Preprints of the
15th IFAC Symposium on System Identification
Saint-Malo, France, July 6-8, 2009

On-line river flow forecasting with 'Hydromax' : successes and challenges after twelve years of experience ⋆

Georges Bastin [*] Luc Moens [*] Philippe Dierickx [**]

[*] *CESAME, Department of Mathematical Engineering, Université catholique de Louvain, 4, Avenue G. Lemaître, B-1348 Louvain-la-Neuve, Belgium (e-mail: Georges.Bastin@uclouvain.be, Luc.Moens@uclouvain.be)*
[**] *SETHY, Direction générale opérationnelle Mobilité et Voies Hydrauliques, SPW, 8, Boulevard du Nord, B-5000 Namur, Belgium (e-mail: Philippe.Dierickx@spw.wallonie.be)*

Abstract: "Hydromax" is the river flow forecasting system for the early warning of extreme hydrological events (floods as well as low waters) in the Meuse river basin and its tributaries. Hydromax provides in real-time short-term predictions of river flows based on past rainfall and river flow measurements, and long-term flood forecasting based on meteorological forecasts. It is now successfully in routine operation for more than twelve years. The purpose of this communication is to give a general description of Hydromax and to report on its performance with typical experimental examples and statistical assessments.

Keywords: ARX model, Conceptual model, Forecasting, Hydrology

1. INTRODUCTION

The early warning for extreme hydrological events in the Meuse river basin is one of the main missions of the Service of Hydrologic Studies (SETHY) of the Walloon Public Administration. The Meuse river is a navigable river which flows through France, Belgium and the Netherlands, see Fig.1. The tributaries of the Meuse are notable for their very varied characteristics, with very high flows in winter and extremely low waters in summer.

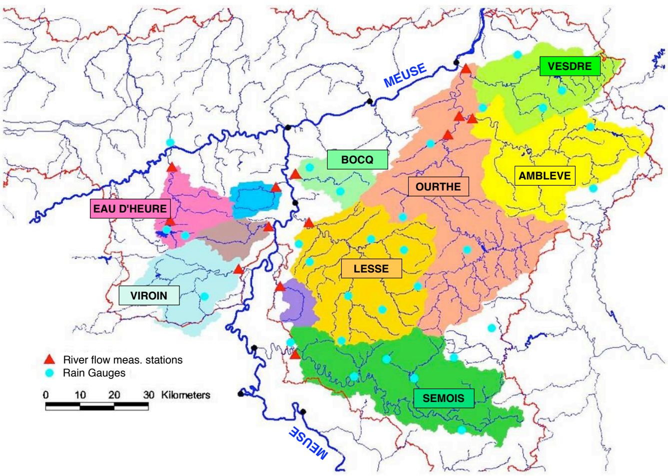

Fig. 1. The Meuse river basin in Wallonia

"Hydromax" is the application that we have developed in collaboration with SETHY for river flow forecasting and flood alarms. Hydromax provides in real-time short-term predictions of river flows based on rainfall and past river flow measurements, and long-term flood forecasting based on meteorological forecasts. Hydromax has been developed to be user friendly and to fulfil the real-time forecasting requirements. It is now successfully in routine operation for more than twelve years in the Meuse river basin (Walloon region, Belgium) and its main tributaries. The purpose of this communication is to give a general description of Hydromax and to report on its performance with typical experimental examples and statistical assessments.

Hydromax is a part of a global integrated forecasting system at SETHY. About 230 hydrologic stations scattered in an area of approximately 20000 km^2 are under permanent on-line monitoring. This forecasting system involves a reliable telemetering network and a data acquisition system able to achieve frequent on-line field measurements of rainfall depths in raingauges, weather radar data, water levels and discharges in rivers as well as positions of mobile weirs in navigable rivers. All measurements are stored in a reliable data base with advanced management tools for data reconciliation, data quality control and data exploitation (visualisation, warnings and internet access). Hydromax is hosted by a computer connected to the data base. In its present state, Hydromax provides on-line river flow forecasting at the outlets of 24 different catchments in the Meuse river basin with areas ranging from 83 to 20140 km^2 (see Table 1). The forecasts are computed with a conceptual/statistical model on the basis of rainfall measurements recorded in 88 automatic telemetered raingauges and two weather radars.

