Towards an increased performance of flood forecasting through assimilation of remotely sensed soil saturation levels in conceptual rainfall-runoff models

P. Matgen [a,*], J.-B. Henry[b], G. Drogue[a], F. Fenicia[a], R. van den Bos[a], P. de Fraipont[b], L. Hoffmann[a], L. Pfister[a]

[a] Centre de Recherche Public Gabriel Lippmann, 41 rue du Brill, L-4422 Belvaux, Luxembourg – matgen@lippmann.lu
[b] Service Régional de Traitement d'Image et de Télédétection, Bd. Sébastien Brant, BP 10413, F-67412 Illkirch Cedex, France

Abstract –Owing to the non-linearity of the rainfall-infiltration-runoff relationship, soil water content in the river basin represents a key parameter to be monitored for flood management purposes. Remote sensing observations can be used in hydrologic models as a source of time varying hydrologic state data that allows constraining model predictions. The analysis of a series of ERS-1 SAR images showed that the mean backscattering coefficient of selected soil parcels is strongly correlated with a ground-based wetness index the so-called soil saturation index (SSI). This paper shows that SSI values obtained via remote sensing can be used to update the internal saturation states of rainfall-runoff models through the sequential assimilation of the soil moisture information. The assimilation procedure is based on an extended Kalman filter as both simulated and observed saturation states are prone to errors. The magnitude of the correction thus depends on the ratio of errors on the observations and the model. Further research is needed to reduce the uncertainties that remain over the reliability of SAR to provide soil moisture information with a sufficient level of accuracy.

Keywords: Data assimilation, Kalman Filter, flood forecasting, soil moisture, SAR, ERS

1 INTRODUCTION

Rainfall-runoff as well as flood propagation models both greatly benefit from the availability of spatially distributed Earth observation (EO) data especially in ungauged basins. Remote sensing observations can be used as parametric input data, as initial condition data and as time-varying hydrologic state and flux data (Walker, 2005). Soil moisture observations that were derived from radar imagery have been successfully used in the past to improve hydrologic model-based discharge predictions (Pauwels et al., 2001; Aubert et al., 2003). This has been achieved by the assimilation of the statistics of remotely sensed soil moisture patterns in lumped conceptual and spatially distributed physically-based rainfall-runoff models. Soil moisture determines the partitioning of the precipitation into saturated overland flow, saturated subsurface flow and unsaturated subsurface flow. As a matter of fact, the monitoring of this environmental variable during successive wetting and drying up phases helps assessing the readiness of a river basin to generate storm runoff during rainfall events. However, large-area soil moisture networks barely exist at the high-frequency and fine spatial resolution that would be required (Houser et al., 1998). SAR data could be a valuable alternative to retrieve soil moisture but because of the concurrent effect of soil moisture, surface roughness and vegetation on the backscattered signal as well as the speckle inherent to all SAR data sets, there is still significant uncertainty over the reliability of SAR to provide accurate soil moisture data (Schmugge et al., 2002). Moreover, even under optimal conditions, SAR remote sensing can only be used to retrieve soil moisture in the first few centimeters of soil whereas runoff generation is more strongly controlled by deeper layers, especially in regions with a temperate oceanic climate. Hence, the scope of this study has been (1) to establish empirical relationships between radar backscattering and an estimate of the saturation degree of the basin and (2) to assimilate this information into a conceptual rainfall-runoff model in order to increase the reliability of its discharge predictions.

2 STUDY AREA AND AVAILABLE DATA

At its outlet in Hesperange (Grand-Duchy of Luxembourg, Europoe), the upper part of the transnational Alzette river basin has a drainage area of 292 km^2. Discharge measurements are available from 1997 to present. Based on daily rainfall information collected at 5 raingauges, the basin averaged rainfall amounts were calculated. Potential evapotranspiration was estimated with daily meteorological data measured at the synoptic station of Luxembourg airport. Meteorological data are also necessary to guarantee that on the days of satellite overpass the radar signal return is not influenced by frozen soils or high wind velocities. Based on the measurement of the water table depth at 10 piezometric stations scattered throughout the basin's alluvial plain, the wetness of the basin is estimated by the means of a soil saturation index explained in more detail hereafter. The topography of the floodplain is characterized by small elevation changes and an average width of 2.5 km. A SPOT derived land cover classification shows that land use is very homogeneous in the floodplain with permanent pasture largely dominating. As a matter of fact, good conditions are provided for signal change detections based on SAR satellite images. The EO database is comprising 13 ERS-1 and ERS-2 images, acquired on descending pass, with 9 of them during the ERS-1 *Ice Phase*, from 20/11/1993 to 23/02/1994. During this phase the usual repeat cycle of 35 days was shortened and ERS-1 operated with a repeat cycle of only 3 days. Three flood events were covered that occurred by the end of December 1993, in early January 1994 and in January 2003. Care was taken to sample images that represent a broad range of possible moisture conditions. The SAR instrument on board of the ERS satellites is a C band (5.3 GHz) radar, operating in VV

