 m. The catchment area is dominated by grit and rocks; a very small part of the region in the valley is copse and grassland. A meteorological observatory is located in the watershed, at Lerkihlid, about 20 km from the outlet of the watershed, and it is situated 150 m above sea level. No meteorological observatory is located in the highlands which could have given information about the weather condition in the catchment area there. The data used are diurnal averages of the discharge, diurnal averages of the temperature and the total precipitation for the past 24 hours. Fig. 2 shows the discharge, the temperature and the precipitation for the whole period of 8 years, starting 1st of September 1976 and ending 31st of August 1984.

3. The stochastic model

The stochastic model proposed is a simple smooth threshold model with a snow routine, and the basic idea is similar to the idea behind the HBV and NAM models, see Bergström and Fossman (1973); Bergström (1975); Nielsen and Hansen (1973), and

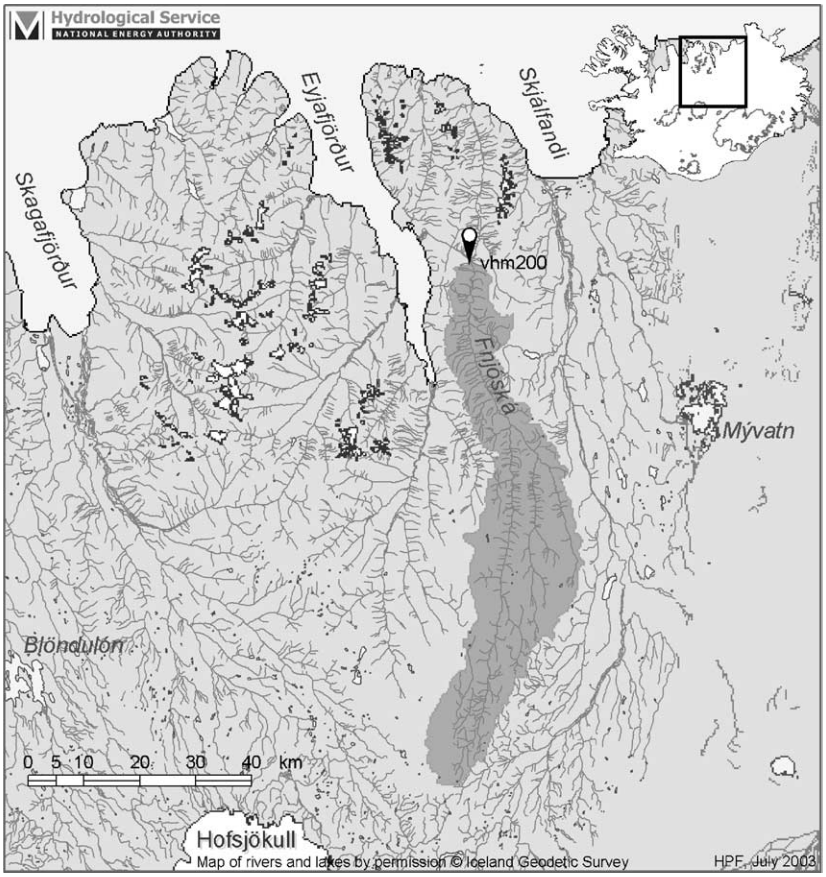

Fig. 1. The watershed of the river Fnjóská is about 1132 km^2 and located in Northern Iceland.

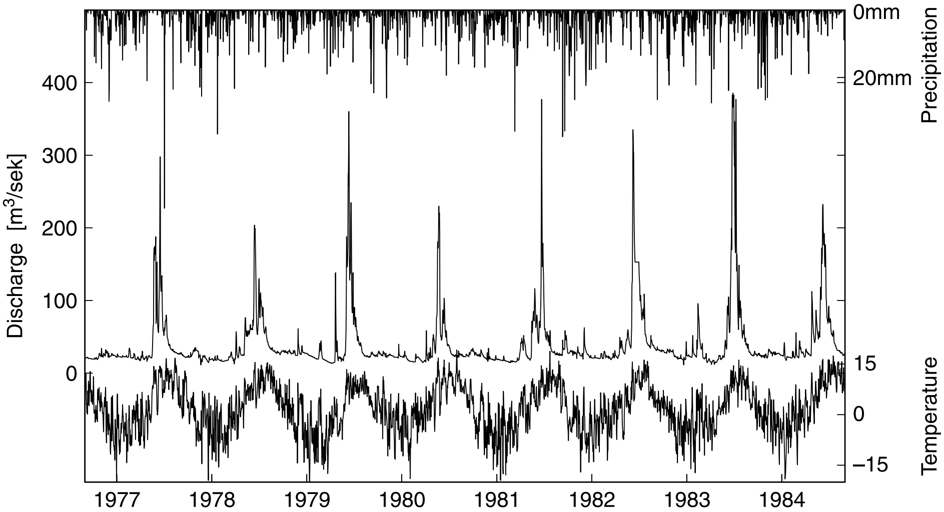

Fig. 2. The data series for discharge, temperature and precipitation starting 1st of September 1976 and ending 31st of August 1984.

Gottlieb (1980). A diagram of the model structure is shown in Fig. 3.