⋆ Research supported by the Belgian Programme on Interuniversity Attraction Poles (IAP VI/4 DYSCO).

River	Catchment outlet	Area (km^2)	Prediction horizon (h)
Ourthe	Nisramont	728	6
Ourthe	Tabreux	1607	6
Ourthe	Comblain	2746	5
Ourthe	Angleur	3607	10
Amblève	Martinrive	1068	4
Vesdre	Chaudfontaine	683	6
Semois	Membre	1226	8
Semois	Tintigny	381	6
Lesse	Gendron	1286	8
Lesse	Daverdisse	302	4
Viroin	Treignes	548	6
Eau d'Heure	Walcourt	191	3
Eau d'Heure	Jamioulx	323	6
Bocq	Yvoir	230	3
Molignee	Warnant	125	3
Hermeton	Hastière	166	5
Houille	Felenne	123	3
Messancy	Athus	63	2
Chiers	Longlaville	152	2
Hantes	Wiheries	142	6
Mehaigne	Huccorgne	305	4
Meuse	Chooz	10120	8
Meuse	Amay	16400	12
Meuse	Lixhe	20440	12

Table 1. River Catchments

2. THE 'HYDROMAX' FORECASTING MODEL

In the recent hydrologic literature, three different types of mathematical models are most often considered for river flow forecasting design : mechanistic models, statistical "black-box" models and conceptual models. The Hydromax forecasting model is a combination of the conceptual and the statistical approaches. The model structure involves three data-based submodels : a conceptual reservoir submodel, a statistical ARX prediction submodel and a long-term forecasting submodel. The data are collected with a basic time-step $\Delta t = 1$h. Hourly rainfall and river flow measurements over a period of several years (including big floods) were thus available for the identification of the models for each cachment. Obviously the basic time-step Δt must be much smaller than the natural response time for each considered watershed.

2.1 The conceptual reservoir submodel

The input of the model is the mean areal rainfall over the considered watershed. It is denoted PB and estimated by Kriging (see Bastin et al. [1984] for details). The main task of the conceptual reservoir submodel (see Fig.2) is to transform the mean areal rainfall PB into an effective rainfall PN which is supposed to reach the watershed outlet as surface runoff. The model describes the balance of water volumes during time intervals Δt. During each time interval the amount of precipitated water is decomposed as follows:

$$PB(t) = PN(t) + E_1(t) + W(t)$$

with t the discrete time index. $E_1(t)$ represents the part of the rainfall $PB(t)$ that directly evaporates during the current time interval. $W(t)$ represents the part of the precipitated water which is stored in the basin under various

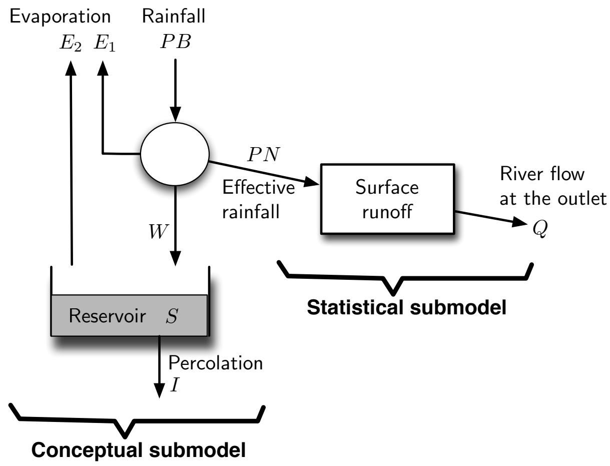

Fig. 2. Structure of the short term forecasting model

forms (vegetation interception, superficial depressions, soil moisture, etc). The storage is represented by a linear reservoir with inflow $W(t)$ described by the difference balance equation