* Corresponding author
** This study is supported by the 'Ministère Luxembourgeois de la Culture, de l'Enseignement Supérieur et de la Recherche' and the French Space Agency (CNES)

polarization with a spatial resolution of 30 m and a pixel size of 12.5 m. The incidence angle is 23 degrees. Speckle noise is reduced using the Frost filtering with a 5x5 kernel and a coefficient of variation that equals 5.

3 METHODOLOGY

The sequential assimilation of remote sensing observations represents a possible step forward in order to improve the accuracy of the production function in lumped conceptual hydrologic models. The assimilated measurements therefore need to be representative of the whole catchment at a given time. Whereas the time variation of the water content in the first few centimeters of soil is only loosely connected to the time variation of the water budget over the entire basin, fluctuations of the water table depth in the floodplain are more representative of the time variation of the basin hydric state (Pfister et al., 2003). However, unlike more aggregate components such as river discharge, water table levels strongly reflect small-scale characteristics of the basin. Since individual point measurements of water table depth do not fully account for the state of saturation at a large scale, a spatial averaging procedure is required. Based on the recorded minimum and maximum water table depth at each available piezometer, a regional mean soil saturation index is computed (Matgen et al., submitted). When the water table reaches its all time minimum depth, the SSI is 100% i.e. we assume that the soil is completely saturated. The SSI decreases linearly until the measured water table depth reaches its all time maximum value (SSI=0%). This study is limited to the highly permeable floodplain area of the river Alzette because in the shallow groundwater areas a strong bound exists between the water table and the water content in the first few centimeters of soil (Chen and Hu 2004). The simultaneous evolution of the water budget at a catchment scale and the mean SSI led us consider that the estimation of a floodplain based SSI provides some valuable information on the expected runoff generation during upcoming storm events.

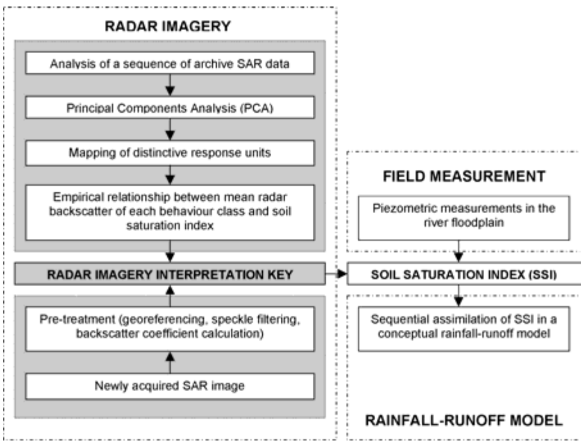

Figure 1. Assimilation of remote sensing observations.

The methodology that was adopted is summarized in Figure 1. Principal Components Analysis (PCA) was applied on archive SAR data in order to outline hydrological response units with distinctive empirically obtained "backscattering-SSI" functions. This preliminary work provides a radar interpretation key that allows deriving the SSI from a newly acquired SAR scene. The SSI that is obtained with SAR or with field measurements is assimilated into a conceptual rainfall-runoff model in order to improve the reliability of its discharge predictions.

4 REMOTELY SENSED SATURATION LEVELS

The PCA is applied to the time series of ERS SAR scenes of the river Alzette floodplain. These images cover markedly different wetness conditions during several winter seasons in order to study the decrease of the radar backscattering signal during drying-up phases following important flood events. At the floodplain scale, with homogeneous land use and constant topography, the first principal components (PC) are mainly dominated by the variance related to the changing areas. The PCs are thus mainly controlled by subsurface and surface water dynamics. A classification scheme, based on the principal components and k-means algorithm, leads to the segmentation of the floodplain into several hydrological behaviour classes with distinctive responses versus changing moisture conditions (Figure 2). To validate this classification method with ground based estimations, the relation between the mean backscattering values of groups of 50x50 m microplots within each PCA-derived class and the water table measurements, expressed by the means of the SSI, are evaluated. Results show that each class of microplots is characterized by the slope of the "backscattering-SSI" function and by the SSI threshold value at which groundwater resurgence appears. When the SSI approaches 100%, the outcropping of the water table in the hollow of the concave footslopes produces groundwater resurgence over an expanding area. The water ponding implies very low signal return due to the specular backscattering effect on the water surface.