The water is stored in reservoirs and the outflow of the reservoirs are routed to the stream with different time constants. The main distinction between the non-stochastic HBV and NAM models and the stochastic model suggested here is that the water flow is modelled as a function of only precipitation and temperature, and there are no factors for evaporation and infiltration into the ground. On the other hand no manual calibration is required since the stochastic model allows for statistical methods for parameter estimation. The total precipitation is divided into snow and rain using a smooth threshold function $\phi(T(t))$, where $T(t)$ is the air temperature. The threshold function is formulated as the sigmoid function

$$\phi(T(t)) = \frac{1}{1+\exp(b_0 - b_1 T(t))} \tag{1}$$

The same smooth threshold is used for the melting process, where the melting $M(t)$ is formulated using the positive degree day method

$$M(t) = \text{pdd}\ \ T(t)\ \ \phi(T(t)) \tag{2}$$

where pdd is the positive degree day constant, which typically is calibrated. No attempt is made to model the actual physical process of melting, i.e. the fact that in the beginning of the melting process the water first stays in the snow pack and is not released until the snow pack is wet enough. However, in order to take this into account the temperature $T(t)$ is low pass filtered

$$\mathrm{d}T_s(t) = [-aT_s(t) + aT(t)]\mathrm{d}t + \mathrm{d}w(t) \tag{3}$$

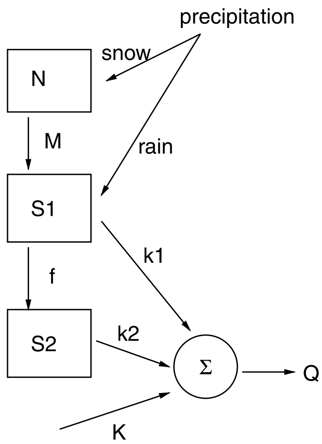

Fig. 3. The model structure. The precipitation is divided into snow and rain, N is snow a container, S_1 and S_2 are upper and lower surface reservoirs, M is melting, f is filtration between the reservoirs and k_1 and k_2 are the routing constants. K is a constant representing the base flow and Q is the discharge.

The consequence is that a single warm day does not give as much impact as a warm day followed by another warm day.

The suggested stochastic state space model is:

$$dT_s(t) = [-aT_s(t) + aT(t)]dt + \sigma_1 dw_1(t) \quad (4)$$

$$dN(t) = [-pddT_s(t)\phi(T_s(t))\psi(N(t)) + (1-\phi(T_s(t)))cP(t)]dt + \sigma_2 dw_2(t) \quad (5)$$

$$dS_1(t) = [pddT_s(t)\phi(T_s(t))\psi(N(t)) - (f+k_1)S_1(t) + (\phi(T_s(t))cP(t))]dt + \sigma_3 dw_3(t) \quad (6)$$

$$dS_2(t) = [fS_1(t) - k_2S_2(t)]dt + \sigma_4\, dw_4(t) \quad (7)$$

$$Y(t) = k_1S_1(t) + k_2S_2(t) + K + e_1(t) \quad (8)$$

where N is the amount of snow in the snow container in meters and the function ψ is a smooth indicator function, controlling whether there is snow to melt or not,

$$\psi(N) = M\exp(-b\exp(-kN)) \quad (9)$$

S_1 and S_2 are water content reservoirs, f is filtration from the upper surface reservoir to the lower surface reservoir, k_1 and k_2 are the routing constants, and c is a precipitation correction factor. $\sigma_1,\ldots,\sigma_4$ are constants representing the variances of the system noise and the noise terms $dw_1(t),\ldots,dw_4(t)$ are assumed to be independent standard Wiener processes and all are assumed independent of measurement noise $e_1(t)$. The base flow is assumed to be constant. An extension of the model with a ground water reservoir would improve the physical reality of the model and it might be a task for future research.

4. Parameter estimation

In this section a maximum likelihood method for estimation of the parameters of the continuous–discrete time stochastic state space models is outlined. The procedure is implemented in a program called CTSM (Continuous Time Stochastic Modelling), and for a further description of the mathematics and numerics behind the program, see Kristensen et al. (2003), and Kristensen et al. (2004).

The hydrological model described by Eq. (4)–(8) is a continuous-discrete time stochastic state space model. The stochastic differential equations describe the dynamics of the system in continuous time as stated by Eq. (4)–(7), and the algebraic equation Eq. (8) describes how the measurements are obtained as a function of the state variables at discrete time instants. Using a slightly different and more compact notation, the mathematical formulation of the continuous-discrete time stochastic state space model is

$$dx_t = f(x_t, u_t, t, \theta)dt + \sigma(u_t, t, \theta)d\omega_t \quad (10)$$

$$y_k = h(x_k, u_k, t_k, \theta) + e_k \quad (11)$$

where $t \in \mathbb{R}_+$ is time, $x_t \in \mathbb{R}^4$ is a vector of the state variables (since $x_t = [T_s(t), N(t), S_1(t), S_2(t)]^T$), $u_t \in \mathbb{R}^2$ is a vector of the input variables (since $u_t = [T(t), P(t)]$). $\theta \in \mathbb{R}^p$ is a vector of the unknown parameters. The vector $y_k \in \mathbb{R}$ is a vector of measurements (i.e. the discharge). The notation $x_k = x_{t=t_k}$ and $u_k = u_{t=t_k}$ is used. Furthermore, the functions $f(\cdot) \in \mathbb{R}^4$, $\sigma(\cdot) \in \mathbb{R}^{4\times4}$ and $h(\cdot) \in \mathbb{R}$ are nonlinear functions, ω_t is a 4-dimensional standard Wiener process and $\boldsymbol{e}_k \in N(0, \boldsymbol{S}(\boldsymbol{u}_k, t_k, \boldsymbol{\theta}))$ is a Gaussian white noise process. With this model formulation the parameters are constants and estimated *off-line* or in a batch form. An *on-line* estimation of (some) parameters is possible by extending the state vector with the relevant parameters. As mentioned, Eq. (10) is known as the system equation and Eq. (11) is known as the measurement equation.