$$S(t) = S(t-1) + W(t) - E_2(t) - I(t)$$

where $S(t)$ denotes the stock of water in the river basin, $I(t)$ is the amount of water drained by percolation and $E_2(t)$ is the part of stored water evapotranspirating during the current time interval. The percolation term $I(t)$ is represented by a linear function of the available water stock:

$$I(t) = \alpha(S(t-1) + W(t))$$

with α a specific percolation parameter. The evapotranspiration terms $E_1(t)$ and $E_2(t)$ are computed as

$$\begin{aligned} E_1(t) &= \min\left[PB(t), ETP(t)\right] \\ E_2(t) &= \max\left[0, \min\left(ETP(t) - PB(t), S(t-1)\right.\right. \\ &\qquad \left.\left. + W(t) - I(t)\right)\right] \end{aligned}$$

where $ETP(t)$ is a periodic forcing input of the model which represents an estimate of the seasonal potential evapotranspiration for the considered catchment (see Wery [1990] for details). It is furthermore assumed that there is a physical upper limit $S_{\max}$ of the amount of stored water $S(t)$ in the river basin. The water storage term $W(t)$ is then expressed as a function of $S(t)$ and $PB(t)$ in order to

- guarantee the condition $0 \leqslant S(t) \leqslant S_{\max}\ \forall t$;
- verify the hydrological principle that the effective rainfall $PN(t)$ increases with both rainfall intensity $PB(t)$ and water stock $S(t)$.

The following function (proposed in Lorent and Gevers [1974]) satisfies these requirements:

$$W(t) = [S_{\max} - S(t)]\left[1 - \exp\left(-\beta \frac{PB(t) - E_1(t)}{S_{\max} - S(t)}\right)\right]$$

with β a specific runoff parameter. The conceptual submodel thus involves three positive parameters $(S_{\max}, \alpha, \beta)$ which are calibrated for each considered watershed. In addition, we must have $0 \leqslant \beta \leqslant 1$ in order to guarantee $PN(t) \geqslant 0$.

2.2 Statisitical ARX prediction submodel

At each time step t, a forecast $\hat{Q}(t+h)$ is computed for the future time instant $(t+h)$ (i.e. with a prediction horizon of

