Hydrol. Earth Syst. Sci. Discuss., 6, 729–766, 2009
www.hydrol-earth-syst-sci-discuss.net/6/729/2009/

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Combining semi-distributed process-based and data-driven models in flow simulation: a case study of the Meuse river basin

G. Corzo[1], **D. Solomatine**[1,3], **Hidayat**[1,5], **M. de Wit**[2], **M. Werner**[1,2], **S. Uhlenbrook**[1,4], **and R. Price**[1,3]

[1]UNESCO-IHE Institute for Water Education, Delft, The Netherlands
[2]Deltares, Delft, The Netherlands
[3]Water Resources Section, Delft University of Technology, Delft, The Netherlands
[4]Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[5]Centre for Limnology, Indonesian Institute of Sciences, Cibinong, Indonesia

Received: 15 December 2008 – Accepted: 15 December 2008 – Published: 10 February 2009

Correspondence to: G. Corzo (corzogac@yahoo.es)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

One of the challenges in river flow simulation modelling is increasing the accuracy of forecasts. This paper explores the complementary use of data-driven models, e.g. artificial neural networks (ANN) to improve the flow simulation accuracy of a semi-distributed process based model. The IHMS-HBV model of the Meuse river basin is used in this research. Two schemes are tested. The first one explores the replacement of sub-basin models by data-driven models. The second scheme is based on the replacement of the Muskingum-Cunge routing model, which integrates the multiple sub-basin models, by an ANN. The results showed that: (1) after a step-wise spatial replacement of sub-basin conceptual models by ANNs it is possible to increase the accuracy of the overall basin model; (2) there are time periods when low and high flow conditions are better represented by ANNs; and (3) the improvement in terms of RMSE obtained by using of ANNs is greater than that when using sub-basin replacements. It can be concluded that the presented two schemes based on the analysis of seasonal and spatial weakness of the process based models can improve performance of the process based models in the context of operational flow forecasting.

1 Introduction

It is a common practice to use semi-distributed conceptual models in operational forecasting for meso-scale catchments. These models are based on the principle of mass conservation and simplified forms of energy conservation. Conceptual models, however, may not represent all sub-basins with the same accuracy. Inaccurate precipitation data and the need for its averaging for the lumped models may seriously influence the accuracy of modelling. Due to the limited representation of the full rainfall-runoff process, the complexity of the model integration and the identification of the lumped parameters, the proponents of fully distributed detailed models argue that there many situations when the accuracy of conceptual models is not sufficient. However, the simplicity of these models and the high processing speed is an advantage for real time operational systems and often makes such models the first choice.

Precipitation forecasts are normally available for low resolution grids which are close to the size of the modelled sub-basins. It has been shown that there are situations when such models are more accurate than the fully spatially distributed physically based and energy based models (Seibert, 1997; Linde et al., 2007). Diermansen (2001) presented an analysis of spatial heterogeneity in the runoff response of large and small river basins, and an increase of error is observed with the increase of the level of detail in the physically based model. An alternative to fully-distributed models is the class of intermediate models, the so-called semi-distributed conceptual models, as the most appropriate modelling approach for meso-scale operational forecasting. In this research the IHMS-HBV model (Lindström et al., 1997) belonging to this class is used (http://www.smhi.se). In this paper it will be called simply HBV, and will refer to the initial hydrological model formulation used as a hydrological prototype module in the flood early warning system for the rivers Rhine and Meuse.

Ashagrie et al. (2006) presented a long term analysis for the effects of climate change and land use change on the Meuse river basin using the HBV model. This analysis showed that the agreement between the observed and measured discharge is generally good, in particular flood volumes and the highest peak are simulated well. However, there are some problems with the medium flow (shape and peak values), and a systematic deviation between for certain observed periods (i.e. 1930–1960) was also observed. de Wit et al. (2007b) explored the impact of climate change on low-flows. They found high accuracy for the monthly average discharge and for the highest (January) and lowest discharge (August), but there was an overestimation and underestimation observed in spring and autumn, respectively. Many performance calibration techniques with different types of models have been used for the Meuse. Booij (2005) presented the manual calibration and validation of the HBV based on expert tuning of model parameters. The problems mentioned above still remain unresolved and under investigation by a number of authors.

An alternative approach to flow forecasting is using data-driven models (DDM). The most common DDMs used in hydrological forecasting are artificial neural networks (ANN) (ASCE, 2000), but other types of models, for example M5 model trees Solomatine and Dulal (2003) or support vector machines (Dibike et al., 2001) are used as well. Traditionally, modellers build a general model that covers all the processes of the natural phenomenon studied (overall model). Hydrological forecasting data-driven models are not exceptions in this sense: they tend to be developed on the basis of using a comprehensive overall model that covers all the processes in a basin (ASCE, 2000; Dibike and Abbott, 1999; Abrahart and See, 2002; Dawson et al., 2005). However, such models do not encapsulate much of the knowledge that experts may have about the system, and in some cases suffer from low extrapolation capacity (generalization capability). In many applications of data-driven models, the hydrological knowledge is "supplied" to the model via a proper analysis of the input/output structure and the choice of the adequate input variables (Solomatine and Dulal, 2003; Bowden et al., 2005). These models are less sensitive to precipitation and temperature information in hydrological systems where high autocorrelation is found in streamflows. Therefore, in operational systems where missing data is an issue, such models can replace local sub-basin models. Additionally, a complex distributed water system requires local model evaluations and integration of models. So an alternative is to build an overall DDMs for the whole basin, and for semi-distributed hydrological modelling a combination of hydrological process-based and data-driven models can be used. Additionally, the routing model integrating sub-basin models can be replaced by a DDM as well. These two approaches are explored in this paper.

Finding adequate combinations of the mentioned model types (conceptual models and DDMs) is a relatively new area of research and has been studied only in the recent years. Anctil and Tape (2004) presented a successful combination of conceptual models where the information from the time series of soil moisture is fed into a network. Their study concentrated on using the daily time series for flow forecasting purposes. However in the same study, the problems of using potential evapotranspiration and antecedent precipitation index as input to the ANN models are reported. The work presented by Nilsson et al. (2006) shows that not only information about soil moisture but also about snow accumulation may bring improvement into the ANN modelling process. Their results were based on monthly data with the purpose of having more accurate forecasts for power production, dam safety and water supply. Although in both papers integration of models is employed, none of them use all the information from the conceptual model. Also the catchment complexities have been limited to basin size which in both studies has a maximum of 1400 km^2.

Chen and Adams (2006) presented the description of sub-basin models and their routing through the use of an ANN-based integration model. The basin area was around 8500 km^2, with a division into three sub-basins based mainly on the river network system. Calibration process included two stages: first, the whole catchment was considered (no sub-basin discharge information was available), and, second, with the use of output discharges from the basins to the outlet. This approach is similar to the one tested in the present paper, but we considered more complex basin, compared the model with the ANN routing integrator with a full basin hybrid model involving ANN submodels, and performed additional analysis of the variations of the models seasonal performance.

The approaches presented in this paper follow the general framework of integrating hydrological concepts and data-driven models using modular models that is being developed in our recent publications (Corzo and Solomatine, 2007, 2005; Fenicia et al., 2007; Solomatine and Price, 2004).

The objectives of this paper are: (i) to analyse the performance of DDMs in their role as sub-basin replacements, in terms of local and overall flow simulation errors; (ii) to explore different data-driven methods as alternative methods for the integration or replacement of sub-basins; iii) draw conclusions about the applicability of the hybrid process-based and data-driven models in operational flow forecasting.

