Hydrological analysis for a distributed and a semi-distributed model

A. A. El-Nasr¹, J. Berlamont & J. Feyen²
¹Laboratory of Hydraulics, K.U.Leuven, Kasteelpark Arenberg 40, 3001 Heverlee, Belgium.

Abstract

This paper presents 2 different methods, increasing gradually in complexity, for predicting the rainfall-runoff of a midsize catchment. The methods applied in this study are (1) a lumped conceptual semi-empirical method, being the NAM-module of the MIKE 11 model, and (2) a fully distributed, physically-based deterministic catchment model, the MIKE SHE model. The two methods are applied to the 465 km² large Jeker basin, situated in the loamy belt region of Belgium. The landscape is rolling, and the soils are varying from sandy-loam to clay-loam. A 6-year continuous series of average daily hydrologic data were used for the calibration and validation. Mainly for reasons of workability of the distributed modelling approach the DTM is 600x600 m. The distributed model was calibrated and validated using a split-sample (SS) and a multi-site (MS) test, while only the split-sample test was used for the lumped model.

In this paper the performance of two hydrologic models, a lumped conceptual and a distributed physical-based model are compared using hydrologic data of a medium sized catchment in the loamy region of Belgium. The performance of both models is compared after calibration and validation.

1 Introduction

The comparison of hydrological models has been reported in several papers. Michaud and Sorooshian (1994) compared a lumped conceptual model (SCS), a distributed conceptual model (SCS with 8 sub-catchments, one per rain gauge)
and a distributed physical-based model (KINEROS) for the simulation of storm events with application on a catchment, 150 km² in size. The models were validated based on the comparison of the peak flow, the time to the peak and the runoff volume. The study revealed that with calibration the accuracy of the two distributed models was similar. Without calibration the distributed physically-based model performed better than the distributed conceptual model. In both cases the lumped conceptual model performed most poor. Refsgaard and Knudsen (1996) validated and compared three different models on three catchments, a lumped conceptual modelling system (NAM), a distributed physically-based system (MIKE SHE) and an intermediate approach (WATBAL). A theoretical framework for model validation, based on the methodology proposed by Klemes (1985, 1986) was presented, where four basic categories of typical modelling tests, including the split-sample test (SS), the differential split-sample test (DSS), the proxy-basin test (PB) and the proxy-basin differential split-sample test (PB-DSS) were applied. The study was applied on two large catchments and a medium one (1090, 1040 and 254 km²). The authors concluded that all models performed equally well when at least 1 year of data was available for calibration, while the distributed models performed marginally better for cases where no calibration was executed.

2 Materials

For the study use was made of a continuous time series of 6-year (06/01/1986 – 05/01/1992) average daily hydrological data. The Jeker catchment (465 km²) is located in the mid-east of Belgium, partly located in the Walloon and Flemish region of the country (Fig. 1). The linguistic border, affecting enormously the effort required in the collection of the basic catchment data, divides it. The dominant soil type is loam and the land use is rain fed arable land, occupying 70% of the total catchment area, and fragmented urban area, covering 17% of the basin area. The landscape is rolling and the original DTM resolution is 30x30 m. The topography of the basin varies from 59 m above sea level in the north to 200 in the south. In the south of the catchment at a depth of 50 m two man-made galleries having a diameter of 1 m drain the aquifer system. The length of the two galleries is 14 and 26 km, respectively. Time series of the daily main hydrological data, i.e., rainfall, discharge, potential evapotranspiration, groundwater and piezometric levels were composed for 6 years. The data sets were split into two periods of three years (06/01/1986 – 05/01/1989 and 06/01/1989 – 05/01/1992), for model calibration and validation, respectively. The discharge data of the outlet station Kanne was used to calibrate and validate the simulated river discharges for both models. The rainfall recorded in seven rainfall stations, distributed in and around the catchment (see Fig. 1), was used to account for the spatial distribution of the rainfall. More specifically, the Thiessen polygon method was used to estimate the spatially averaged rainfall distribution over the catchment.
3 Methods