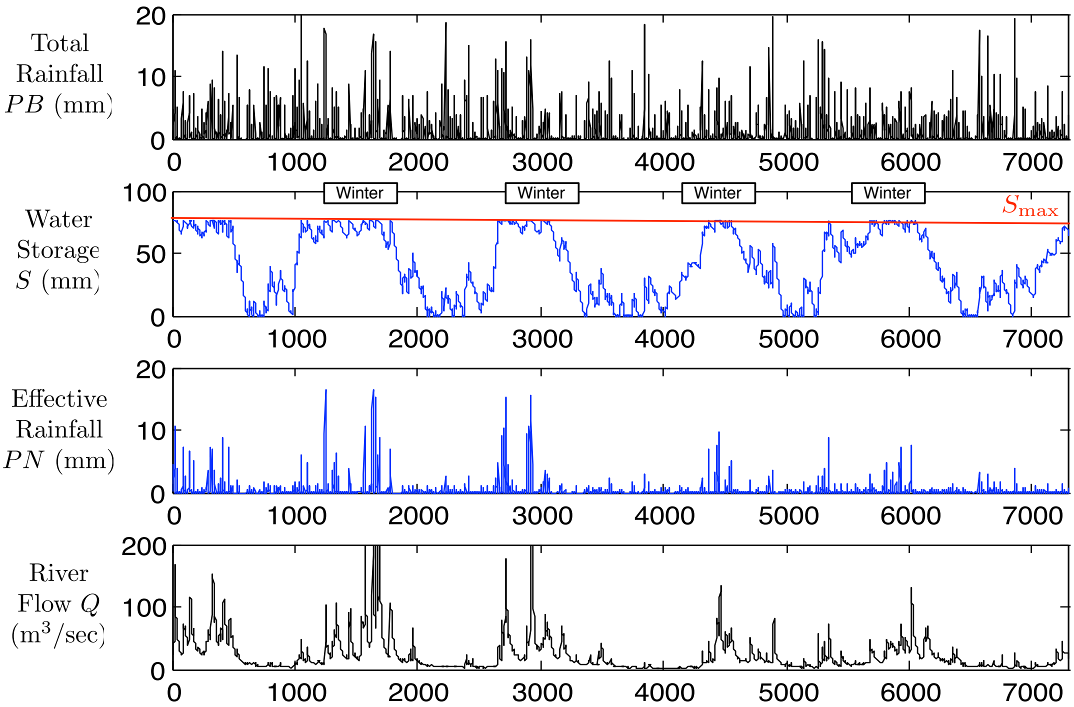

Fig. 3. Illustration of the performance of the conceptual reservoir submodel for the Ourthe river at outlet Tabreux. Data of the period from 1 January 2001 to 31 December 2005. 1 time unit = 6 hours.

h measurement time steps) as a linear combination of past river flow measurements and past effective rainfall values, with a linear regression model (ARX model) of the form

$$\hat{Q}(t+h) = \sum_{i=1}^{n} a_i Q(t-(i-1)h) + \sum_{j=1}^{m} b_j PN(t-(j-1)h)$$

where $Q(t-(i-1)h)$ denotes the river flow measurements at the past time instants $(t-(i-1)h)$ while $PN(t-(j-1)h)$ represents the effective rainfall cumulated over h successive time steps and computed with the conceptual submodel. For each catchment, the values of the coefficients a_i, b_j are determined from experimental data by linear regression. To get accurate forecasts, the prediction horizon h must obviously be smaller than the natural response time of the river basin. As a rule of thumb, it is selected between the one fifth and the one third of the peak time of the unit hydrograph (see Table 1).

2.3 Long term forecasting submodel

The goal here is to compute river flow forecasts over prediction horizons that are significantly larger than the natural response time of the river basin. This requires to anticipate the future rainfalls by using meteorological informations. Such long-term river flow forecasts are computed by iterating the short-term prediction model as $\hat{Q}(t+kh) =$

$$\sum_{i=1}^{k-1} \left(a_i\hat{Q}(t-(i-1)h) + b_i\hat{PN}(t-(i-1)h)\right)$$
$$+\sum_{j=k}^{n} a_j Q(t-(j-1)h) + \sum_{j=k}^{m} b_j PN(t-(j-1)h)$$

where $\hat{Q}(t+(i-1)h)$ are successive iterated river flow forecasts and $\hat{PN}(t+(i-1)h$ are effective rainfall forecasts to be provided by the user.

3. HYDROMAX PERFORMANCES

Hydromax has now been in operation at SETHY for more than twelve years. Hydromax was set up for the first time during the big flood of 1995 (Fig.4). From that time, Hydromax has been run without interruption. In this section the excellent performances of Hydromax during the period 1995-2008 will be outlined. We shall especially focus on years 2001-2002 characterized by two extreme floods in the Meuse river and its tributaries.

3.1 Performance of the conceptual reservoir submodel

The performance of the conceptual reservoir submodel is illustrated here with the case of the Ourthe River basin (see Fig.1 and Table 1). The outlet of the river basin is located at Tabreux (1607 km^2) and four raingauges are used for the mean areal rainfall estimation. Hourly rainfall-discharge data during 2 years (1992-1993) were used to calibrate the model. The estimated model parameters are

Maximum stock: $S_{max} = 7.6$ cm,
Runoff coefficient: $\beta = 0.86$,
Percolation parameter: $\alpha = 0.00065$.

The performance of the conceptual reservoir model is illustrated in Fig.3 for the period 2001 - 2005. In this figure, the experimental data of the mean areal rainfall $PB(t)$ and of the flowrate $Q(t)$ at the outlet are presented. Moreover the two main quantities that are computed by the model are also presented namely the stock $S(t)$ of water stored in the catchment and the effective rainfall $PN(t)$. Three key observations emerge from this figure.