The measurements y_k are in discrete time. It is well known that the likelihood function for time series models is a product of conditional densities (see e.g. Restrepo and Bras (1985)). By introducing the notation

$$\mathcal{Y}_k = (\boldsymbol{y}_k, \boldsymbol{y}_{k-1}, \ldots, \boldsymbol{y}_1, \boldsymbol{y}_0) \quad (12)$$

where $(y_k, y_{k-1}, \ldots, y_1, y_0)$ is the time series of all measurements up to and including the measurement at time t_k. The likelihood function can be written as

$$L(\boldsymbol{\theta}; \mathcal{Y}_N) = \left(\prod_{k=1}^{N} p(\boldsymbol{y}_k|\mathcal{Y}_k, \boldsymbol{\theta})\right) p(\boldsymbol{y}_0|\boldsymbol{\theta}) \quad (13)$$

In order to obtain an exact evaluation of the likelihood function, the initial probability density $p(y_0|\theta)$ must be known and all subsequent conditional densities can be determined by successively solving Kolmogorov's forward equation and applying Bayes's rule, Jazwinski (1970). This approach is not feasible in practice. However, since the diffusion term in the system equation, Eq. (10), in the continuous–discrete state space model is a Wiener process, which is independent of the state variables, and the error term in the measurement equation, Eq. (11), it follows that for LTI (Linear Time Invariant) and LTV (Linear Time Variant) models the conditional densities are Gaussian, Jazwinski (1970). In the NL (Non Linear) case it is reasonable to assume that under suitable regularity conditions, the conditional densities can be well approximated by the Gaussian distribution, Kristensen et al. (2004). This assumption can be tested after the estimation e.g. by considering the sequence of residuals. Thus, assuming that the conditional densities $p(\boldsymbol{y}_k|\mathcal{Y}_k,\theta)$ are Gaussian the likelihood function becomes

$$L(\boldsymbol{\theta}|\mathcal{Y}_N) = \left(\prod_{k=1}^{N} \frac{\exp\left(-\frac{1}{2}\varepsilon_k^T \boldsymbol{R}_{k|k-1}^{-1}\varepsilon_k\right)}{\sqrt{\det(\mathbf{R}_{\mathrm{k|k-1}})}\left(\sqrt{2\pi}\right)^l}\right) p(\boldsymbol{y}_0|\boldsymbol{\theta}) \tag{14}$$

where $\varepsilon_k = \boldsymbol{y}_k - \hat{\mathrm{y}}_{k|k-1} = \boldsymbol{y}_k - E\{\boldsymbol{y}_k|\mathcal{Y}_{k-1},\boldsymbol{\theta}\}$ is the one step prediction error and $\hat{\boldsymbol{R}}_{k|k-1} = V\{\boldsymbol{y}_k|\mathcal{Y}_{k-1},\boldsymbol{\theta}\}$ is the associate conditional covariance. For given parameters and initial states, ε_k and $\boldsymbol{R}_{k|k-1}$ can be computed by means of a Kalman filter in the linear case or an extended Kalman filter in the nonlinear case. The continuous-discrete Kalman filter equations are (Kristensen et al. (2004) or Jazwinski (1970)):

$$\hat{\mathrm{y}}_{k|k-1} = \boldsymbol{h}(\hat{\mathrm{x}}_{k|k-1}, \boldsymbol{u}_k, t_k, \theta) \quad \text{(Output prediction)} \tag{15}$$

$$\boldsymbol{R}_{k|k-1} = \boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^{\mathrm{T}} + \boldsymbol{S} \quad \text{(Output variance prediction)} \tag{16}$$

$$\varepsilon_k = \boldsymbol{y}_k \hat{\mathrm{y}}_{k|k-1} \quad \text{(Innovation)} \tag{17}$$

$$\mathrm{K}_k = \mathrm{P}_{k|k-1} C^{\mathrm{T}} \mathrm{R}_{k|k-1} \quad \text{(Kalman gain)} \tag{18}$$

$$\hat{\mathrm{x}}_{k|k} = \hat{\mathrm{x}}_{k|k-1} + \boldsymbol{K}\boldsymbol{\varepsilon}_k \quad \text{(Updating)} \tag{19}$$

$$\boldsymbol{P}_{k|k} = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}_k \boldsymbol{R}_{k|k-1} \boldsymbol{K}_k^{\mathrm{T}} \quad \text{(Updating)} \tag{20}$$

$$\frac{\mathrm{d}\hat{\mathrm{x}}_{t|k}}{\mathrm{d}t} = \boldsymbol{f}(\hat{\mathrm{x}}_{t|k}, \boldsymbol{u}_k, t_k, \theta) \quad t \in [t_k, t_{k+1}] \tag{21}$$

(State prediction)

$$\frac{\mathrm{d}\hat{\boldsymbol{P}}_{t|k}}{\mathrm{d}t} = \boldsymbol{A}\hat{\mathrm{P}}_{t|k} + \boldsymbol{P}_{t|k}\boldsymbol{A}^{\mathrm{T}} + \sigma\sigma^{\mathrm{T}} \quad t \in [t_k, t_{k+1}] \tag{22}$$

(State var. pred.)

where

$$\boldsymbol{A} = \frac{\partial f}{\partial x_t}\Big|_{\boldsymbol{x}=\hat{\mathrm{x}}_{k|k-1},\boldsymbol{u}=\boldsymbol{u}_k,t=t_k,\boldsymbol{\theta}}, \qquad \mathrm{C} = \frac{\partial h}{\partial x_t}\Big|_{\boldsymbol{x}=\hat{\mathrm{x}}_{k|k-1},\boldsymbol{u}=\boldsymbol{u}_k,t=t_k,\boldsymbol{\theta}} \tag{23}$$

and

$$\sigma = \sigma(u_k, t_k, \theta), \qquad S = S(u_k, t_k, \theta) \tag{24}$$

Given information upto and including time t the prediction $\hat{\mathrm{x}}_{k+1|k} = E\{\boldsymbol{x}_{t_{k+1}}|\boldsymbol{x}_{t_k}\}$ and $\boldsymbol{P}_{k+1|k} = E\{\boldsymbol{x}_{t_{k+1}}\boldsymbol{x}_{t_{k+1}}^{\mathrm{T}}|\boldsymbol{x}_{t_k}\}$ are needed for the Kalman filter equations. Eq. (22) is a linear differential equation, which can be solved analytically. This analytical solution is used to calculate the 'initial' problem $P_{k+1|k}$. On the other hand equation Eq. (21) is nonlinear with a nontrivial solution. The software CTSM offers three options for handling this:

Linearization by first order Taylor, the linear equation is solved analytically, iteratively in a subsampled interval.