2 HBV-M model for Meuse river basin

The conceptual hydrological model HBV was developed in the early 1970s (Bergström and Forsman, 1973) and its versions have been applied to many catchments around the world (Lindström et al., 1997). HBV describes the most important runoff generating processes with simple and robust procedures. In the snow routine, snow accumulation and melt are determined using a degree temperature-index method. The soil routine divides the forcing by rainfall and meltwater, into runoff generation and soil storage for later evaporation. The runoff generation routine consists of one upper non-linear reservoir representing fast and intermediate runoff components, and one lower linear reservoir representing base flow. Runoff routing processes are simulated using a simplified Muskingum approach and/or a triangular equilateral transfer function.

HBV is a semi-distributed model and the river basin can be subdivided into sub-basins (HBV-S). This model simulates the rainfall-runoff processes for each sub-basin separately with a daily or hourly time step. Each sub-basin is divided into homogenous elevations which are then divided into vegetation zones. Further details about the HBV model can be found in Lindström et al. (1997) and Fogelberg et al. (2004).

2.1 Characterisation of the Meuse river basin

The Meuse river originates in France, flows through Belgium and The Netherlands, and finally drains into the North Sea (Fig. 1). The river basin has an area of about 33 000 km^2 and covers parts of France, Luxemburg, Belgium, Germany and The Netherlands. The length of the river from its source in France to the North Sea at the Hollands Diep (an estuary of the Rhine and Meuse rivers) is about 900 km. Major tributaries of the Meuse are the Chiers, Semois, Lesse, Sambre, Ourthe, Amblve, Vesdre and Roer. The hydrological model of the Meuse basin upstream of Borgharen is subdivided into 15 sub-basins, covering an area of 21 000 km^2 (Fig. 1). For more detailed information about catchment geological and hydrological properties see Berger (1992) and de Wit et al. (2007a).

In general terms the land use in the basin is made up of 34% arable land, 20% pasture, 35% forest and 9% built up areas (source: CORINE). Tu et al. (2005) found the coverage of forest and agricultural land relatively stable over the last ten years, but the forest type and management practices have changed significantly. In addition to this it seems that intensification and upscaling of agricultural practices and urbanization are the most important land changes in the last century.

As far as the hydrologic properties are concerned the Meuse can roughly be split into three parts Berger (1992):

1. The upper reaches (Meuse Lorraine), from the Mesuse source to the mouth of the Chiers. Here the catchment is lengthy and narrow, the gradient is small and the major bed is wide. Because of that the discharge up to the mouth of the Chiers has a comparatively calm course.

2. The central reaches of the Meuse (Meuse Ardennaise), leading from the Chiers to the Dutch border near Eijsden. In that section the main tributaries are Viroin, Semois, Lesse, Sambre and Ourthe. Here the Meuse transects rocky stone, resulting in a narrow river and a great slope. The poor permeability of the catchment and the steep slope of the Meuse and most of the tributaries contribute to a fast discharge of the precipitation. The contribution of the area to flood waves is great, the contribution to low flows is small.

3. The lower reaches of the Meuse, corresponding to the Dutch section of the river. The lower reaches themselves may again be split into the stretches from Eijsden to Maasbracht and from Maasbracht to the mouth. In the former part the slope is still relatively high. For the greater part the river has no weirs here. In the section the Meuse has no dikes. For those reasons the stretch above Maasbracht is occasionally reckoned as part of the Meuse Ardennaise, which in that case flows from Sedan to Maasbracht. It may be remarked that the stretch that forms the border with Belgium is called the Grensmaas (Border Meuse) in the Netherlands, and Gemeenschappelijke Maas (Common Meuse) in Flanders.

2.2 Data validation

The validation of the data sets presented in this paper is based on the results obtained from different researches (Booij, 2002; Deursen, 2004; Ashagrie et al., 2006; Leander and Buishand, 2007; de Wit et al., 2007b). The overall model error obtained in the validation was ±5%. de Wit et al. (2007a) concluded that the average correlation of the HBV predictions and measured data is around 0.9, and the Nash-Sutcliffe efficiency is 0.93.

Hereafter HBV-M (HBV-Meuse) refers to the instantiation of the HBV rainfall-runoff model for the whole of the Meuse basin. The calibration and validation data sets used in HBV modelling were constructed in such a way that the observed and simulated discharges in both data sets in terms of flow volumes, and the number of flood peaks and the overall shape of the hydrographs are similar. However, initially no specific low-flow indices are used neither for calibration nor validation. Therefore in this study the results of the hydrological simulation of the Meuse discharges done by de Wit et al. (2007b) where the model was specifically validated against low-flow indices derived for the period 1968–1998.

Complementary information on data validation can be found in the research done by de Wit et al. (2007a). Their work presents the complete and detailed description of the hydrological data used for the model development.

3 Methodology

In this study two hybrid modelling schemes were tested. In the first one, some HBV-S (sub-basin) models were replaced by data-driven model representations. The second scheme is based on the replacement of the Muskingum-Cunge flow routing model by an ANN model integrating the outputs of the sub-basin models.

3.1 HBV-M model setup

The HBV-M model simulates the rainfall-runoff processes for each sub-basin separately. The sub-basins are interconnected within the model schematization and HBV-M simulates the discharge at the outfall. The schematisation and parameter optimization is derived from the approach proposed by Deursen (2004).

The Meuse basin model has been calibrated and validated using daily temperature (T) and precipitation (P) for 17 locations interpolated from measurement stations, the calculated potential evapotranspiration (E_{pot}) per subbasin, and the discharge (Q) at Borgharen. The interpolation of the different locations was performed using kriging (Stein, 1999).

HBV-M has been run on daily basis using daily temperature, precipitation, potential evapotranspiration and discharge data for the period 1968–1984 (calibration) and 1985–1998 (validation) by Booij (2002, 2005) and fine-tuned (with more detailed data) by Deursen (2004).

The model results in this study have been evaluated against the observed discharge records using (a) the volume errors (mm/yr), (b) the coefficient of efficiency (CoE) for the gauging stations along the Meuse and the outlets of sub-basins and the root mean squared error (RMSE, Eq. 1); (c) the normalised RMSE (NRMSE, Eq. 2) (for comparing the sub-basin models with considerably different flows).

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Q_s^i - Q_o^i)^2}{n}} \tag{1}$$

$$\mathrm{NRMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Q_s^i - Q_o^i)^2}{\sum_{i=1}^{n}(Q_o^i - \overline{Q}_o)^2}} \tag{2}$$

where Q_s is simulated discharge (m^3/s); Q_o is observed discharge (m^3/s); n is the number of observations, and is average observed discharge (m^3/s) over the whole period (all summations are run from time $i=1$ to n).

3.2 Scheme 1 sub-basin models replacement

3.2.1 HBV-S sub-basin models

The objective of the further analysis is to determine the average error contributions of the different sub-basin models (referred to as HBV-S) to the total error of HBV-M, and hence to identify the candidate sub-basin models that would need improvement or replacement. In this modelling exercise the river basin behaviour during different seasons and flow regimes will be also taken into account.

3.2.2 Sub-basin error contribution

The relative error contribution from a particular HBV-S (sub-basin) model is calculated as follows. First, the HBV-M model is run and its root mean square error (RMSE) at the outlet is calculated. Then, according to a given replacement scenario a number of input discharges that are fed into the HBV-M from the HBV-S models are replaced by the corresponding observed discharges. The HBV-M model is run once for each scenario. The resulting RMSEs for each scenario are compared to the RMSE of the standard HBV-M. This gives the possibility of identifying the overall error of a sub-basin model. Such an error contribution is calculated for the different flow conditions (e.g. dry and wet seasons).