3.1 The NAM module (MIKE 11 model)

NAM forms part of the rainfall-runoff module of the MIKE 11 river modelling system (Havno et al., 1995). It can be characterized as a deterministic, lumped, conceptual model with moderate input data requirements. A mathematical hydrological model like NAM is a set of linked mathematical equations describing, in a simplified quantitative form, the behaviour of the land phase of the hydrological cycle. NAM represents various components of the rainfall-runoff process by continuously accounting for the water content in four different and mutually interrelated storages. Each storage represents different physical elements of the catchment. NAM can be used either for continuous hydrological modelling or for simulating single events (DHI, 1999b). The storages considered in the model are surface water storage, snow storage, root zone storage and groundwater storage.

3.2 The MIKE SHE model

MIKE SHE (Refsgaard and Storm, 1995) is a deterministic physical-based distributed model for the simulation of the hydrological processes of the land phase of the hydrological cycle, with obvious large amount of input data. The SHE model was developed to model the spatial distribution of basin parameters, hydro-meteorological inputs and hydrological response in 3 dimensional forms.
This means that it represents the basin horizontally by an orthogonal grid network and it uses a vertical column at each horizontal grid square. MIKE SHE encompasses a number of components describing the flow within different parts of the hydrological cycle. They can be combined depending on the scope of the study (DHI, 1999a). The hydrological processes are modelled by finite difference representations of the partial differential equations for the conservation of mass, momentum and energy, in addition to some empirical equations. The major flow components (processes) considered in the model are flow in the saturated zone, flow in the unsaturated zone, evapotranspiration and overland channel flow.

3.3 Criteria for assessment of the model performance

For the calibration and validation of models, and the comparison of models, one needs quantitative criteria to measure how good the model is simulating the real world. In this study the river discharge data measured at the outlet of the basin were used for assessing the model performance. The approaches used are:
- Closure of the water balance of the catchment, i.e., the total evaporation over a certain period should correspond to the accumulated net precipitation minus runoff;
- Agreement of the overall shape of the time series of discharge together with the total accumulated volumes, i.e., between observed and simulated discharge, and this for both the calibration and validation periods;
- Value of statistical performance indices (Nash and Sutcliffe, 1970; Loague and Green, 1991; Gupta et al., 1998; Xevi et al., 1997; Legates and McCabe, 1999, Feyen et al., 2000) such as: the relative root mean square error (RRMSE), the mean absolute error (ABSERR), the coefficient of determination (CD), the modelling efficiency (EF) and the goodness of fit (R²);
- Graphical representation of observed and simulated (i) hydrograph maxima, (ii) baseflow minima, and (iii) extreme value distribution of rainfall-runoff discharges. By plotting the modelled discharge maxima for all hydrograph events during both the calibration and validation period against the observed maxima, the performance of the model to describe the hydrograph peak discharges is tested. Compared with the scatter plot of all daily discharges, this plot has the advantage that points are independent and that the statistical performance indices are calculated in a more explicit way. To give an equal weight to all the maxima or minima a Box-Cox (BC) transformation was applied (Bandy and Willems, 2000). In this transformation, the scattering of the higher and lower discharges is forced to be equal. The foregoing is achieved applying the following formula:

\[ BC(q) = \frac{q^\lambda - 1}{\lambda} \]  

(1)

where \( q \) is the rainfall-runoff discharge and \( \lambda \) is a calibration parameter.

In the same way as for the hydrograph maxima, the performance to describe the baseflow can be tested by plotting modelled against observed minimum discharges at the end of baseflow recession periods. On the basis of a third plot,
the extreme value distribution plot finally can be used to test the capability of the model to describe extreme events as a function of the return period.