(1) Although the intensity of the precipitations is rather uniformly distributed all along the year, the variations of the flow rate have a strongly marked seasonal behaviour with several big floods in winter and very low waters in summer. Some of the winter floods, as in 2001 and 2002, may be extreme with a flow rate exceeding the *pre-alarm* (160 m^3/s) and alarm (200 m^3/s) thresholds.

(2) The seasonal fluctuations are still much more apparent when looking at the water storage $S(t)$ which is characterised by an annual cycle with oscillations between saturation ($S(t) = S_{\max}$) in winter and dry periods ($S(t) \approx 0$) in summer. Obviously the river floods occur only when the basin reservoir is saturated because it is a necessary condition to have high effective rainfalls and high runoff. But in contrast, it is clear also that reservoir saturation, possibly for rather long periods, does not always induce big floods even in case of important rainfalls as it can be seen in 2003-2005.

(3) Fig.3 also illustrates the efficiency of the computation of the effective rainfall $PN(t)$ whose production appears to be well related with the occurrence of high flow rates : in simple terms we can say that the total rainfall is filtered in such a way that there is a production of effective rainfall only when it is necessary to produce a river flood.

3.2 Performance of the statistical short-term prediction submodel

It is an evidence that the most important objective of the development of Hydromax is to have a tool able to predict accurately the extreme floods that are recurrent in the Meuse river basin. Extreme floods are defined as the floods that exceed the **pre-alarm threshold** which means that the water starts to overflow the banks of the river. When the **alarm threshold** is exceeded, the inundation starts to flow in the urban areas near the river. It is clear that a precise forecast of the flood rise is especially important. For the 24 catchments under the supervision of Hydromax, about 150 such extreme floods have occurred during the period 1995-2008. Computing the prediction errors for all the catchments during the rising periods of all the extreme floods, the global performance of Hydromax may be assessed by considering the

$$\text{performance index} = \sqrt{1 - (\sigma_e^2/\sigma_Q^2)} = 0.962$$

where σ_e^2 denotes the prediction error variance and σ_Q^2 the flow rate variance. All the pre-alarm warnings were correctly issued within the preceding time horizon (no false alarms). Furthermore 82 % of the warnings were issued with a temporal prediction accuracy smaller than 1 hour.

The operationality and the efficiency of Hydromax are also illustrated here with the specific example of the flood of February 2002 for the catchment of the Meuse river at Chooz. The selected prediction horizon is $h = 8$ hours. It must be stressed that the model identfication has been

Fig. 4. The flood of January 1995 in the Meuse river basin. (copyright SPW)

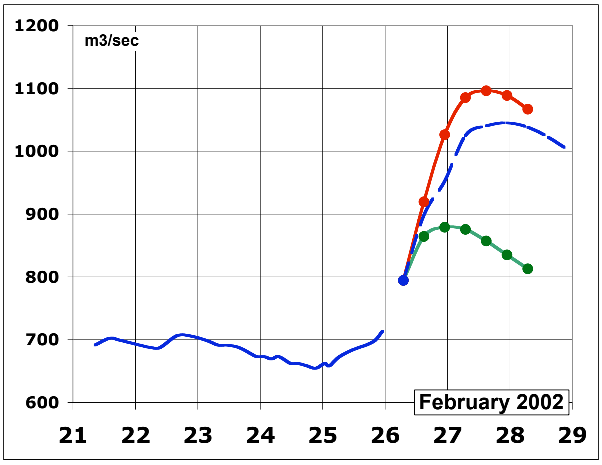

Fig. 5. Forecasting for the Meuse river at Chooz.

based on the data of the period 1995-1997. Thus the flood of February 2002 is not in the model calibration data set and can be viewed as a validation data set.