The replacements of the sub-basin models is performed in sequence: starts with the Lorraine Sud in the direction downstream towards Borgharen, then one more sub-basin model is replaced, then yet another one, until all selected sub-models are replaced (ending at Borgharen). It is important to stress that the independent replacements of sub-basins will not allow for seeing the accumulative error reduction, which is necessary to have an overall idea of the total error of accumulative areas. Two important assumption are made to be able to visualize the error contribution. First, is that the compensation of errors when adding the basin is minimal in comparison to the error of the basin contribution. The second assumption is based on the additive linear error propagation along the river basin. Assuming non-linear error propagation may lead to complications of interpreting the contributions since there are temporal dynamics that affect the non-linearity.

3.2.3 Data-driven sub-basin models

After the error contribution of the HBV-S models are identified, data-driven models (DDM) can be built for each of the sub-basin models under consideration. Various data-driven techniques are compared to select the representative and accurate DDM.

As candidates for data-driven modelling, several statistical and computational intelligence techniques were tested: ANNs, linear autoregressive models and M5 model trees. Their performances were compared to that of the existing HBV-M model. Apart from that, an attempt was made to recalibrate a number of local HBV models; however, the overall performance obtained was lower than that after the calibration of HBV-M as a whole, and these experiments are not presented here. A detailed reference of the algorithms used can be found in Haykin (1999) and Witten and Frank (2000)

In the case study, before identifying the relative error contribution of various sub-basin models, several types of the DDMs were compared for the 8 of 15 sub-basins. This made it possible to judge if DDM are useful as HBV-S replacements.

Each data-driven rainfall-runoff model for the sub-basins uses precipitation and measured discharge as inputs, and the response discharge of the basin is generated for the moment T time steps ahead. The general DDM forecast formulation can be represented as follows:

$$Q_{t+T} = f(R_t, R_{t-1}, R_{t-2} \ldots R_{t-L}, Q_t \ldots Q_{t-M}) \tag{3}$$

where the optimal lags L for precipitation and M for discharge are obtained through model optimization (these can be different for various forecast horizons T); f is the data-driven regression model, and T is the forecast horizon (e.g. 1 day). In this research several data-driven models are tested; including linear regression model (LR, Kachroo and Liang, 1992), artificial neural networks (ANN, Dawson et al., 2005) and M5 model trees (MT, Solomatine and Dulal, 2003).

Neural network are all trained using the same random seed, with (Levenberg, 1944; Marquardt, 1963). The learning rate was set to 0.1. networks where setup to be homogenous in the use of one hidden layer with sigmoid function, and one linear transfer function in the output layer. Most of the ANN models have different nodes in the hidden layers according to the different number of input variables selected.

Building model trees followed the procedure presented by Witten and Frank (2000). The size of the trees is controlled by fixing of the minimum number of instances in linear regression models at leaves (e.g. four).

3.3 Scheme 2 Integration of sub-basin models

Routing is a common way to integrate sub-basin models of a meso-scale catchment. However, river routing models include hydrodynamic conditions that require a large number of physical measurements. The accuracy is determined by the availability and the quality of these measurements and of the models. Since the cost of the measurements is high, often simplified routing equations are used. In HBV the sub-basin models use simple transfer functions that represent the routing process. However, to link HBV-S sub-basins models the Muskingham-Cunge equation (albeit simplified as well) is used. The routing equation is applied to river reaches where the distance between the outlets of the basins is significant.

The main idea of the Scheme 2 is the replacement of the traditional runoff routing by a more accurate non-linear function, in this case the multi-layer perceptron ANN (ANN-MLP). The output discharges from the fifteen HBV-S sub-basin models are lagged and used as inputs to this model. The lags are determined using the correlation and average mutual information analysis involving different sub-basin flows and the final outflow at Borgharen. The ANN-MLP model has the following input-output structure:

$$Q_{\text{Borgharen}}^{t+T} = f(Q_{t-l_1^1}^1, Q_{t-l_2^1}^1, \ldots, Q_{t-l_M^1}^1, Q_{t-l_1^2}^2, \ldots, Q_{t-l_M^2}^2, \ldots, Q_{t-l_M^N}^N) \tag{4}$$

where the lower sub-index T represents forecast horizon, N is the total number of sub-basins, and l the lag and M^i the number of lags taken per sub-basin i. All basins in the model are lagged with respect to the current flow at Borgharen.

4 Application of Scheme 1: data-driven models for sub-basin representation

4.1 Inputs selection and data preparation for DDMs

Each data set is split into a training set (70%; some data is used for cross-validation as well) and a verification (30%) sets. This procedure is performed in a way that ensures that the training data contains the maximum and minimum values of each variable to reduce the possible extrapolation problems. Additionally, the statistical similarity of each set was verified by comparing its probability density function graph.

The first step in developing data-driven models for the Meuse sub-basins was to identify the most appropriate inputs for predicting the future discharges. Two approaches were used to select the appropriate input variables and their lags: correlation analysis and the average mutual information (AMI), as it was done, for example, by Solomatine and Dulal (2003). A lag is defined as the number of time steps by which a time series is shifted relative to itself (when autocorrelated), or relative to the corresponding time values of another time series (when cross-correlated).

The correlation coefficient and AMI was calculated for 10 lag values. The variables compared where discharge, precipitation and evapotranspiration. Since the correlation analysis reflects only linear relationships and the phenomena are highly non-linear, the analysis based on AMI was employed as well. The AMI between two measurements

x_i and y_j drawn from sets X and Y is defined by:

$$I_{XY} = \sum_{\substack{x_i \\ y_j}} P_{(X_i,Y_j)} log_2 \left[\frac{P_{XY}(x_i, y_j)}{P_X(x_i)P_Y(y_j)} \right] \tag{5}$$

$$P_{(X_i,Y_j)} = \int\int_{XY} f(x,y)dxdy \tag{6}$$

where $P_{(X_i,Y_j)}$ is the joint probability density for measurements X and Y resulting in values x and y, and are the individual probability density for the measurements of X and Y. If the measurements of a value from X resulting in x_i is completely independent of the measurement of a value from Y resulting in y_j then the average mutual information I_{XY} is zero. The probabilities where calculated with different bin sizes and the results were similar. Figure 2 shows the AMI for the Ourthe (a) and Lorraine Sud (b), which represent the sub-basins with faster and slower precipitation-discharge response respectively. The maximum AMI of precipitation-discharge time lag for the Ourthe corresponds to a three-day lag (P_{t-3}), and for the Lorraine Sud sub-basin (Meuse source) up to a four-day lag (P_{t-4}).

Based on a similar analysis as the one presented in the Fig. 2, the following model structure was adopted for eight basins:

$$Q_t = f(P_t, P_{t-1}, P_{t-2}, P_{t-3}, Q_{t-1}) \tag{7}$$

The models were built for: Semois, Viroin, Lesse, Ourthe, Ambleve, Vesdre, Mehaigne, Chiers, Meuse Source; see location Fig. 1. The data used to build each sub-basin model (except Vesdre) covered the period from 1989–1995 for the training set and the period 1996–1998 for testing. Due to the availability of data, for Versdre the period used was from 1992-1996 for training and from 1997–1998 for testing.