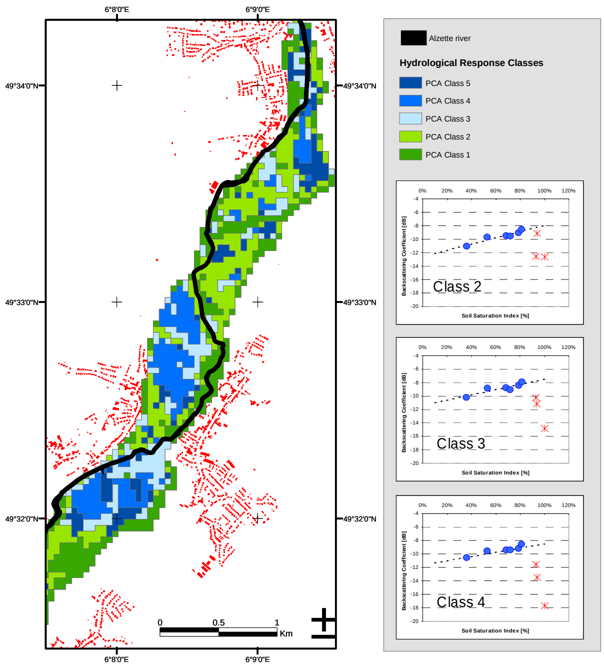

Figure 2. Hydrological response units and the variation of radar backscattering with changing moisture conditions.

By using the empirical SSI-backscatter model (Matgen et al., submitted), a global soil saturation index for the day of the satellite overpass is computed. This empirical model was found to give good results when plots with non-significant relationships with moisture conditions were no longer considered. The empirical model is based on the assumption that soil moisture near the drainage network and the piezometric recordings are strongly correlated and that the water table depth much more than the point measurement of soil moisture, represents an aggregate measurement of the basin wide saturation level. Hence, the SSI can be related to the saturation level of the soil reservoir of the rainfall-runoff model. The latter determines the infiltration during storm events and the evaporation and drainage between storm events. The relationship between the backscattering and the hydrologically relevant SSI still needs to be strengthened. At this point the uncertainties of SAR derived moisture indices remain high. The results presented hereafter focus on the potential application of the SSI as it can be obtained by the means of such a SAR interpretative regression model.

5 ASSIMILATION PROCEDURE

5.1 Rainfall-runoff Model

The hydrologic model used in this study is a 11-parameter lumped conceptual model, which simulates daily discharge using rainfall and potential evapotranspiration as input fluxes. The model is a version of the widely used HBV model (Bergström, 1995) that has been adapted to the particular environment of the modelled catchment. The conceptual model was further modified to allow for the assimilation of saturation state variables. The soil reservoir is characterized with a parameter of maximal storage capacity, Smax [mm], a parameter of non-linearity, b [-], describing the production function of runoff and, finally, the maximum percolation rate percmax [mm/d]. The rainfall is divided into two terms: a first part fills the soil reservoir which is drained by deep percolation and evapotranspiration and a second part, the net rainfall, fills the two routing reservoirs (linear baseflow reservoir and non-linear fast runoff reservoir) according to a ratio depending on the soil saturation. The recession rate of the two routing reservoirs differs. As a continuous model it can predict antecedent wetness (i.e. summer versus winter conditions in a humid temperate hydrological regime).

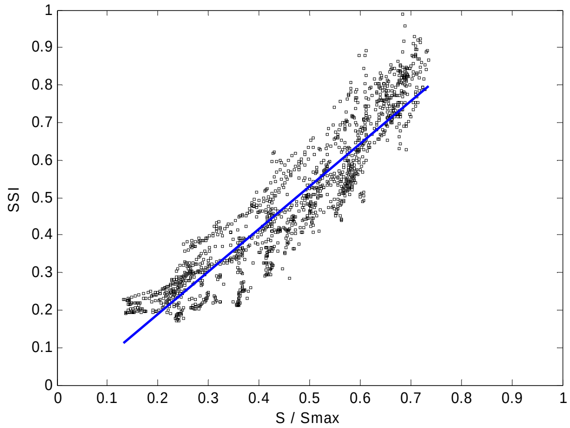

Figure 3. Relationship between the state of the soil reservoir (S) and the observed floodplain based SSI.