In Fig.5, a typical example of on-line forecasting with Hydromax is shown. In this figure, the blue solid line represents the measured past river discharge. The forecasting is made on 26 February at 8 a.m. Hydromax computes a short term prediction of 790 m^3/s for 4 p.m. (i.e. 8 hours in advance) represented by a blue dot in the figure. The red and green lines represent long-term predictions that will be discussed in the next section. Fig.6 compiles a set of predictions performed from 8 February to 6 March 2002. We can see that two successive extreme floods have occured. The blue solid line represents the measured flow rate. The blue dots represent a set of 8 h-ahead predictions issued from Hydromax. The excellence of the forecasts is evident here. In particular, it can be observed that the pre-alarm and alarm thresholds as well as the peak values of the floods are predicted with a very high accuracy.

3.3 Performance of the long-term prediction submodel

Let us come back to Fig.5. In this figure, we can see also two long-term prediction profiles computed with the model described in Section 2.3. The red profile is a "pessimistic" forecast which is computed on 26 February under the assumption that rather heavy rainfalls (about 13 mm/day)

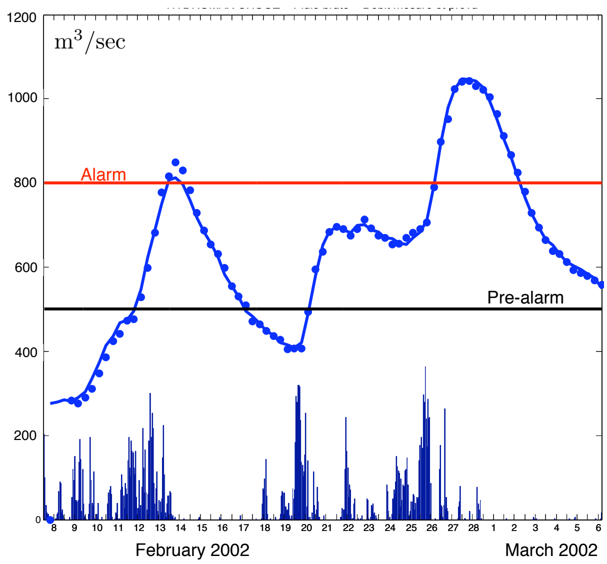

Fig. 6. Short-term forecasts at Chooz in 2002.

will persist for another two days. The green profile is a very "optimistic" forecast which is computed under the assumption that the rainfall will immediately stop forever. The blue dashed line is the record of the flow rate that eventually occured on 27 and 28 february 2008. As it can be seen in Fig.5, the reality in February 2002 was finally not very far from the pessimistic forecast !

4. CHALLENGES AND OPEN ISSUES

4.1 Linear versus nonlinear statistical prediction model

As we have illustrated above, in Hydromax we use a linear statistical ARX model for the forecasting of floods. From the seventies, many successful applications of statistical models based on linear regressions have been reported in the hydrological literature. However, it has been recognised in a number of publications that nonlinear statistical models may provide better results in some instances. The most popular nonlinear models are based on Artificial Neural Networks or Wavelet Transforms, see e.g. Wu et al. [2005], Leahy et al. [2008], Adamowski [2008] for recent references. In Hydromax, we have recently considered another very natural extension of linear regression models by investigating the interest of using so-called NARX models that combine an ARX structure with appropriate static nonlinearities.

A NARX model for rainfall-river flow modelling may be defined as a linear regression model relating nonlinear transformations of inputs (effective rainfalls) and outputs (river flows) as follows:

$$f(\hat{Q}(t+h)) = \sum_{i=1}^{n} a_i f(Q(t-(i-1)h)) + \sum_{j=1}^{m} b_j g(PN(t-(j-1)h))$$

where $\hat{Q}(t+h)$ is the river flow forecast for time $t+h$, $Q(t-(i-1)h)$ are the past river flow measurements and $PN(t-(j-1)h)$ are the past effective rainfall measurements. The functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ and $g : \mathbb{R}_+ \to \mathbb{R}_+$ are monotonically increasing.