4.2 Data-driven sub-basin models

The performance of the HBV-S models was compared with that of several data-driven models (LR, M5P, ANN) (Fig. 3); NRMSE was used as the error measure.

Both MLP and M5P data-driven models outperform the HBV-S models. Only for the Lesse, Ourthe, Ambleve, and Vesdre HBV-S model error is relatively low, but even then it is not comparable with that of the data-driven models. According to Berger (1992), Ourthe sub-basin together with Vesdre and Ambleve are the most important tributaries for flood forecasting, relating area percentage and response time. HBV-M results for Semois, Viroin, and Mehaigne show high NRMSE. The error graphs show that the M5P and ANN models outperform the HBV model for all the considered sub-basins.

However, this does not mean that DDM is unconditionally superior to the conceptual modelling approach. The conceptual model aims to represent the processes of the modelled phenomena (albeit roughly), and the DDM is based on the analysis of historical data. Since the conceptual model only uses the forcing information (precipitation, temperature, etc.), weather forecast information can be effectively used for the longer lead times. Other variables like measured discharges are incorporated in operational systems through the use of external post-processes like data assimilation.

Note that the extended forecast made by DDMs need the previous simulation discharges. Three important implications have to be mentioned here, one is that the use of previous simulation discharge iteratively decreases the quality of the forecast. Second, if we assume that the measured information is a perfect forecast, the HBV average performance will not decrease for the higher forecast horizons. The thirs consequence is that the DDM is not representing the basin behaviour and instead is acting more as an autoregressive model.

The data-driven model, which tends to generate high weight values for input from previous discharges in its structure, underestimate the use of other variables that are poorly correlated with the output. In this sense data-driven models (DDM) can simulate the flow quite accurately (only on average, however, and not in the beginning of a high precipitation event) even without the use of the variables that really drive the phenomena (precipitation and temperature).

The ANN-MLP model outperforms HBV in more cases than M5P does and therefore is selected for the replacement experiments. The results show that DDMs can serve as accurate replacement models for sub-basins. However, when more and more sub-basin models are replaced, there will be less and less hydrological knowledge (encapsulated in process models) left. In addition, for the extended forecast scenarios the weather information is highly important. Therefore an analysis of the overall performance of the model under different replacements is made below. Since there is a large number of possible scenarios of replacing various numbers of models, it is necessary to analyze the river basin behaviour and the relative quality of the individual HBV-S sub-basin models with respect to the overall basin measurements given by the discharge at Borgharen.

4.3 Analysis of HBV-S simulation errors

The changes in the overall model performance (RMSE) on the verification data set as a result of various replacements with measured discharge data are shown in Fig. 4 and Table 1.

The replacement order can be followed by reading Fig. 4 from top to bottom. From the total RMSE of 83.84, Chiers has the largest relative error contribution of 4.53% (10.5% of the total area), followed by Lorraine Sud (referred as Meuse source St Mihiel in Fig. 1) and Lesse sub-basins with the error contribution of 3.86% (12.10% of the total area) and 2.81% respectively. Chiers is the second largest sub-basin of the Meuse and it is known that it commonly influences floods generated by its slow response, Lorraine Sud is also a slow responding basin. Vesdre, Ambleve, Viroin and Ourthe basins closer to the outlet are the most accurate in the HBV-M model and are the ones directly responsible for floods.

Hydrological data is available for 52% of the basin area; however, only 20% of the total errors seem to be attributed to this area. The rest of the error contribution can be associated with the other variables in the system, the modelling capacity of the HBV, as well as the different uncertainties in modelling of the basin. It would also be interesting to identify the error contribution of the Sambre, the largest sub-basin, but this was not carried out due to data unavailability. In Fig. 4, the RMSE contributions obtained by each sub-basin replacement are associated with the measured discharge values.

Since it is well known that seasonality influences this river basin, the error contributions of the HBV-S models in summer (May–October) and winter (November–April) seasons are calculated in terms of the percentage of error with respect to the total HBV-M error; see Fig. 4. The results in Fig. 4 show that there is a homogeneous error contribution from Chiers in both seasons. The model for Lorraine Sud basin has a higher error contribution for summer and a small overall contribution in the winter. Clearly the calibration of the model is well suited for summer conditions where the slow response of the catchment is important for the average discharge in these periods. This is congruent with the size (2540 km^2), which represents approximately 10% of the considered area.

In terms of flood forecasting at Borgharen the most sensitive basins for the HBV-M model distribution are Ourthe, Vesdre and Ambleve. The analysis shows that the Ourthe and Ambleve stream flows do not influence the model in the summer period, but together make a significant contribution to the error generated in the winter season. The contributions of the Mehaigne and Viroin sub-basins do not depend on the season: they have a small and similar error percentage for both seasons.

4.4 Replacements of sub-basin models by ANNs

There are numerous replacement scenarios and these should be identified based not only on the previous error analysis, but also taking into account the river basin behaviour during the different seasons and the different flow regimes. The total number of possible replacement scenarios (combinations of the sub-basin models with the data availability) is too high and it is not feasible to analyze them all. The experiments to replace a sub-basin model were carried out using only 8 scenarios as shown in Table 2.

The scenarios reflect mainly the fact that sub-basins with slow and fast flow responses contribute to different components of the resulting streamflow (mainly low and high flows, respectively). Characterisation of the eight scenarios (R1–R8) is as follows:

- R1: The sub-basin (Chiers) with the largest error contribution, and a slow runoff response.
- R2: Three sub-basins which include the Meuse tributaries upstream of Chooz. These are the highest elevation areas with relatively low slope and slow response during flood situations.
- R3: The two fast responding sub-basins that have high contributions during floods (Berger, 1992).
- R4: The same sub-basins as in R3, but together with the slow response Lesse sub-basin whose model has a high error in summer and a low error in winter.
- R5: The same sub-basins as in R3, but together with the slow responding Semois whose model has a high error in summer and low error in winter.
- R6: Combination of slow and fast responding sub-basins.
- R7: Combinations of slow and fast responding sub-basins, but with a larger area covering 35% of the basin.
- R8: Slow responding sub-basins with a large total area.

Table 2 presents the HBV-M model performance changes as a consequence of the different ANN-S replacement strategies. The following statements describe the interpretation of some of the results:

The effectiveness of the models replacements can be evaluated by analysing the changes in the overall HBV-M RMSE. The last column presents the percentages of the maximum reduction possible in case of implementing a particular replacement scenario.

Comparing sub-basins with similar area and similar discharge we can see where the replacement of models was more successful. For example. R1 and R3 have similar percentage of area (11 and 13, respectively), also similar average discharge contribution (10 and 15, respectively). However, the R1 (ANN-S) model gives a RMSE reduction (85%), which is higher than that for the scenarios R3, corresponding to larger areas and higher average discharge. This is an indicator that low flows play a significant role in the overall process, and also reflects the weakness of the HBV-S models currently used in simulating low flows. Other similar case can be seen when R7 replaced a bigger area (35%) than R8 (28%), however, the efficiency for the latter replacement is significantly higher (95%). In terms of discharge, R8 has a smaller average discharge and therefore less contribution. For the scenarios R6 and R8 results show a similar error reduction after the replacement. They have approximately the same average discharge percentage contribution to the basin and a similar area, however, their seasonal error contribution is different (Fig. 4).