Numerical solution of the ODE equation Eq. (21) by using a Predictor/Corrector scheme, also occasionally referred to as Gears method or Adams method (see e.g. Dahlquist and Björck (1988)).

Numerical solution of the ODE equation Eq.(21) by using BDF (Backward Difference Formula) (see e.g. Dahlquist and Björck (1988)).

For a detailed description of all the methods see Kristensen et al. (2003). The BDF formula demands a Newton-like method for solving a nonlinear zero-point equation and is thus the most time consuming algorithm. However, for stiff systems the BDF formula is the most reliable method (Dahlquist and Björck (1988)) and consequently this option has been used in the following.

The hydrological model described by Eq. (4)–(8) is a model with four states, the low pass filtered temperature T_S, the snow container N, and upper and lower surface containers S_1 and S_2. It is found important to include the snow container as a state variable and not to treat the water from snow melt as an input as some times is done, e.g. Refsgaard et al. (1983). Treating the snow-melt as an input requires manual control of the snow balance. However, by including the snow container into the state vector leads to numerical complications. The system is singular during summer, fall, and most of the winter or more accurately while snow is not melting. The system is non-stiff during spring floods, i.e. when melting is significant and extremely stiff during the transition points in between.

On of the strengths of using the maximum likelihood method for parameter estimation it that it follows from the central limit theorem that the estimator $\hat{\theta}$, is asymptotically Gaussian with mean θ and covariance

$$\hat{\Sigma}_{\hat{\theta}} = \boldsymbol{H}^{-1} \tag{25}$$

where the information matrix H is given by

$$h_{ij} = -E\left\{\frac{\partial^2}{\partial\theta_i\partial\theta_j} ln(L(\boldsymbol{\theta}|\mathcal{Y}))\right\} \quad i,j = 1,\ldots,p. \tag{26}$$

An approximation of H can be obtained by evaluating $h_{ij} = \partial^2/(\partial\theta_i\partial\theta_j)\ln(L(\boldsymbol{\theta}|\mathcal{Y}))$ in the point $\theta = \hat{\theta}$. The asymptotic Gaussianity of the estimator also allows marginal t-test to be performed like a test for the hypothesis:

$$H_0 : \theta_j = 0 \quad H_1 : \theta_j \neq 0 \tag{27}$$

The one step prediction of the output $\hat{y}_{k|k-1}$, the state update $\hat{x}_{k|k}$, and the state prediction $\hat{x}_{k|k-1}$, corresponding to each time instant t_k are generated by the (extended) Kalman filter. A simulation $\hat{x}_{t|0}$ and $\hat{y}_{t|0}$ can be obtained using the (extended) Kalman filter equations without the updating.

5. Some comments on parameter estimation in hydrological models

Model calibration has been a topic in hydrology since the computer evolution in 1960 and since then parameter optimization has been practiced. A solution to a rainfall-runoff prediction problem is to optimize the parameters such that the model performs the ‘best’ fit to data. On the other hand what is best the fit to data? This is a selective question with a selective answer. Best fit can be such that the sum of squared simulation error is minimized, or the sum of squared prediction error is minimized, or models, which conserve the water balance best, or those who have the best timing of flood peaks. In the recent years multi objective calibration and Pareto optimality have been applied in rainfall-runoff modelling, see e.g. Madsen (2000). However, two estimation methods have frequently been used in hydrology. Those are, the Output Error method (OE), and the Prediction Error method (PE). The OE method minimizes the sum of squared simulation error and is used in white box modelling but also in other contexts. This method is always off line. The PE method minimizes the sum of squared one step prediction error, this method offers both off-line and on-line estimation. In order to allow for a comparison between the methods the off-line method is considered in the following. Young (1981) gives an overview and comparison of parameter estimation for continuous time models, which includes PE and OE principles. The maximum likelihood method as presented here is a PE method, whereas the OE method, can in statistical implications include Maximum Likelihood terms for the case where there is no system noise, Young (1981). Using the Kalman filter notation, the sum of squares of the error terms for the OE method is written as $\Sigma(y_k - \hat{y}_{k|0})^2$. This corresponds to a state space representation without system noise and all the errors incorporated in the measurement noise, which means prediction without updating, i.e. a simulation.

Comparing the computational time for the two modelling approaches, the state space formulation and the Kalman filter in general involve more calculations since a state filtering through the whole data series is needed for each evaluation of the objective function. It is thus questionable whether this time-consuming estimation method is worth the time. Kristensen (2002) performed a simulation study for continuous discrete models by comparing the PE method as implemented in the program CTSM and the OE method as implemented by Bohlin and Graebe (1995). The calculations for the OE method were performed by using the MoCaVa software (Bohlin (2001)), which runs under Matlab. Some of the results and discussions are also demonstrated in Kristensen et al.