In Hydromax, we have shown that a NARX model outperforms the ARX model in the specific case of low waters while for the floods a linear model is the best choice. This is obviously not surprising since, when the floods occur, the river basin is saturated and the runoff is constituted by the totality of the precipitations. In that case, it is indeed well known in hydrology that the river watersheds have often a quasi-linear dynamical behaviour. In contrast, for low waters, the interplay between water storage, effective rainfalls and flow rates becomes more complicated and may therefore require a nonlinear structure not only for the conceptual reservoir part but also for the runoff part. So far, the best model that we have found for low waters is a NARX model with logarithmic f functions and linear g functions.

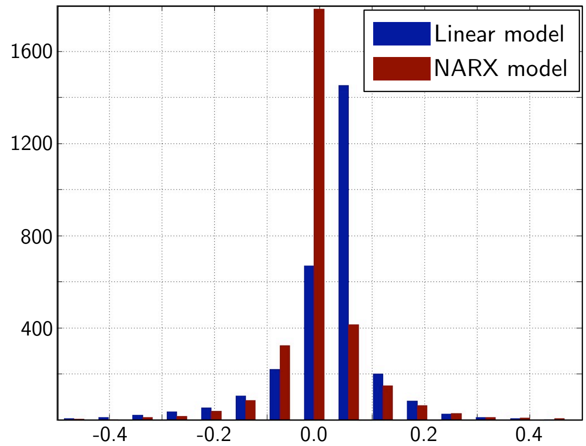

Fig. 7. Distribution of the residual prediction errors with an linear ARX and a nonlinear NARX prediction model.

The performance improvement is illustrated in Fig.7 where the distribution of the residual prediction errors is represented for the Ourthe river watershed at the outlet Tabreux over the period from 1/01/2001 to 30/06/2002 with a prediction horizon $h = 6$ hours (i.e. 2177 short-term predictions). The blue distribution is obtained with the linear ARX prediction model whose performance for the floods has been emphasized in Section 3.2. The red distribution is obtained with a composite model which is made up of the combination of a linear ARX model during the flood periods and a nonlinear NARX model during the low water periods. It can be observed that the blue distribution is not symmetric with a clear bias towards positive prediction errors corresponding to systematic discharge underestimations during the low waters periods. In contrast, it is very apparent that the error distribution with the nonlinear prediction model is symmetric and well centered.

4.2 Rain gauge versus weather radar data

As we have mentioned in Section 2, in order to run the Hydromax forecasting model we need to compute the input mean areal rainfall $PB(t)$ at each time step. This signal is not directly available and is estimated from telemetered local rain gauge data. Rain gauges provide accurate pointwise measurements of the rainfall field. But rain gauges may also be quite scattered in an area of the size of a watershed. For instance in the Ourthe catchment at Tabreux and the Semois catchment at Membre the density is about one rain gauge per 175 km^2. Such a low density may prevent detecting correctly the spatial variability of the precipitation field as is illustrated in Fig.8 for the Semois catchment. In contrast, weather radars can

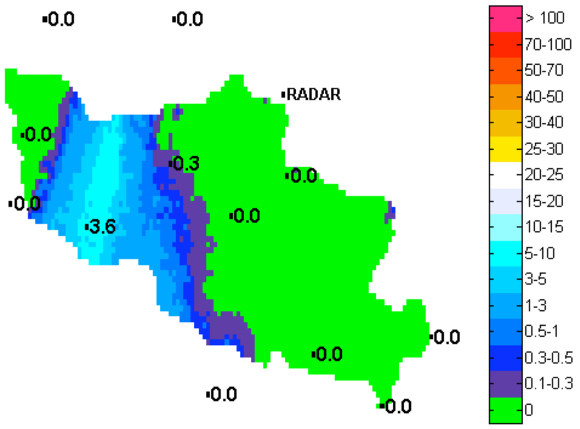

Fig. 8. Semois watershed at outlet Membre : hourly radar rainfall (mm/h) and local values at several rain gauges (1 may 2002 at 9 pm)

detect precipitations with a good spatial resolution. For the Meuse river basin, the Wideumont radar, operated by the Royal Meteorological Service (RMI), produces measurements every five minutes with a resolution less than one km^2. However it is a well known fact that the quantitative rainfall measurements provided by weather radars are often inaccurate and biased. One of the main error sources is due to non uniformity of the vertical reflectivity profile. Therefore both rainfall measurement systems (rain gauges and weather radar) seem to have complementary advantages. (see e.g. Cole and Moore [2008]).

