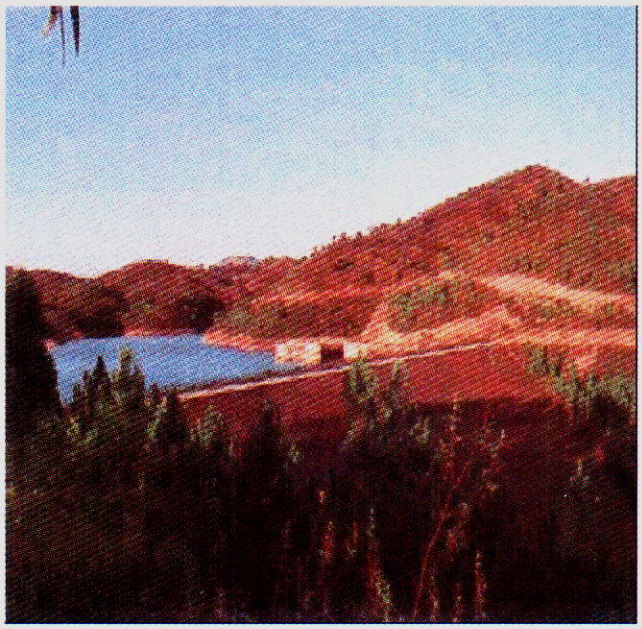
"AIMWATER"
Contract n° ENV4-CT98-0740 January 1st, 2002
Final report

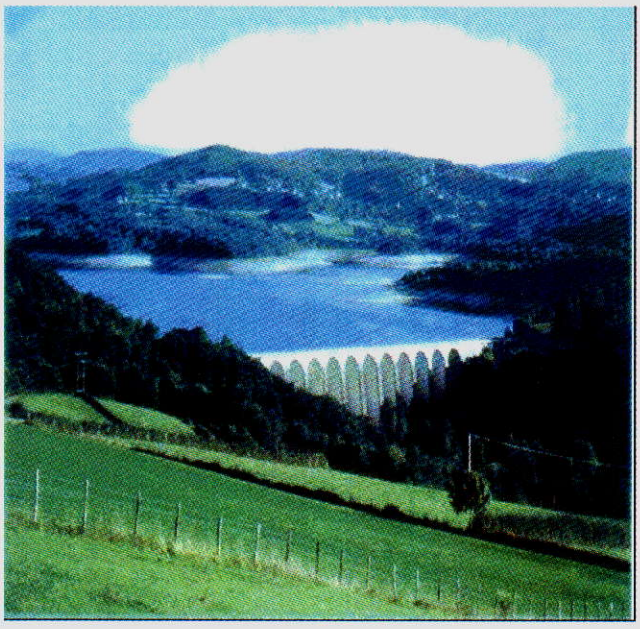
(The Seine Reservoir, IIBRBS)

(The Marne Reservoir, IIBRBS)

(The Arade Reservoir, ARSLP)



(The Yonne Reservoir, IIBRBS)



AIMWATER
Contract n° ENV4-CT98-0740(DG 12- ESCY)
FINAL REPORT

Origin: CEMAGREF/CETP/CEH/U.Valencia/U. Independente/ ARBSLP

TITLE: ***ANALYSIS, INVESTIGATION AND MONITORING OF WATER RESOURCES FOR THE MANAGEMENT OF MULTI-PURPOSE RESERVOIRS.***

**Shared cost project, part-funded under Area 3.3 (CEO Programme)
Environment and Climate Programme, European Commission.**

PARTNERS:

PA1. Cemagref	FR	ROR
PA2. CETP/CNRS	FR	ROR
PA3. Univ. of Valencia	SP	EDU
PA4. Univ. Independente	PO	EDU
PA5. CEH Wallingford	GB	EDU
PA6/CU1. Ass. Reg. Ben.	PO	IND

CUSTOMER:

CU2. IIBRBS	FR	IND

STARTING/ FINISHING DATE: 01/10/98 – 01/01/2002

DURATION: 39 months

CONTACT PERSON :

o Dr. Cécile Loumagne (co-ordinator)
CEMAGREF Parc de Tourvoie
BP 44 – 92163 Antony Cedex- France
Tel : 0033 (0) 1 40 96 61 21
Fax : 0033 (0) 1 40 96 61 99
Emailcecile.loumagne@cemagref.fr

o Dr. Sylvie Le Hégarat-Mascle
CETP/CNRS 10-12 Avenue de L'Europe
78140 Vélizy France
Tel : 0033 (0) 1 25 49 34
Fax : 0033 (0) 1 39 25 49 22
Email Sylvie.Mascle@cetp.ipsl.fr

o Dr. Ragab Ragab CEH Wallingford,
Oxon OX10 8BB United Kingdom
Tel : +0044-1491-838800
Fax : 0044-1491-692424
Email: rag@ceh.ac.uk

o Pr. Maria Simas Univ. Independente
Dep. Rec. Nat., Av. Marechal Gomes da
Costa, Lt-9, 1800 Lisboa, Portugal
Tel : +00351-1467-5247 Fax : 00351 1 859 2311
Email: mjguerreiro@mail.telepac.pt

o Pr. José Moreno Univ. of Valencia
– Facultade de Fisica
C/ Doctor Moliner nº 50 Burjassot
46100 Comunidad Valenciana España
Tel : 0034 (0) 6 39 83 112
Fax: 0034 (0) 6 36 42 345
Email: moreno@uv.es

o Mr. José Cabrita Correia
Associacao de Regantes e
Beneficiaros de Silves, Rua Manuel
Arriaga, 148300 Silves, Portugal
Tel : +351 82 44 21 45
Fax : +351 82 44 22 01
Email: jcscorreia@mail.telepac.pt

o Mme Laurence Lejeune
Institution Interdépartementale des
Barrages Réservoirs du Bassin Seine
8 rue Villiot 75012 Paris France
Tel : +33-01 44 75 29 29
Fax : +33-01 44 75 29 30
Email: iibrbs@wanadoo.fr

Origin: C. Loumagne/Cemagref
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ ARBSLP/IIBRBS/CEE

REMOTE SENSING GROUP MEMBERS:

Jose F. Moreno (Coordinator) [1]
Maria-Carmen Gonzalez [1]
Luis Alonso [1]
Carlos Cuñat [1]
Sylvie Le Hégarat-Mascle [2]
Fazia Alem [2]
Michel Normand [3]
Jean Louis Rosique [3]
José Cabrita [4]

[1] University of Valencia, Spain
[2] CNRS/CETP, France
[3] Cemagref, Anthony Research Centre, Paris, France
[4] Associaçao de Regantes , Silves, Portugal

MODELLING GROUP MEMBERS:

R. Ragab (Co-ordinator) [1]
John Bromley[1]
Ian Littlewood [1]
Martin France [2]
Cecile Loumagne [3]
Claude Michel [3]
David Aubert [3]
Anne Weisse [3]
Ludovic Oudin [3]
Charles Perrin [3]
Marine Riffard [3]

[1] Centre for Ecology & Hydrology, Wallingford, Oxfordshire, UK
[2] IT Solutions and Services, NERC , Wallingford, Oxfordshire, UK
[3] Cemagref, Research Centre, Antony, FR

CUSTOMER REQUIREMENTS GROUP MEMBERS:

H. Morel-Seytoux (Co-ordinator) [1]
M. Simas [2]
J. Cabrita [3]
C. Loumagne [4]

[1] Hydrology Days publications, Atherton, CA, US
[2] Univ. Independente de Lisboa, Sides, Lisboa, PO
[3] Associaçao de Regantes , Silves, PO
[4] Cemagref, Research Centre, Antony, FR

Contents

INTRODUCTION

The AIMWATER project, deals with important aspects of multipurpose reservoir operation under two contrasting European climates, humid temperate and semi-arid Mediterranean. AIMWATER attempts to access a dependable forecasting tool to assist reservoir managers in making decisions based on the assumption that soil moisture state is unanimously recognized as a key hydrological variable among the hydrological community.