The influence of changing Ourthe and Ambleve for Lorraine Sud shows that most of the errors arise in the low flow modelling. The Lorraine Sud (location of the Meuse source) is the most distant basin with relatively mild slope, and therefore its contribution to flash flood (fast flow and runoff) is minimal. This is consistent with the results of de Wit et al. (2007b), who showed that the peak discharges of Vesdre and Ourthe basins are larger than those of Chooz. The results point to a partial explanation based on the differences in precipitation depths of the region and on the difference in hydrogeological conditions. On the other hand, the basins Ourthe and Ambleve (central part) are closer to the outlet and their individual performances are more sensitive for short time lags and fast phenomena.

The results of simulations for the verification period (last three years) are evaluated by calculating the RMSE and Coeff of efficiency (Fig. 6). A typical section of the hydrograph is extracted in Fig. 5. The shape of the hydrograph with ANN-S replacement is mainly driven by the overall hydrological model. It is possible to see that after the replacement R8 the flows (under 600 m^3/s) are closer to the observed discharge. For flows above 600 m^3/s the HBV-M is hardly affected due to the low influence of the replaced basins during the peak flow events (Fig. 5). This shows that the replacement affects mainly the low flow simulation periods.

5 Application of Scheme 2: integrating sub-basin models by ANN

To build a neural network model for routing preprocessing and input variable identification is required. For this the AMI and cross-correlation analysis were carried out to identify the relation (time lag) between the local sub-basins discharge calculated by the HBV model and the measured discharge at Borgharen. For most of the sub-basins the maximum value of AMI related to the observed discharge at Borgharen is a time lag of 1 day. Exceptions are sub-basins 9 (Sambre), 10 (Ourthe), 13 (Mehaigne), and 15 (Jeker) since the corresponding AMI is at maximum for lags less than 1 day. These results are in agreement with recent researches de Wit et al. (2007a), where it was found that the travel time between the measuring stations of the Sambre and Mehaigne, Ourthe and Jeker to Borgharen is less than half a day. More precise time lags can be obtained with hourly data. The average travel time of the flow between the Semois measuring station (sub-basin 5) and Borgharen is 1 day (Berger, 1992).

The results of the model can be visualized by correlation graphs. High correlations are found between the observed and simulated discharge both for training and verification sets (Fig. 7).

Figure 8 shows the observed and simulated discharges at Borgharen from 2 December 1990 (record 700) to 20 June 1991 (record 900). On average the integrated HBV-ANN model outperforms the original HBV-M model. The recession curve of the hydrograph is clearly closer to the measured curve and what was viewed as the systematic error in the recession curve of the HBV-M model is now corrected. An interesting phenomenon can be observed close to the measured peak: the measurement value goes up and down before it reaches its maximum value. This peak change in the hydrograph is reproduced by the ANN routing model with a relatively small underestimation.

For a 3 years error analysis the HBV-ANN gives RMSE of 58.66 m^3/s. An extended error analysis of nine years verification period shows that the RMSE for the HBV-M and HBV-ANN are 86 m^3/s and 55 m^3/s respectively which is the 36% improvement (Table 3). The coefficient of efficiency is also improved from 0.918 for the HBV-M to 0.967 for HBV-ANN model. For both winter and summer seasons it is clear that the use of ANN for integrating the sub-basin models improves the accuracy.

The application of the ANN-based integration method (Scheme 2) outperformed the sub-basin replacement (Scheme 1). This approach does not only consider the integration of streamflow process, it can be also seen as a data assimilation approach (error corrector). Note that the ANN-MLP model does not target the accurate representation of the physical system dynamics; instead, target is aimed to reproduce the measured value (discharge).

6 Discussion

In this section the results for each scheme are discussed and compared.

6.1 Scheme 1

The results show that the replacing some of the conceptual sub-basin models with data-driven models clearly improves the overall model performance. The low flow error results confined to some of the sub-basins can be corrected without any deterioration in the high flows performance. The operational forecasting system requires more variables information in real time and, therefore, this approach may bring operational advantages.

The knowledge representation can be preserved for the basins which have fast response times. This is justified by the fact that their high flows are mainly driven by precipitation phenomenon. This, and the fact that detailed analysis can be conceptualized on the model makes some of the rules that can be used to define the replacements.

The use of Scheme 1 may well be suited for simulation, but comparative tests with data-assimilation and data-driven approaches of the whole basin may be needed to determine whether the use of data-assimilation in operational system is more accurate or suitable than a simple ANN model of a basin. This analysis will be conducted in the further studies.

6.2 Scheme 2

Applying the integrating ANN model (Scheme 2) leads to a more accurate calculation of the overall river discharge, if compared to both Scheme 1 and to the simplified routing scheme employed in the HBV-M model. Our results in this experiment are in agreement with the work by Chen and Adams (2006) where an ANN model was used to integrate the three basin models (Xinanjiang model, Tank model, and Soil Moisture Accounting model).

The use of physical conceptual and data-driven models in operation should consider the dynamics of the basin. The dynamics of the Meuse basin has hardly changed during the last decade (Tu et al., 2005), so the combination of models seems to be reliable under relative long periods of time (e.g. 3 to 5 years as the validation period of the models presented).

It should be noted that the experiments presented in this paper are based on daily data and are aimed at improving the HBV-M hydrological model. In the subsequent studies it is planned to explore the usefulness of the approaches above under a more detailed and complex framework (daily forecast with hourly data and precipitation forecast information). The challenge in extending of these concepts to hourly-based models relate not only to the non-linearity and dynamics, but also to the influence of human interventions at weirs, sluices, canals, power plants, etc. These aspects are not included in HBV-M model and are part of the motivation to use the data-driven techniques, and, possibly, rule-based techniques allowing for multiple regimes of model operation.

7 Conclusions

This paper explored two schemes of introducing data-driven model components into a semi-distributed process based rainfall-runoff model. The first scheme explored the replacement of HBV sub-basin models by ANN-MLP models using several scenarios. The results show that such approach improves the discharge simulation both in terms of reducing the RMSE and increasing the model efficiency. The improvement was mainly observed for the summer periods for low flows.

The second scheme used the replacement of the routing model (combining the individual sub-basin models) by an ANN, and lead to higher gains in terms of the overall error than the first scheme. Nevertheless, it is important to stress that this latter scheme does not only reproduces the flow, but also the noise in the system. The use of an ANN for routing replacement is not only a simulation tool but also captures the variation in the time series. Therefore, its results can not be interpreted as the accurate representation of the river routing but more as a tool to combine the models and which acts as an error corrector as well.

In general it can be concluded that the both combination schemes have a clear potential in improving the accuracy of the considered class of hydrological models.

As one of the following steps, it is planned to move from daily to hourly data, and from the one-step ahead forecasts to the models that forecast the flow several time steps (hours) ahead. In this case the set of inputs of DDMs may not include the previous values of flow (since thay cannot be measured), and their performance may deteriorate if compared to the conceptual models that do not need the discharge as input and are fed with the precipitation forecast only. A possible answer could be in using the architecture of DDM that would use estimates of flow, or to use an ensemble of conceptual and data-driven models.

There is also a problem of limited and inaccurate data for most of the sub-basins, and this affects the performance of operational systems. A possible way to alleviate this, is to use autoregressive models which are not sensitive to the precipitation, temperature and evapotranspiration. Yet another issue is the estimation of the models uncertainty associated with the inaccuracies in data and model structures. It is planned to explore all these issues and possibilities in the further studies.