(2004). The results show that the PE estimation method gives significantly less biased estimate of the parameters than the OE method. For simulations with no system noise the methods were similar, but the more noise the greater is the difference between the two methods resulting in larger bias for the OE method. Moreover, the PE estimation method provides uncertainty information in terms of standard deviations of the estimates and other statistical tools for model evaluation. Therefore, for the purpose of short-term prediction such as in flood warning systems it is truly recommended to use the PE method even though the method is more computational demanding. Once the parameters are estimated the output prediction is not time consuming. Only if the model focus on good long-term prediction capabilities, the OE method is to be preferred Kristensen et al. (2004). It must, however, be kept in mind that the input, i.e. precipitation and temperature are always needed as input and long-term prediction for precipitation and temperature variables are not particularly precise and for that reason long-term prediction might not be so reliable. Last, but none the least, in a state space formulation it is easy to handle missing values in observations automatically, and this prevents the user from having to resort other models (e.g. black box models) to fill in gaps in the data.

It is worth mentioning that Rajaram and Georgakakos (1989) presented a parameter estimation of stochastic hydrologic models formulated in a continuous-discrete state space form, with the parameters estimated in a batch form. Their method mainly differs from the one presented here in two ways. Firstly, the filtering, or the state prediction is calculated by a fourth order predictor–corrector scheme, while here a BDF method is used. Secondly, and probably the most important difference, in the methodology presented here the system error, $\sigma(u_t,t,\theta)\mathrm{d}\omega_t$ is estimated. Conversely in the methodology presented by Rajaram and Georgakakos (1989), the estimation of the state error $\sigma(u_t,t,\theta)\mathrm{d}\omega_t$ demands a human input. In Rajaram and Georgakakos (1989) the state error $\sigma(u_t,t,\theta)\mathrm{d}\omega_t$ is decomposed into three error terms; error term from input, error term associated with estimation of uncertain constants (such as topographic or rating curve constants) and error term in model structure. Only the last term is estimated, the two first must be set as a degree of believe by the trained hydrologists if they are not exactly known, Rajaram and Georgakakos (1989).

6. Results

The stochastic model in Eq. (4)–(8) is used to investigate how the parameter estimation method performs for the hydrological problem described in Section 2. The parameters are estimated by using the first 6 years of the data while the last two years are used for validation. For a comparison between the PE and OE method the optimization was performed using both methods. First in a PE setting by estimating *all* the parameters, including the system noise, and then in an OE settings by fixing the system noise term parameters to a small value. The former parameter values are optimal for prediction and the latter for simulations. The estimated parameter values are shown in Table 1.

The units for the snow container N and the upper and lower reservoirs S_1 and S_2 are given in meters. The total volume is calculated by multiplying with the watershed area. Fig. 4 illustrates the results from the PE method and Fig. 5 illustrates the results from the OE method.

Note from Fig. 5, that the OE formulation produces the same prediction and simulation and hence the coefficient of determination (Nash and Sutcliffe, 1970), is the same, $R^2=0.69$, in both cases. Conversely, the PE method produces very different results for prediction and simulation with coefficients of determination as $R^2_{\text{prediction}}=0.93$ and $R^2_{\text{simulation}}=0.43$, respectively. Furthermore, it is interesting to compare some of the estimated parameter values yielded by the two different estimation techniques. The precipitation correction factor c, the threshold parameter b_0 (for snow/rain) and the positive degree-day constant pdd are much larger for the OE estimation than for PE estimation; the difference being almost factor 2. The routing constants k_1 and k_2 are, however, smaller in the OE estimation, whereas the filtration f is similar. Note also that the PE method estimates some memory in the temperature, i.e. $a=1.475$ while the OE estimates no memory in the temperature i.e. $a=4.939$. Finally, the total noise is incorporated in the measurement noise in the OE settings, resulting in larger prediction error

Table 1
Estimation results and comparison of the PE and OE method

	PE method		OE method		Unit
Par.	Estimate	Std. dev.	Estimate	Std. dev.	
N_0	0.000		0.000		*m*
S_0	0.00099	0.00020	0.00010	3×10^{-7}	*m*
S_1	0.00037	0.00022	0.00016	5×10^{-7}	*m*
T_{s0}	3.000		3.000		°C
b_0	4.511	0.012	8.131	0.073	°C
b_1	1.000		1.000		
M	1.000		1.000		
B	100		100		
K	200		200		
c	1.518	0.00075	2.788	0.054	
pdd	0.00342	9×10^{-7}	0.00585	6×10^{-6}	m/°C day
F	0.031	0.00065	0.049	0.002	1/day
k_1	0.674	0.05704	0.216	0.010	1/day
k_2	0.097	0.00424	0.049	0.070	1/day
A	1.475	0.01411	4.939	0.023	
σ_{Ts}	10^{-8}		10^{-8}		°C
σ_N	0.0074	0.00012	1×10^{-6}		*m*
σ_{S_1}	0.0011	0.00042	1×10^{-6}		*m*
σ_{S_2}	0.0008	0.00002	1×10^{-6}		*m*
K	0.00198	0.00005	0.00175	0.00003	m/day
s_1	2.2×10^{-10}	7.6×10^{-12}	1.4×10^{-6}	5.0×10^{-8}	m/day
$R^2_{pred.}$	0.93		0.69		
$R^2_{sim.}$	0.43		0.69		

and hence, the confidence band around the prediction is larger than in the PE settings.