Recent studies have shown that information derived from Earth Observation (EO) data could reliably provide spatialized and temporal soil moisture information. The integration of this information into catchment scale hydrological models could significantly improve the reservoir operation decisions. The purpose of the project was then to set up an approach to derive catchment soil moisture from Earth Observation data using microwave space-borne Synthetic Aperture Radar (SAR) images from ERS satellites and to study the improvements brought about by an assimilation of this information into hydrological models.

These methods have been put forward for use in the Seine basin upstream of Paris (France) and in the Arade basin in South Portugal where dams are operated for two different purposes, flood control and irrigation water supply. These methodologies are based on previous work done by different teams participating in the consortium and have been thoroughly described in papers presented in a special session dedicated to AIMWATER of the 5th International Symposium of Remote Sensing and Hydrology (Loumagne *et al,* 2001, Ragab *et al,* 2001, Moreno *et al*, 2001, Le Hegarat-Mascle *et al,* 2001, Morel-Seytoux *et al,* 2001a,b, Oudin *et al,* 2001, Aubert *et al,* 2001).

The first issue of this approach was to analyze customer requirements and then assess the suitability of modeling approaches to respond to their expectations. The second issue was to set up operational methodologies to retrieve soil moisture from microwave SAR data and derive hydric indicators over the Seine sub-catchments and the Arade basin. The third issue was to develop assimilation methodologies in order to integrate soil moisture information into hydrological models. The last issue was to evaluate the adequacy of these methodologies regarding user requirements and make some suggestions as how to implement these methodologies in an operational context.

The main results of these different issues will be presented in this report along with an overview of the project including some important management issues. Deliverables due at the end of this last period, which are included in this report, concern the final report on user requirements, the validation of operational methodologies for soil moisture monitoring, the comparison of assimilation methodologies, suggestions for making changes to current reservoir operation along with perspectives on the use of EO data in the reservoir community.

1. OVERVIEW OF THE PROJECT

This chapter provides an overview of management and technical issues addressed in the project. The different tasks and the work done are displayed on a Gantt Chart (figure 1).

1.1 Management issues WP 1000, WP 6000, WP 7000

• Quality Plan

A quality plan with guidelines for quality control was set up and agreed on by all partners. The partners were required to abide by this quality plan; this ensured the quality and the reliability of the data collected, measurements conducted and results obtained during the lifetime of the project.
A frame for the procedures was proposed to the consortium.
Procedures describing the EO and Hydrometeorological database, assimilation process, model suitability and derivation of EO indicators have been written and procedures on the comparison of assimilation methodologies, adequacy of model results and validation of EO indicators have been completed. Each partner was expected to write the procedures, which are their responsibility, and to provide Brussels with these procedures, if they are required.

• Promotion of the project

- *Web site:*

A Unix station dedicated to the AIMWATER project puts up the web site (http://dataserv.cetp.ipsl.fr/AIMWATER/) to the awareness and the promotion of the project. A link to the INFEO web site pages describing the project is available (http://www.infeo.org/). On the web site dedicated to the project the visitor can find the main objectives and results of the project including the project presentation, the list of participants, the schedule, the studied watersheds, hydrological models and database, remote sensing data, reports and publications.

- *Publications and conferences:*

Publications and conferences related to the AIMWATER project and presented during the lifetime of the project are displayed below:

1999

Moreno J., 1999. Presentation of the AIMWATER project, 12-13/03/99, Envisys Workshop, Spain.
Weisse A., Normand M., 1999. Presentation of the AIMWATER Project, 29 /03/99, French Research Ministry Meeting, R.S./ Hydrology, France.
Quesney A., Weisse A., Le Hégarat-Mascle S., Normand M., Loumagne C., 1999. Assimilation of Soil Moisture Index into an Hydrological Model, 22-23/04/99, Annales Geophysicae, Vol 1, n°2, pp 351, EGS, Netherlands
Montfort M., Weisse A., Loumagne C., Normand M., Quesney A., Le Heragat-Mascle S., Alem F., 1999. Integration of remote sensing data into hydrological models for reservoir management purposes. 16-20/09/ 99, 19th Hydrology days, AGU, Colorado.