Acknowledgements. Measurement data for each sub-basin of the HBV model was collected from the Direction gnrale des Voies hydrauliques Belgian (http://voies-hydrauliques.wallonie.be) and from the Ministre de lecologie et du dveloppement durable (http://www.hydro.eaufrance.fr). The authors appreciate the support through the Delft Cluster research programme of the Dutch Government (project Safety against flooding) and of the UNESCO-IHE Institute for Water Education for supporting this research. We would also like to thank Mr. A. G. Ashagrie who provided us with some valuable data sets.

References

Abrahart, R. J. and See, L.: Multi-model data fusion for river flow forecasting: an evaluation of six alternative methods based on two contrasting catchments, Hydrol. Earth Syst. Sci., 6, 655–670, 2002, http://www.hydrol-earth-syst-sci.net/6/655/2002/. 732

Anctil, F. and Tape, D. G.: An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition, J. Environ. Eng. Sci, 3, Supp. 1, S121–S128, 2004. 732

ASCE: Artificial Neural Networks in Hydrology, II: Hydrologic Application, J. Hydrol. Eng., 5, 124-136, 2000. 732

Ashagrie, A. G., de Laat, P. J., de Wit, M. J., Tu, M., and Uhlenbrook, S.: Detecting the influence of land use changes on discharges and floods in the Meuse River Basin - the predictive power of a ninety-year rainfall-runoff relation?, Hydrol. Earth Syst. Sci., 10, 691–701, 2006, http://www.hydrol-earth-syst-sci.net/10/691/2006/. 731, 736

Berger, H. E. J.: Flow Forecasting for the River Meuse, Technische Universiteit Delft, 13–24, 1992. 734, 735, 743, 748

Bergström, S. and Forsman, A.: Development of a conceptual deterministic rainfall-runoff model, Nord. Hydrol., 4, 147-170, 1973. 734

Booij, M. J.: Modelling the effects of spatial and temporal resolution of rainfall and basin model on extreme river discharge, Hydrolog. Sci. J., 47, 307-320, 2002. 736, 737
Booij, M. J.: Impact of climate change on river flooding assessed with different spatial model resolutions, J. Hydrol., 303, 176-198, 2005. 731, 737
Bowden, G. J., Dandy, G. C., and Maier, H. R.: Input determination for neural network models in water resources applications, Part 1: background and methodology, J. Hydrol., 301, 93-107, 2005. 732
Chen, J. and Adams, B. J.: Integration of artificial neural networks with conceptual models in rainfall-runoff modeling, J. Hydrol., 318, 232-249, 2006. 733, 750
Corzo, G. and Solomatine, D.: Knowledge-based modularization and global optimization of artificial neural network models in hydrological forecasting, Neural networks, 20, 528–536, 2007. 733
Corzo, G. A. and Solomatine, D. P.: Baseflow separation techniques for modular artificial neural networks modelling in flow forecasting, Hydrolog. Sci. J., 52, 491-507, 2007. 733
Dawson, C. W., See, L. M., Abrahart, R. J., Wilby, R. L., Shamseldin, A. Y., Anctil, F., Belbachir, A. N., Bowden, G., Dandy, G., and Lauzon, N.: A comparative study of artificial neural network techniques for river stage forecasting, in: Neural Networks, 2005, IJCNN '05, Proceedings, 2005 IEEE International Joint Conference on Publication Date: 31 July–4 August 2005, 4, 2666–2670, 2005. 732, 740
de Wit, M. J. M., Peeters, H. A., Gastaud, P. H., Dewil, P., Maeghe, K., and Baumgart, J.: Floods in the Meuse basin: Event descriptions and an international view on ongoing measures, International Journal of River Basin Management, 5(4), 279 pp., 2007a. 734, 736, 748
de Wit, M. J. M., van den Hurk, B., Warmerdam, P. M. M., Torfs, P., Roulin, E., and van Deursen, W. P. A.: Impact of climate change on low flows in the river Meuse, Climatic Change, doi:10.1007/s10584-006-9195-2, available at: http://www.knmi.nl/publications/showBib.php?id=3068, 2007b. 731, 736, 747
van Deursen, W.: Calibration HBV model Meuse, Report to Rijkswaterstaat RIZA, Carthago Consultancy, Rotterdam, 2004 (in Dutch). 736, 737
Dibike, Y., Velickov, S., Solomatine, D., and Abbott, M.: Model Induction with Support Vector Machines: Introduction and Applications, J. Comput. Civil Eng., 15, 208-216, 2001. 732
Dibike, Y. B. and Abbott, M. B.: Application of artificial neural networks to the simulation of a two dimensional flow, J. Hydraul. Res., 37, 435-446, 1999. 732

Diermansen, F.: Physically based modelling of rainfall-runoff processes, Ph.D. thesis, TuDelft, 123-150, 2001. 731
Fenicia, F., Solomatine, D. P., Savenije, H. H. G., and Matgen, P.: Soft combination of local models in a multi-objective framework, Hydrol. Earth Syst. Sci., 11, 1797–1809, 2007, http://www.hydrol-earth-syst-sci.net/11/1797/2007/. 733
Fogelberg, S., Arheimer, B., Venohr, M., and Behrendt, H.: Catchment Modelling of Nitrogen Flow with two Different Conceptual Models, Proceedings of Nordic Hydrologic Conference, Tallin, available at: http://www.euroharp.org/diss/proceedings.htm, August 2004. 734
Haykin, S.: Neural networks: a comprehensive foundation, Prentice Hall, 2nd edn., 156–161, 1999. 739
Kachroo, R. and Liang, G.: River Flow Forecasting. Part 2. Algebraic Development of Linear Modelling Techniques, J. Hydrol., 133(1/2), 17–40, 1992. 740
Leander, R. and Buishand, T. A.: Resampling of regional climate model output for the simulation of extreme river flows, J. Hydrol., 332(3–4), 487–496, 2007. 736
Levenberg, K.: A method for the solution of certain problems in least squares, Quart. Appl. Math., 2, 164-168, 1944. 740
te Linde, A. H., Hurkans, R., Aerts, J., and Dolman, H.: Comparing model performance of the HBV and VIC models in the Rhine basin, Geophys. Res. Abstr., 9, 04234, 2007. 731
Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergtröm, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272-228, 1997. 731, 734
Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters, SIAM J. Appl. Math., 11, 431-441, 1963. 740
Nilsson, P., Uvo, C. B., and Berndtsson, R.: Monthly runoff simulation: Comparing and combining conceptual and neural networks models, J. Hydrol., 321, 344-363, 2006. 733
Seibert, J.: Estimation of parameter uncertainty in the HBV model, Nord. Hydrol., 28, 247-262, 1997. 731
Solomatine, D. P. and Dulal, K. N.: Model tree as an alternative to neural network in rainfall-runoff modelling, Hydrolog. Sci. J., 48(3), 399–411, 2003. 732, 740, 741
Solomatine, D. P. and Price, R. K.: Innovative approaches to flood forecasting using data driven and hybrid modelling, in: 6th International Conference on Hydroinformatics, World Scientific Publishing Company, Singapore, 1639–1646, 2004. 733
Stein, M.: Interpolation of Spatial Data: Some Theory for Kriging, Springer, 6–14, 1999. 737

Tu, M., Hall, M. J., de Laat, P. J., and de Wit, M.: Extreme floods in the meuse river over the past century: agravated by land-use changes?, Phys. Chem. Earth, 30, 267-276, 2005. 735, 750
Witten, I. H. and Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann, 62–82, 2000. 739, 740

Table 1. RMSE error contribution to the HBV overall simulation.