The model's parameter values were not observed and had to be estimated through model calibration. A Monte Carlo framework is adopted to sample the parameter space and the predictions of individual model parameterisations are weighted according to the model fit to the observed discharge and water table depths. Hence, the performance measure is an additive combination of the Nash evaluation criterion and the coefficient of determination of a linear regression analysis of the SSI-S(t) relationship (Figure 3). The latter expresses the agreement of fit between the simulated water content in the soil reservoir and the floodplain based SSI. Although dotty plots reveal many good fits across the parameter ranges implying that the individual parameter values are less important than the parameter set as a whole, only the best performing parameter set is retained.

5.2 Sequential assimilation

The states of the model that represent the storage of water in the root zone, can be updated with field measurements. It is based on the assumption that a better simulation of the model states at day j will also improve the accuracy of the model states at days j+1, j+2, etc. (Aubert et al., 2003). After each time step of simulation, coincidental observations are sought. A control simulation without any data assimilation can be considered as the baseline run i.e. we assume that the real-time observations contain no valuable information. On the contrary, if a direct assimilation takes place, it is assumed that the model contains no information at all. Refsgaard (1997) states that among different data assimilation methods, the state updating methodology is the best suited for non-linear models. Hereafter, the method consists in correcting the internal state of the model that accounts for soil moisture, i.e. S(t), whenever a field measurement of SSI is available. The risk of severe model failures due to a wrong assessment of the antecedent moisture conditions will become less high because even in calibrated models, without any assimilation of ground data, the internal state data of the model are inherently uncertain. Thus, besides improving the reliability of the model forecast, the sequential assimilation allows increasing the internal consistency of the conceptual rainfall-runoff model. Pauwels et al. (2001) come to a somewhat similar conclusion and state that one of the main reasons to do data assimilation is to reduce the need for model calibration and to reduce the effect of uncertainty in certain parameter sets on the model results.

The first step in data assimilation consists in establishing the relationship between the observed SSI and the level of the soil reservoir S. Figure 3 shows the SSI derived from the piezometric recordings against the simulated level of the soil reservoir without assimilation. A linear relationship is chosen to relate the SSI to the model state whenever a measurement is available. A coefficient of determination of 0.88 shows that the linear regression is acceptable. In this short summary we only consider a particular case of the Kalman filtering (forced mode) where we assume that the uncertainties on the observations are very small compared to those that are associated to the model. Thus with each assimilation step the current level of the soil reservoir is changed to become the value derived from the ground observation. Generally, however, both forcing terms and model output are known to be uncertain. The model gives a state estimate with high temporal resolution, but the values are altered by the accumulation of errors. Measurements give an alternative estimate that is usually more accurate, but they are sparsely distributed in space. Hence, representation error is typically the main error source that needs to be considered (Sorensen and Madsen, 2004). The magnitude of the correction should therefore depend on the ratio of the errors on the observations and the model and the best estimate of the true saturation state should be obtained based on the available information from the two sources of information.

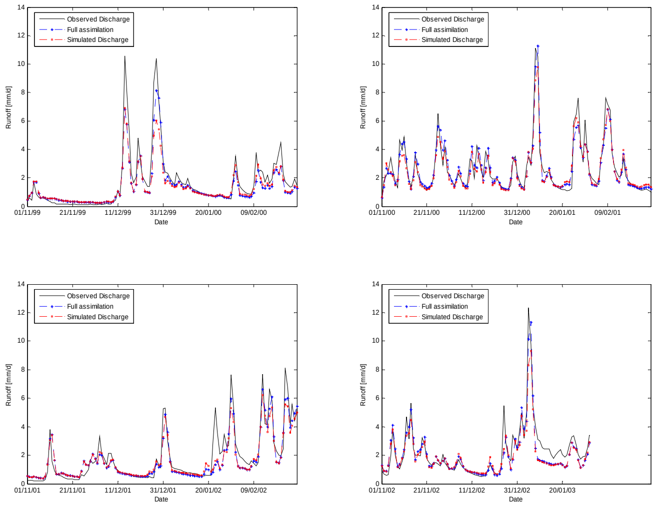

Figure 4. Simulated discharges compared to the observed discharges with and without data assimilation