### 3.4 Model calibration and validation

#### 3.4.1 The NAM module (MIKE 11 model)

The calibration was carried out in two major steps: an automatic and manual calibration (DHI, 1999b). First the automatic calibration was executed to speed up the calibration process and to limit and constrain the most important parameters to a range of acceptable values. The best result of the automatic calibration was further improved in a manual way, by interactively fine-tuning the most sensitive model parameters. The automatic calibration (Madsen, 2000) routine is based on a multi-objective optimisation strategy in which four different objectives are optimised simultaneously. For those objectives the following numerical performance measures were used: (i) agreement between the average simulated and observed catchment runoff (overall volume error), (ii) overall agreement of the shape of the hydrograph (overall root mean square error), agreement of peak flows (average RMSE of peak flow events) and agreement of low flows (average RMSE of low flow events). An automatic calibration has however the disadvantage that it does not use knowledge and experience-based information from the modeller. Therefore, a manual improvement is pursued with a trial-and-error parameter adjustment of the most sensitive parameters. Checking some properties of the discharge time series, which were not controlled during the automatic calibration, does this. In a first step the time constants of the subflows (overland flow, interflow and base flow) are adjusted based on the recession velocities in different parts of the decreasing flanks of the hydrographs. The second step is to adjust the peak discharges by controlling the overland flow runoff coefficient (allowing small time shifts to occur due to rainstorm movements). Storage parameters are finally improved by evaluation of the water balance in a continuous way (based on a plot of continuous runoff volumes versus time).

#### 3.4.2 The MIKE SHE model

A grid resolution of 600x600 m was chosen. The choice was based on the assumption that a 600x600 m grid cell is accurate enough for modelling the runoff process and ground water levels in an adjacent catchment to the Jeker in Belgium (Feyen et al., 1999). The following three parameters were optimised in the calibration: (i) the drainage level, used to calculate the slope of the drains to define the drainage reference system, which links each single computational grid in the saturated zone to a drainflow recipient; (ii) the time constant, which determines the drainage rate i.e. the height of the peaks and the tailing of the recession on the river discharges; and (iii) the vertical and horizontal hydraulic conductivity of the geological layers in the saturated zone. The hydraulic conductivity was calibrated in a distributed way, depending on the sequence and spatial distribution of the geological layers in the study area. The most significant
effect was recorded by changing the hydraulic conductivity of the fractured and compacted zones of the Cretaceous Chalk layer. The compacted zone is characterised by the presence of many abstraction wells and two large extraction galleries draining water by gravity. Introducing around the area of the galleries zones with different hydraulic conductivity mimicked the effect of the draining galleries. Further a constant abstraction along the galleries was assumed. The pre-defined six geological layers were reduced to three computational layers in the model structure merging similar layers. The foregoing was done with the objective to reduce the problems arising from the sparse distribution and the small thickness of some of geological layers that slows down the computational process. To make full use of the available time series of hydrologic data for the calibration and validation the warming up period of the model was reduced (one year). Based on the available time series of ground water observation wells within the catchment, spatially distributed initial conditions were derived using the average ground water levels at each point of the catchment. The foregoing was possible by properly defining the conditions at the water divide and the initial spatial conditions. Based on a-priori knowledge that a constant flux of water flows into the catchment along the southern boundary, a constant boundary condition was introduced in the south whereas elsewhere impermeable boundaries were assumed.

4 Results and discussion

The comparison between observed and simulated discharges, using both model approaches, are graphically presented in the Figs. 2 and 3 where each figure shows the result of one-year simulation. The total accumulated volumes for the entire period of the three years for both cases are also indicated. To assess the performance of both models with respect to their capacity to simulate the maximum and minimum values of the discharge, a set of maximum and minimum values were chosen. To this end in the calibration and validation period the maximum and minimum discharge of each hydrograph event was selected.

![Figure 2: Observed and simulated average daily discharge in the Kanne limnigraphic station for the last year of the calibration period and observed and simulated accumulated discharges for the entire calibration, using both models](image-url)
Figure 3: Observed and simulated average daily discharge in the Kanne limnigraphic station for the last year of the validation period and observed and simulated accumulated discharges for the entire validation, using both models. The separation of the events is based on an automatic procedure (Bandy and Willems, 2000). In doing so the maximum and minimum discharges can be considered as independent. The selected discharges are plotted in Fig. 4a & b as scatter plots after being transformed using the Box-Cox transformation procedure (see Eq. 1).