Sub-basin	Relative error reduction HBV-M (RMSE)	Area (km^2)	Area (% of total basin)	Observed – simulated (% Volume difference)
Mehaigne	0.87	346	1.65	1.04
Ambleve	1.44	1.050	5.00	1.72
Ourthe	1.89	1.597	7.60	2.26
Lesse	2.36	1.311	6.24	2.81
Viroin	1.08	526	2.50	1.29
Semois	1.35	1.235	5.88	1.61
Chiers	3.79	2.207	10.51	4.53
Lorraine Sud	3.23	2.540	12.10	3.86
Others	67.82	10.188	48.51	80.89
Total HBV error	83.84	21.000	100	100.00

Table 2. Replacement scenarios and the effect of their implementation.

Short name	Replacement of sub-basin	PAR* (%)	ADC** (%)	RMSE reduction with ANN-S	RMSE reduction with ***MD	(ANN-S/MD) (%)
R1	Chiers	11	10	3.84	4.53	0.85
R2	Chiers, Semois, Viroin	19	22	6.17	7.42	0.83
R3	Ourthe and Ambleve	13	15	2.25	3.98	0.57
R4	Ourthe, Ambleve, Lesse	19	21	4.21	6.79	0.62
R5	Ourthe Ambleve, Semois	18	25	3.73	5.58	0.67
R6	Semois, Chiers, Lesse	24	28	7.52	8.39	0.9
R7	Ourthe, Ambleve, Semois, Lesse, Chiers	35	41	8.62	12.92	0.67
R8	Lorraine Sud, Chiers, Semois	28	28	9.47	9.99	0.95

* Percentage of area replaced of the total basin (PAR).
** Average discharge contribution in relation to the total average discharge (ADC). The total average discharge is calculated using the average annual discharge from 1970 to 2000 (280.1 m^3/s).
*** Measured data(MD).

Table 3. Comparison between the HBV-M model and the integrated HBV-ANN model.

	Hydrological year (Nov–Oct)		Winter (Nov–Apr)		Summer (Apr–Oct)	
Model	HBV-1	HBV-ANN	HBV-1	HBV-ANN	HBV-1	HBV-ANN
RMSE(m^3/s)	85 651	54 508	100 023	64 255	71 662	45 559
NRMSE	0.286	0.182	0.273	0.175	0.484	0.308

HESSD
6, 729–766, 2009

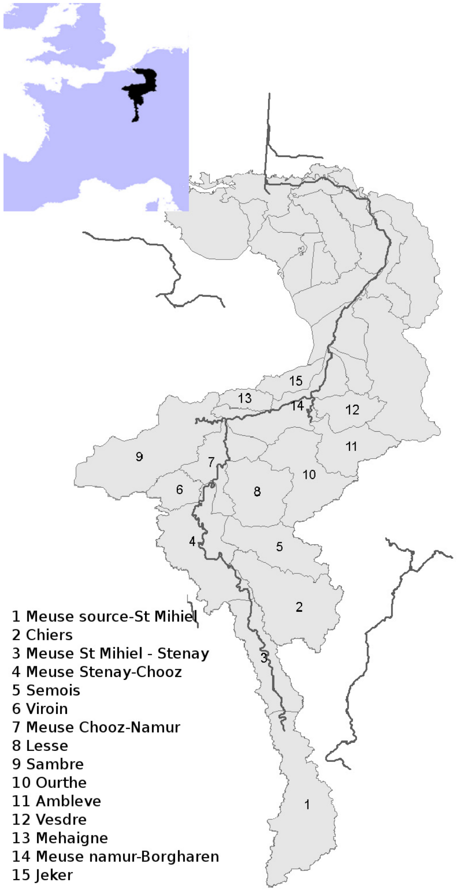

Fig. 1. The Meuse river basin and the sub-basins upstream of Borgharen.

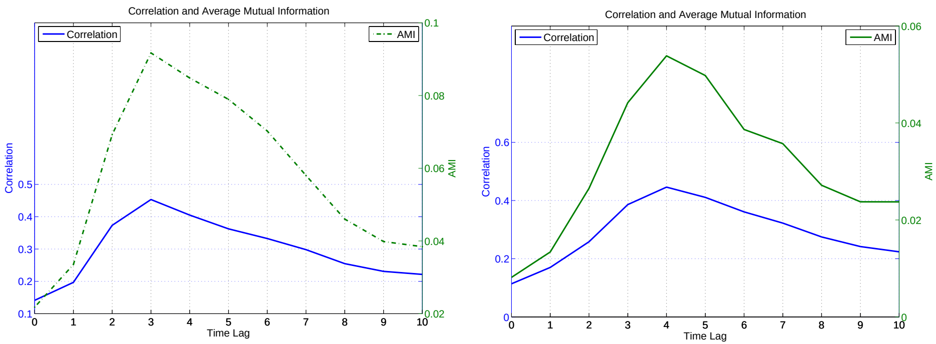

Fig. 2. Average mutual information for between lagged precipitation and discharge for sub-basins Ourthe **(a)** and Lorraine Sud **(b)**.

HESSD
6, 729–766, 2009

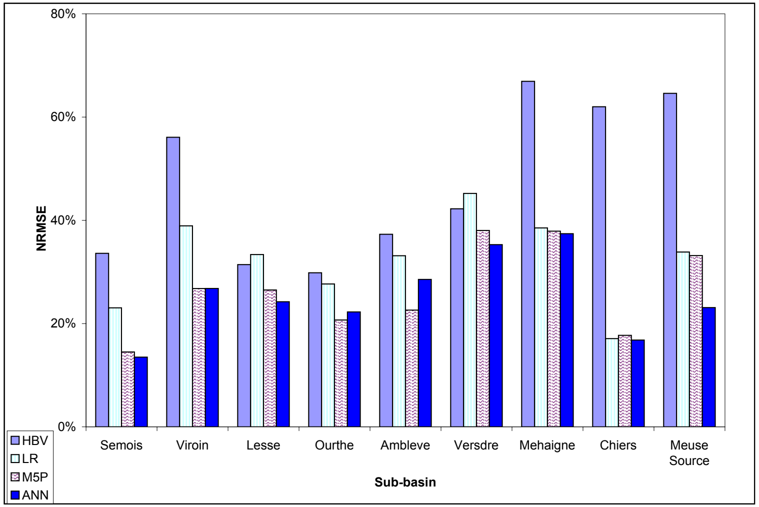

Fig. 3. Comparison of model performance for each sub-basin, expressed in NRMSE of streamflow (Calculated for verification period).

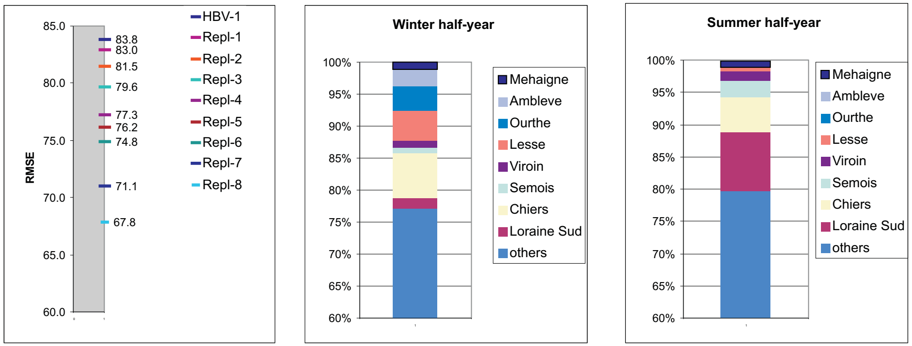

Fig. 4. Reduction in RMSE of the HBV-M due to accumulative replacements of the sub-basin models and seasonal performance.