In the following the results from the PE estimation will be discussed. The routing constants k_1 and k_2, and the filtration f are measured in the unit 1/day, i.e. 24 h. For hourly values the constants can be multiplied by 1/24. The routing constants are rather small but bearing in mind that the size of the watershed is

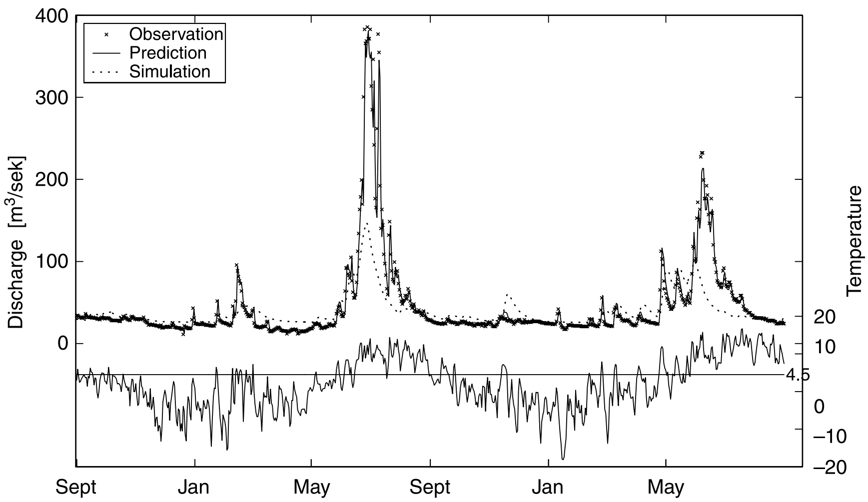

Fig. 4. The results from PE estimation. The validation period from 1st of September 1982 to 31st of August 1984. The figure shows the river discharge the one step prediction and the simulation. The temperature shown is the low pass filtered air temperature.

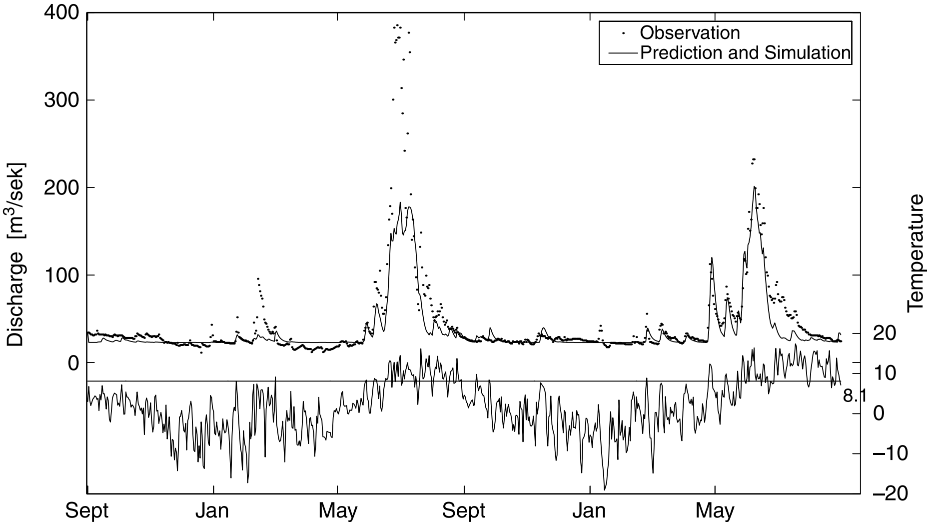

Fig. 5. The results from OE estimation. The figure shows river discharge, the one step prediction and the simulation. The validation data are from 1st of September 1982 to 31st of August 1984.

1132 km^2, these small constants are realistic. The upper routing constant is 0.674/day and thus the corresponding time constants is about 1 and a half day, whereas, the lower routing constant is 0.097/day, and hence the time constant is about 10 and a half day. The filtration constant is 0.031/day and the corresponding time constant about 32 days. Consequently most of the spring flood is delivered through the river via the first reservoir. The threshold function Eq. (1) for dividing precipitation into snow and rain is the same as the threshold function for melting snow. The parameter b_1 controls the steepness and has been set to one, and the parameter b_0 controls the center and is estimated to 4.50 °C. Thus, the threshold function is about zero when the temperature is 0 °C and then no snow is melting and all precipitation is solid. When the temperature is about 9 °C snow is melting everywhere and all precipitation is rain. In between some precipitation is snow and some as rain, and a proportion of the snow is melting (if there is snow in the snow-container). Recall that the watershed is 1132 km^2 with an altitude ranging from 44 to 1084 m and the meteorological observatory is located about 20 km from the watershed outlet at an altitude about 150 m. The center of mass of the watersheds altitude is about 830 m and, if it is assumed that the temperature in altitude 830 m is zero when the temperature is 4.5 °C at the observatory, it leads to a temperature lapse rate of 4.5/6.8 = 0.66 °C/100 m, which is physically realistic. Hence, this smooth threshold function has the effect that it is not necessary to divide the area into elevation zones.

The precipitation correction constant c is estimated as 1.5. This correction is both correcting the underestimate of the rain gauge and the average increase in precipitation due to altitude. The factor c controls the input-output balance of the model. A water balance model with ground water container and evapotranspiration would have had a much larger correction constant. However, it should be mentioned that it is not possible to identify (estimate) both the correction constant and evapotranspiration given only measurement of the precipitation and discharge. Finally, Fig. 6 shows the state estimates of the contents of the snow-container, N, and the upper and lower surface containers S_1 and S_2 as predicted by the model using the PE parameters.

The unit is meter and the total volume is calculated by multiplying with the watershed area, 1132 km^2. The estimated noise terms σ_N, σ_{S_1} and σ_{S_2} have a order of magnitude 10^{-3} and thus the noise terms more or less only have an effect when the states are around zero, with the consequence that the states might become slightly negative. This has not lead to problems in this case. The problem might be solved by transforming the model using the logarithm.