To measure the efficiency of the assimilation procedure, the simulated discharges of the Alzette catchment are compared to the observed discharges with and without data assimilation (Figure 4). It is clearly shown that the updating of the level of the soil reservoir allows improving the simulation of the high flows. Without assimilation the peak discharges of the high flow events were generally underestimated by the model (e.g. floods that occurred in December 1999, January 2001 and January 2003). These results suggest that during flood events following long periods with sustained rainfall, the calibrated model underestimates the antecedent moisture conditions. In this case, the updating of the water content in the soil reservoir through the assimilation of the SSI helped improving the model performance. This is not a general rule however and the difference between simulated and observed discharges sometimes increased after the ground-based estimation of the SSI is assimilated into the model. Floods that occurred following a long period without rainfall, are less well simulated with assimilation than without. This result suggests that during dry periods the water table depth in the floodplain is no longer representative of the overall hydric state of the basin. Obviously, the assimilation procedure has no improving effect at all on the model's performance during low flows.

6 CONCLUSION

It is known that the backscattering of the radar signal emitted by active microwave sensors is highly influenced by the soil surface conditions, especially the water content of the first few centimeters of soil. The present study showed that hydrologically relevant information was derived from SAR imagery and was successfully used to improve discharge predictions through data assimilation. As this index can be derived from SAR imagery, remote sensing techniques show a potential of becoming an important asset for many flood forecasting applications. However, the dispute whether accurate spatially distributed soil moisture values can be obtained from SAR images is wide spread and far from being resolved and, to date, many doubts exist over the reliability of Synthetic Aperture Radar (SAR) to provide accurate soil moisture information. Furthermore the imaging frequency of currently available spaceborne SAR still appears to be insufficient for operational applications. Since several spaceborne SAR instruments providing high spatial resolutions and multi-polarisation capabilities will be mounted on satellites to be launched from 2005 onwards, the upcoming years will certainly decide upon the future of radar imagery in hydrology.

6. REFERENCES

D. Aubert, C. Loumagne, and L. Oudin, "Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model," Journal of Hydrology, vol 280, p.p. 145-161, 2003.

S. Bergström, "The HBV model," Computer Models of Watershed Hydrology, edited by V.P. Singh, Water Resources Publications, Highlands Ranch, Colorado, p.p. 443-476, 1995.

X. Chen, and Q. Hu, "Groundwater influences on soil moisture and surface evaporation,“ Journal of Hydrology, vol 297, p.p. 285-300, 2004.

P.R. Houser, W.J. Shuttleworth, J.S. Famiglietti, H.V. Gupta, K.H. Syed, and D.C. Goodrich, "Integration of soil moisture remote sensing and hydrologic modeling using data assimilation," Water Resources Research, vol 34, p.p. 3405-3420, 1998.

P. Matgen, A. El Idrissi, J.B. Henry, N. Tholey, L. Hoffmann, P. de Fraipont, and L. Pfister, "Patterns of remotely sensed floodplain saturation and its use in runoff predictions,“ Hydrological Processes, submitted.

V.R.N. Pauwels, R. Hoeben, N.E.C. Verhoest, and F.P. De Troch, "The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation," Journal of Hydrology, vol 251, p.p. 88-102, 2001.

L. Pfister, G. Drogue, A. El Idrissi, J. Humbert, J.F. Iffly, P. Matgen, and L. Hoffmann, "Predicting peak discharge through empirical relationships between rainfall, groundwater level and basin humidity in the Alzette river basin, Grand-Duchy of Luxembourg," Journal of Hydrology and Hydromechanics, vol 51, p.p. 210-220, 2003.

J.C. Refsgaard, "Validation and intercomparison of different updating procedures for real-time forecasting," Nordic Hydrology, vol 28, p.p. 65-84, 1997.

T.J. Schmugge, W.P. Kustas, J.C. Ritchie, T.J. Jackson, and A. Rango, "Remote sensing in hydrology," Advances in Water Resources, vol 25, p.p. 1367-1385, 2002.

J.V.T. Sorensen, and H. Madsen, "Data assimilation in hydrodynamic modelling: on the treatment of non-linearity and bias," Stochastic Environmental Research, vol 18, p.p. 1-17, 2004.

J. Walker, "Towards remote sensing for hydrologic prediction in ungauged basins, "Proceedings of the IAHS General Assembly, Foz do Iguacu, Brazil, 2005.