Figure 4: Scatter plots of Box-Cox transformed independent observed and simulated maximum (a) and minimum (b) values of the discharge for the limnigraphic station in Kanne, selected in the calibration and validation period.

Table 1 shows the results of the statistics for the simulated versus observed time series of the discharge at the main outlet station. Better results are characterised by values of RRMSE and MAE closer to zero and R2, CD and EF values closer to unity. From Fig. 4 and table 1, it can be concluded that the NAM model predicts slightly better the maximum and minimum discharges than the MIKE SHE model. Particularly in the validation period the MIKE SHE model underestimate the maximum values, and it underestimates the lower baseflow.
values. The NAM model is slightly overestimating the minimum discharge values in the calibration period.

Table 1: Value of 5 statistical indices, applied to the observed versus the calculated average daily discharges for the calibration and validation period, using the NAM and MIKE SHE model

<table>
<thead>
<tr>
<th>Statistical index</th>
<th>NAM module (MIKE 11 code)</th>
<th>MIKE SHE code</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration</td>
<td>Validation</td>
</tr>
<tr>
<td>RRMSE</td>
<td>0.1979</td>
<td>0.2030</td>
</tr>
<tr>
<td>ABSErr</td>
<td>0.3397</td>
<td>0.2960</td>
</tr>
<tr>
<td>CD</td>
<td>1.3058</td>
<td>1.1216</td>
</tr>
<tr>
<td>EF</td>
<td>0.7422</td>
<td>0.7671</td>
</tr>
<tr>
<td>R²</td>
<td>0.7512</td>
<td>0.7811</td>
</tr>
</tbody>
</table>

Figure 5 was designed to examine how the two models behave for extrapolation of extreme discharges. For the construction of this figure the ‘peak-over-threshold’ (POT) extreme value analysis technique was applied (Willems, 1998 & 2000). The highest observed and simulated independent peak discharges are derived from the entire calibration and validation period. An exponential extreme value distribution was fitted to these discharges. It can be concluded from Fig. 5 that the distributed model predicts better extreme discharges than the lumped model, and the latter model simulates better the less extreme discharges.

Figure 5: Comparison of the observed and simulated extreme value peak discharges for the Kanne limnigraphic station as a function of the return period

The overall results from the goodness-of-fit statistics (Tables 1) and plots (Figures 2, 3 & 4) show that there is no significant difference between the overall behaviour of the two models during the validation period. According to the results presented in the table and the figures in the calibration period the NAM module seems to perform slightly better than the MIKE SHE model. The MIKE SHE model seems to underestimate moderate peak discharges, whereas the NAM
module underestimates extreme high discharge values. The latter conclusions were also derived from the extreme value analysis (Fig. 5).

5 Conclusion

The performance of both models was tested using a qualitative (graphical) and quantitative (statistical) assessment. In comparison to the MIKE SHE model, the NAM module has less model parameters and requires less input. The graphical and statistical tools used indicate that neither of the modelling approaches, used for conversion of rainfall into runoff, statistically performs much better than the other one. The calibration results of both models, although very different in concept and degree of spatial distribution, perform quite close. The foregoing not only might be the consequence that prior to the model comparison most sensitive model parameters was calibrated, but also because only daily values of climate and river discharge variables were available. The analysis also learned that predictions made by the NAM module results in a less variable, more averaged, time series of river discharge as compared to the observed series. This leads to an underestimation of the flood frequency distribution. Due to the higher physically-based level and the more detailed description of the surface topography and roughness of the MIKE SHE modelling code, as compared to the NAM module, the first model better predicts extreme discharge values.

References

[8] Gupta, H.V., Sorooshian, S. & Yapo, P.O., Toward improved calibration of