2000

Le Héragat-Mascle S., Taconet O., Quesney A., Vidal-Madjar D., Loumagne C., Normand M., 2000. Land Cover discrimination from multitemporal ERS images and multispectral Landsat images: a study case in an agricultural area in France. Int. Journ., of Remote sensing, vol 21 n°3, pp 435-456
Le Hégarat-Mascle S., Alem F., Quesney A., Normand M., Loumagne C., 2000. Estimation of watershed soil moisture index from ERS/SAR data. EUSAR 2000, Munich, Germany, May 23-25 2000., pp.679-682

Quesney A., François Ch., Ottlé C., Le Hégarat-Mascle S., Loumagne C., Normand M., 2000. Sequential Assimilation of SAR/ERS data in a surface hydric model coupled to a global hydrological model with an extended Kalman filter. 3-7/04/ 00, RS & Hydrology Symposium, IAHS/ICRS, Sta Fé, NM, USA.

Loumagne C., Weisse A., Normand M., Riffard M., Quesney A., Le Heragat-Mascle S., Alem F, 2000, Integration of remote sensing data into hydrological models for flood forecasting. 3-7/04/ 00, RS & Hydrology Symposium, IAHS/ICRS, Sta Fé, NM, USA.

Quesney A., Le Héragat-Mascle S., Taconet O., Vidal-Madjar D., Wigneron JP, Loumagne C., Normand M., 2000, Estimation of Watershed Soil Moisture index from ERS/SAR data. Remote Sens. Environ. 72:290-303

Riffard M., Loumagne C., Weisse A. , Normand M., Quesney A., Le Hégarat –Mascle S., Alem F., 2000. Etat hydrique des bassins versants, observation spatiale et prévision des débits: AIMWATER un projet européen sur le bassin de la Seine. Milieux poreux et transferts hydriques. Bulletin du GFHN, N°45, (in press).

Normand M., Loumagne C., Ottlé C., Le Heragat-Mascle S., Alem F., Quesney A., 2000. Etat hydrique des sols et hydrologie: Approche par télédétection pour la prévision des débits. Colloque PNRH Mai 2000, Toulouse France. Actes 5pg.

Loumagne C., 2000. Démarche qualité pour un projet de recherche. Le projet Aimwater. Séminaire pour la qualité en recherche Cemagref, INRA CEA AFSA. Mai 2000.

Riffard M., Littelwood I., Loumagne C., Weisse A., Perrin Ch., 2000. Continuous daily rainfall-catchment wetness-streamflow simulation models for subcatchments of the Seine basin France. ERB 2000 Monitoring and modeling catchment water quantity and quality, Sept, Ghent Belgium, Actes 8pg.

Weisse A., Loumagne C., Normand M., Aubert D., Le Hégarat-Mascle S., Alem F. 2000. Assimilation variationnelle de mesures d'humidité de surface dans un but de prévision de crue. Ateliers modélisation CNRM, Toulouse, novembre 2000, pp. 31-34.

Aubert D., Loumagne C., Weisse A., Michel C., 2000. Application d'un filtre de Kalman sur un modèle hydrologique pluie-débit. Assimilation de données d'humidité du sol et de débits. Ateliers modélisation CNRM, Toulouse, novembre 2000, 115-118.

2001

Le Hégarat-Mascle S., Alem F., Quesney A., Normand M., Loumagne C., 2001. Surface soil monitoring from ERS/SAR data. Method and Validation over three different watersheds. ISPRS, 8th International Symposium, 8-12 Jan 2001, Aussois, pp. 581-587

Quesney A., François Ch., Ottlé C., Le Heragat-Mascle S., Loumagne C., Weisse A., Aubert D., 2001. Sequential assimilation of SAR/ERS data in a surface hydric model coupled to a lumped rainfall-runoff model with an extended Kalman Filter. ISPRS, 8th International Symposium, 8-12 Jan 2001, Aussois, .pp. 689-693.