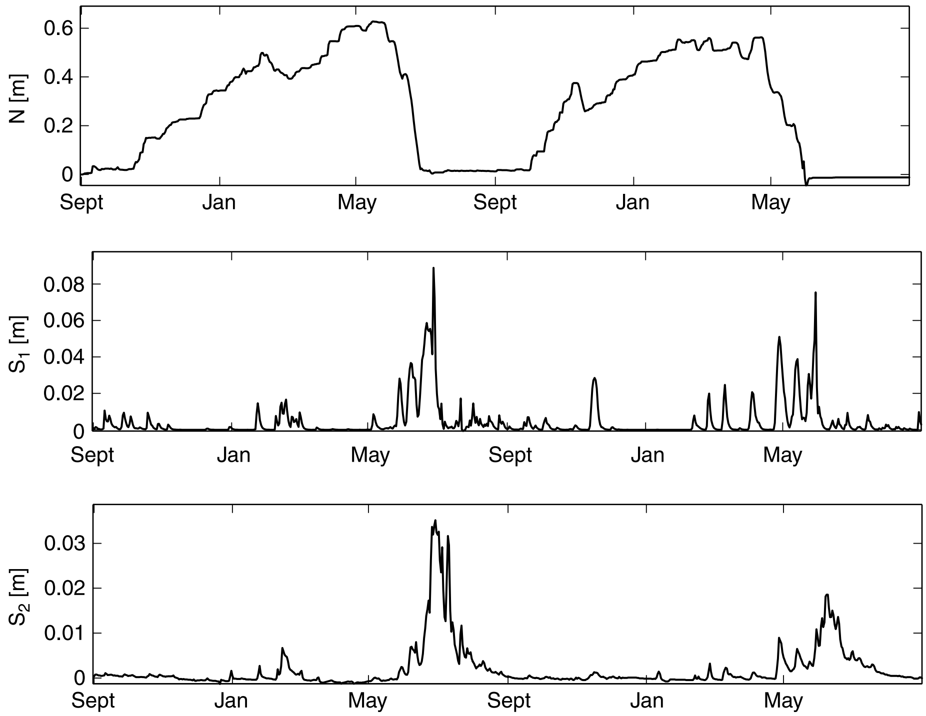

Fig. 6. State estimates of the contents of the snow container and the upper and lower surface containers.

As a flood forecasting model it is concluded that this simple model is satisfactory. Using the simple model makes it possible to estimate all the parameters, thus allowing the data to be used for an automatic calibration. The optimization, using 6 years of data, takes several hours on a PC computer but as mentioned earlier, once the parameters are estimated the update of the Kalman filter and the prediction are not computational demanding. CTSM can be run on a parallel computer using several CPUs and then the computer time will be much lower.

For the purpose of flood forecasting the most interesting development of the model would be to include some of the parameters in the state vector and thus allow for time varying parameters. Particularly since the parameters have been estimated, these estimates could act as good initial states for the time varying parameters. This could particularly be done for the base flow constant.

Finally, it is interesting to point out some revisions, which might improve the performance of simulation using an OE estimation. The large threshold temperature $b_0 = 8.1$ °C indicates that it might be necessary to divide the area into two elevation zones, still using smooth threshold functions but with different centers. The former spring flood is much higher and narrower than the latter and such a narrow flood is difficult to produce. It might be necessary to have three surface containers and thus three time constants for the flow. It would also be interesting to let the pdd constant vary in time. Rango and Martinec (1995) state that the positive degree day factor should gradually increase during the melting season and this could certainly be introduced in the PE settings as well. A time varying pdd might though have larger differences in cases where the melting season is longer such as for glacier rivers.

7. Conclusions

All precipitation runoff models are approximations of the reality and hence they cannot be expected to provide a perfect fit to data. The process is highly non-stationary and the dynamics related to the snow is extremely non-linear. Furthermore, the deviations between the model prediction and the data (the residuals) are almost always serially correlated. This calls for a stochastic model with both system noise and measurement noise.

In this paper a simple conceptual stochastic rainfall-runoff model is suggested. A method for estimation of the parameters of the model is outlined. The estimation method is a generic maximum likelihood method for parameter estimation in systems described by continuous-discrete time state space models, where the system equation consists of stochastic differential equations. Hence, the dynamics are described in continuous time, which allows for a direct use of prior physical knowledge, and the estimated parameters can be physically interpreted directly.

A further advantage of the stochastic state space approach is that the same model structure can be used for both prediction and simulation. It is advocated that the only difference lies in a different parameterization of the system error leading to different parameter values.

The presented model is simple and demands only two input variables, namely precipitation and temperature, and a single output, the discharge. The results for simulation are reasonable but not fully satisfying and it is concluded that a slightly more complicated model is needed even though it is questionable whether it is possible to obtain a better performance due to the poor precipitation data as in this study. However, the results obtained for prediction (flood forecasting) are satisfying.

Acknowledgements

The authors wish to thank the National Energy Authority in Iceland for delivery of discharge series, and The Icelandic Meteorological Office for meteorological data series. Furthermore, we wish the thank PhD Niels Rode Kristensen for technical support with the program CTSM.

References

Abbott, M., Bathurst, J., Cunge, J., O'Connell, P., Rasmussen, J., 1986. An introduction the the European Hydrological System—System Hydrologique Europeen, 'SHE' 2: structure of a physically-based, distributed modelling system. Journal of Hydrology 87, 61–77.

Ashan, M., O'Connor, K.M., 1994. A reappraisal of the kalman filtering technique, as applied in river flow forecasting. Journal of Hydrology 161, 197–226.

Bathurst, J.C., 1986. Physically-based distributed modelling of an uplang catchment using the Systeme Hydrologique Europeen. Journal of Hydrology 87, 79–102.

Bergström, S., 1975. The development of a snow routine for the HBV-2 model. Nordic Hydrology 6, 73–92.

Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources Publication, Colorado.

Bergström, S., Fossman, A., 1973. Development of a conceptual deterministic rainfall-runoff model. Nordic Hydrology 4, 147–170.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sciences Bulletin 24, 43–69.

Bohlin, T., 2001. A Grey-box Process Identification Tool: Theory and Practice. Technical Report IR-S3-REG-0103, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden.

Bohlin, T., Graebe, S.F., 1995. Issues in nonlinear stochastic grey-box identification. International Journal of Adaptive Control and Signal Processing 9, 465–490.