Here we are mainly interested in comparing the flow prediction accuracy obtained with Hydromax when using two different estimators for the hourly mean areal rainfall. The first one is the Kriging estimator based on local rain gauge data and currently used in Hydromax from 1995. The second estimator is obtained from the weather radar data which are cumulated over hourly periods and averaged over the catchment area. Fig. 9 gives an example that both estimators can give very different values with a clear underestimation when the radar data are used. In Table 2, we compare the prediction error variances obtained with two Hydromax models identified using rain gauge and radar rainfall data respectively for the period between 1 February 2002 to 30 November 2003. We see clearly that the bias affecting the raw radar measurements has a negligible impact on the accuracy of the river flow predictions. Obviously this is due to the adaptation of the model parameters that are identified separately for each

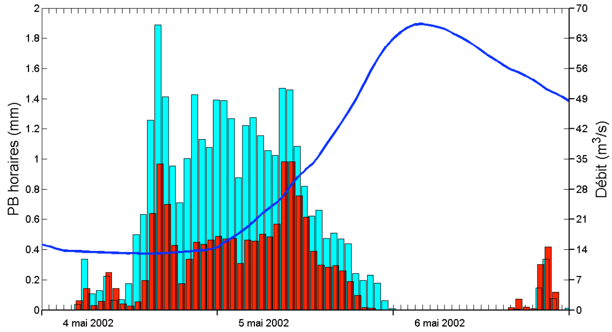

Fig. 9. Mean areal rainfall (mm) estimated from radar data (red) and from 14 rain gauges (blue-turquoise) for the Ourthe catchment at Tabreux (4-6 May 2002). The blue solid line represents the measured river flow (m^3/s).

Watershed	Ourthe at Tabreux	Semois at Membre
Rain gauges	2.542	4.346
Weather radar	2.490	4.583

Table 2. Prediction error variance with Hydromax from rain gauge and radar data $(m^3/s)^2$.

approach. Furthermore, for the considered period and for the two considered basins, we did not find any significant improvement in merging both rain gauge and radar measurements into a single prediction model. However these results need to be confirmed over longer time periods involving extreme floods. Unfortunately the starting of the Wideumont radar was posterior to the last extreme floods!

REFERENCES

J.F. Adamowski. Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. *Journal of Hydrology*, 353:247–266, 2008.

G. Bastin, B. Lorent, C. Duqué, and M. Gevers. Optimal estimation of the average areal rainfall and optimal selecion of raingauge locations. *Water Resources Research*, 20:463–470, 1984.

S.J. Cole and R.J. Moore. Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. *Journal of Hydrology*, 358:159–181, 2008.

P. Leahy, G. Kiely, and G. Corcoran. Structural optimisation and input selection of an artificial neural network for river level prediction. *Journal of Hydrology*, 355: 192–201, 2008.

B. Lorent and M. Gevers. Identificaton of rainfall-runoff processes. In North-Holland, editor, *Proceedings of 4th IFAC Symposium on Identification and System Parameter Estimation,*, pages 735–744, 1974.

B. Wery. *Identification des systèmes hydrologiques - Application à la prévision des crues.* PhD thesis, Universté de Louvain, 1990.

J.S. Wu, J. Han, S. Annambhotta, and S. Bryant. Artificial neural networks for forecasting watershed runoff and stream flows. *Journal of Hydrological Engineering*, 10 (3):216–222, 2005.