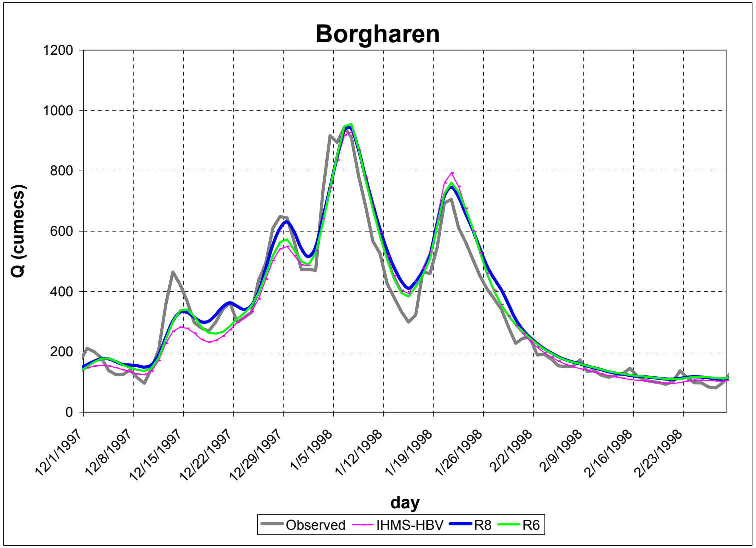

Fig. 5. Hydrograph replacements (R8) and (R6).

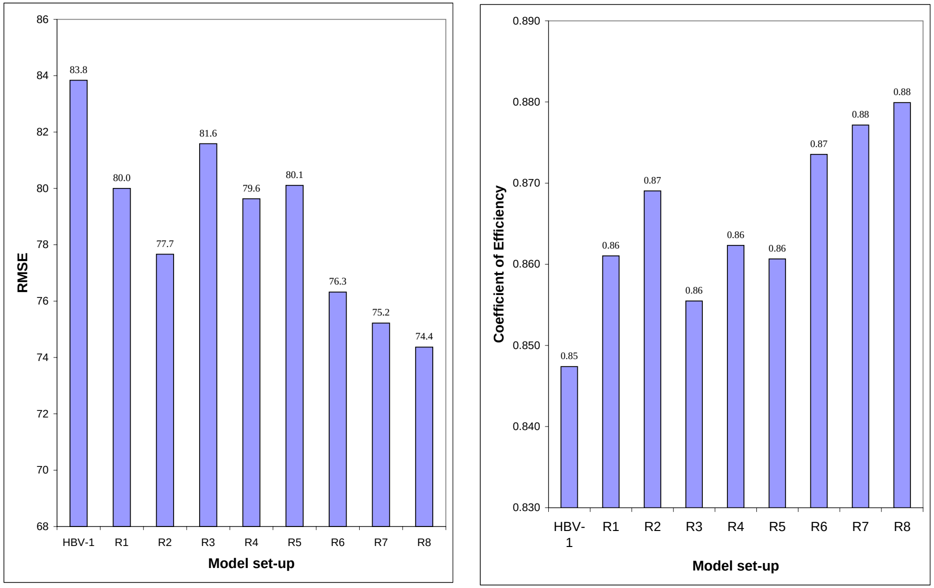

Fig. 6. RMSE (left) and Ceff (right) for different combinations of replacement (evaluated at Borgharen).

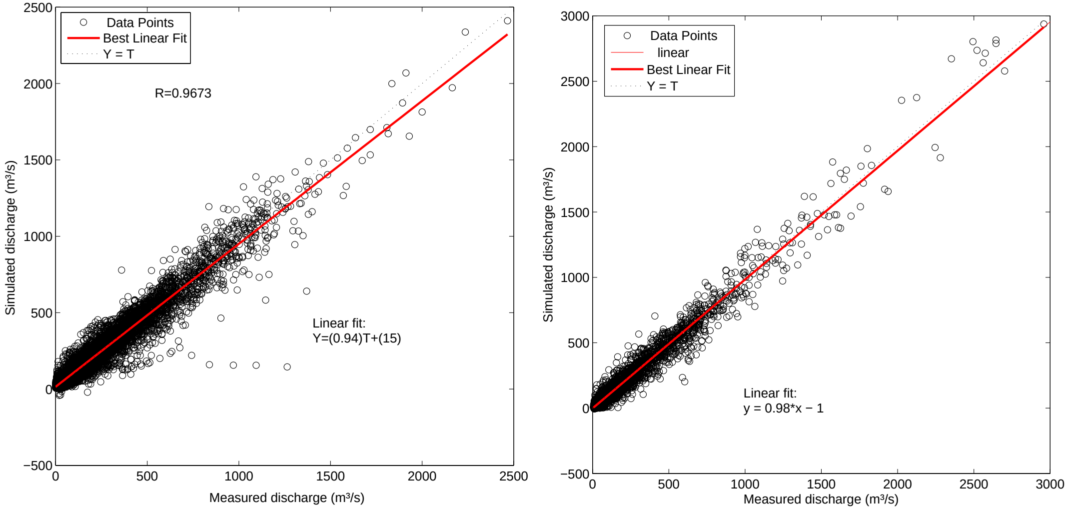

Fig. 7. Scatter plots of target (measured) and ANN model for training and verification period.

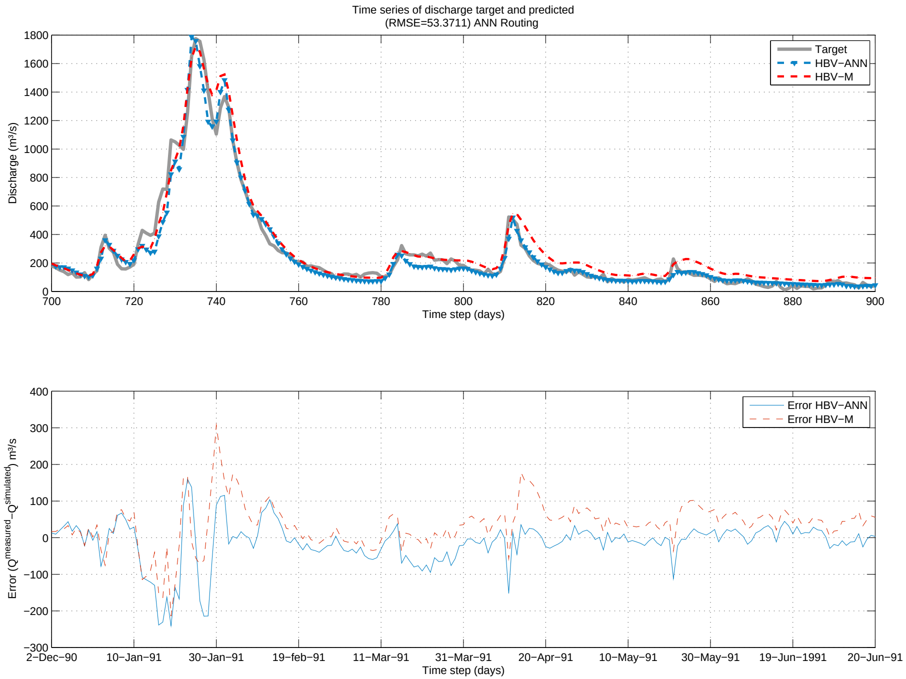

Fig. 8. Hydrograph of the original HBV-M and HBV-ANN integrated models.