Burnash, R., 1995. NWS river forecast system—catchment modeling. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources Publication, Colorado.

Burnash, R.J.C., Ferral, R., McGuire, R., 1973. A Generalized Streamflow Simulation System—Conceptual modeling for Digital Computers. Technical Report, National Weather Service, NOAA and State of California Department of Water Resources, Joint Federal State River Forecast Center Sakramento, California.

Dahlquist, G., Björck, Å, 1988. Numerical Methods. McGraw-Hill, New Jersey.

Førland, E.J., Allerup, P., Dahlström, B., Elomaa, E., Jónsson, T., Madsen, H., Perälä, J., Rissanene, P., Vedin, H., Vejen, F., 1996. Manual for Operational Correction of Nordic Precipitation Data. Technical Report No. 24/96 ISNN 0805-9918, The Norwegian.

Georgakakos, K.P., 1986a. A generalized stochastic hydrometeorological model for flood and flash-flood forecasting I. Water Resources Research 22, 2083–2095.

Georgakakos, K.P., 1986b. A generalized stochastic hydrometeorological model for flood and flash-flood forecasting II. Water Resources Research 22, 2096–2106.

Georgakakos, K.P., Rajaram, H., Li, S.G., 1988. On improved Operational Hydrologic Forecasting. Technical Report IIHR No. 325, Institue of Hydraulic Research, Univeristy of Iowa.

Gottlieb, L., 1980. Development and applications of runoff model for snowcovered and glacierized basins. Nordic Hydrology 11, 255–272.

Haltiner, J., Salas, J.D., 1988. Short term forecasting of snowmelt runoff using armax models. Water Resources Bulletin 24, 1083–1089.

Havnø, K., Madsen, M.N., Dørge, J., 1995. MIKE11 a generalized river modelling package. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources Publication, Colorado.

Jain, S., Storm, B., Bathurst, J., Refsgaard, J., Singh, R., 1992. Application of the SHE to catchments in India Part 2. Field experiments and simulation studies with the SHE on the Kolar subcatchment of the Narmand river. Journal of Hydrology 140, 25–47.

Jazwinski, A.H., 1970. Stochastic Processes and Filtering Theory. Academic Press, San Diego.

Kristensen, N.R., 2002. Fed-Batch Process Modelling for State Estimation and Optima Control. PhD Thesis, Department of Chemical Engineering, The Technical University of Denmark. ISBN 87-90142-83-7.

Kristensen, N.R., Melgaard, H., Madsen, H., 2003. CTSM Version 2.3—A Program for Parameter Estimation in Stochastic Differential Equations. Available from www.imm.dtu.dk/ctsm.

Kristensen, N.R., Madsen, H., Jørgensen, S.B., 2004. Parameter estimation in stochastic grey-box models. Automatica 40, 225–237.

Layrenson, E., Mein, R., 1995. RORB a hydrograph synthesis by runoff routing. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources Publication, Colorado.

Lee, Y.H., Singh, V.P., 1999. Tank model using Kalman filter. Journal of Hydrologic Engineering 4, 344–349.

Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. Journal of Hydrology 235, 276–288.

Martinec, J., 1960. The degree day factor for snowmelt-runoff forecasting. In: General Assembly of Helsinki Commission on Surface Runoff, number 51 in IAHS, pp. 468–477.

Martinec, J., Rango, A., 1986. Parameter values for Snowmelt runoff modeling. Journal of Hydrology 84, 197–219.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models par i—a discussion of principles. Journal of Hydrology 10, 282–290.

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

Rajaram, H., Georgakakos, K.P., 1989. Recursive parameter estimation of hydrologic models. Water Resources Research 25, 281–294.

Rango, A., Martinec, J., 1995. Revisiting the degree-day method for snowmelt computations. Water Resources Bulletin 31, 657–669.

Refsgaard, J., Storm, B., 1995. MIKE SHE. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources Publication, Colorado.

Refsgaard, J.C., Rosbjerg, D., Markussen, L.M., 1983. Application of the Kalman filter to real time operation and to uncertainty anlyses in hydrological modelling. In: Plate, E., Buras, N. (Eds.), Scientific Procedures Applied to the Planning, Design and Management of Water Resources Systems Number 147 in IAHS, pp. 273–282.

Refsgaard, J., Seth, S., Bathurst, J., Erlich, M., Storm, B., Jørgensen, G., Chandra, S., 1992. Application of the SHE to catchments in India Part 1. General results. Journal of Hydrology 140, 1–23.

Restrepo, P.J., Bras, R.L., 1985. A view of maximum-likelihood estimation with large conceptual hydrologic models. Applied Mathematics and Computation 17, 375–404.

Sajikumar, N., Thandaveswara, B., 1999. A non-linear rainfall-runoff model using an artificial neural network. Journal of Hydrology 216, 35–55.

Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modelling. Journal of Hydrology 199, 272–294.

Singh, V.P., Woolhiser, D.A., 2002. Mathematical modeling of watershed hydrology. Journal of Hydrologic Engineering 7, 270–292.

Sugawara, M., 1995. Tank model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water resources Publication, Colorado.

Szollosi-Nagy, S., 1976. An adaptive identification and prediction algorithm for the real time forecasting of hydrological time series. Hydrological Sciences Bulletin 21, 163–176.

Todini, E., 1978. Using a desk-top computer for an on-line flood warning system. IBM Journal of Research and Development 22, 464–471.

Todini, E., 1996. The ARNO rainfall-runoff model. Journal of Hydrology 175, 339–382.

Young, P., 1981. Parameter estimation for continuous-time models—a survey. Automatica 17, 23–39.

Zhao, R.-J., 2002. The Zinanjiang model applied in China. Journal of Hydrology 135, 371–381.

Zhao, R.J., Liu, X., 1995. The Zinanjiang model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publication, Colorado.