Loumagne C., Normand M., Riffard M., Weisse A., Quesney A., Le Hegarat-Mascle S., Alem F.,2001. Methodology for integration of remote sensing data into Hydrological models for reservoir management purposes. Hydrological Science Journal 46(1), 89-102

Weisse A., Michel C., Aubert D., Loumagne C., 2001. Variational assimilation of a watershed soil moisture index in a hydrological model for flood forecasting. SVAT transfer schemes and large scale hydrological models, IASH Red Book Publication.n°270, pp. 249-256

Weisse A., Le Hegarat-Mascle S., Aubert D., Loumagne C., 2001. Le projet Européen AIMWATER: Utilisation de l'humidité des sols mesurée par radar embarqué (ERS/SAR) pour la modélisation pluie-débit, Colloque SHF, 20-21 juin 2001, Toulouse, La Houille Blanche, (accepted).

Weisse A., Oudin L., Loumagne C., 2001. Assimilation des données d'humidité des sols pour la prévision de crues : Comparaison d'un modèle Pluie-débit conceptuel et un modèle intégrant une interface sol-végétation-atmosphère. Revue des Sciences de l'Eau. (submitted).

Le Hégarat-Mascle S., Zribi M., Alem F., Weisse A.., Loumagne C., 2001. Soil moisture estimation from ERS/SAR data : Toward an operational methodology. IEE (submitted).

Aubert D., Loumagne C., Weisse A., Le Hégarat-Mascle S., 2001, Assimilation of Earth Observation data into hydrological models : the sequential method. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 8pg.

Oudin L., Weisse A., Loumagne C., Le Hégarat-Mascle S., 2001. Assimilation of soil moisture into hydrological models for flood forecasting: a variational approach. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 8pg.

Ragab R., Perrin C., Littlewood I, Bromley J., France M., 2001. Prediction of runoff to surface reservoirs using the remotely sensed catchment wetness index. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 8pg.

Le Hégarat-Mascle S., Poirier- Quinot M, Alem F, Weisse A., Loumagne C., 2001. Validation of a methodology to Monitor Soil Moisture fom C-band SAR Spaceborne in an Operational Way. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 9pg.

Morel-Seytoux H., Loumagne C., Simas M., 2001a. Appropriate forecast Models and reservoir operations algorithms: theory. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 7pg.

Morel-Seytoux H., Loumagne C., Simas M., 2001b. Practical changes to current reservoir operations procedures : suggestions. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 6pg.

Moreno J., Cuñat C., Alonso L., Gonzalez MC, Garcia JC., 2001. Operational methodologies to retrieve surface parameters from EO data: Synergistic use of optical and SAR data. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 8pg.

Loumagne C., Aubert D.,Weisse A., Ragab R., Moreno J., Le Hégarat-Mascle S., Morel-Seytoux H., M. Simas, Cabrita J, 2001. AIMWATER, Analysis, Investigation and Monitoring of Water Resources for the Management of Multipurpose Reservoirs: Overview of the Project. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, 10pg.

• Reports and Meetings

During the lifetime of the project 5 progress reports, 1 Final report and financial reports were sent to Brussels. Also, 1 Workshop and 6 meetings were held and organized by the different partners of the consortium. Progress reports and meeting reports can be find on the FTP site dedicated to AIMWATER:

- Meetings

Kick off meeting	19-20 November, 1998, at CEMAGREF, Antony, France
2nd meeting	29-30 April, 1999, at Associaçao de Regantes S, L, P, Silves, Portugal
3rd meeting	04-05 November, 1999, at Universitat de Valencia, Spain
4th meeting	11-12 Mai, 2000, at CETP, Velizy, France
5th meeting	13-14 November, 2000, at Universisdade Independente, Lisboa, Portugal
6th meeting	08-09 Mai, 2001, at CEH, Wallingford, UK

- Final Workshop 02-05 October 2001, in Montpellier, France

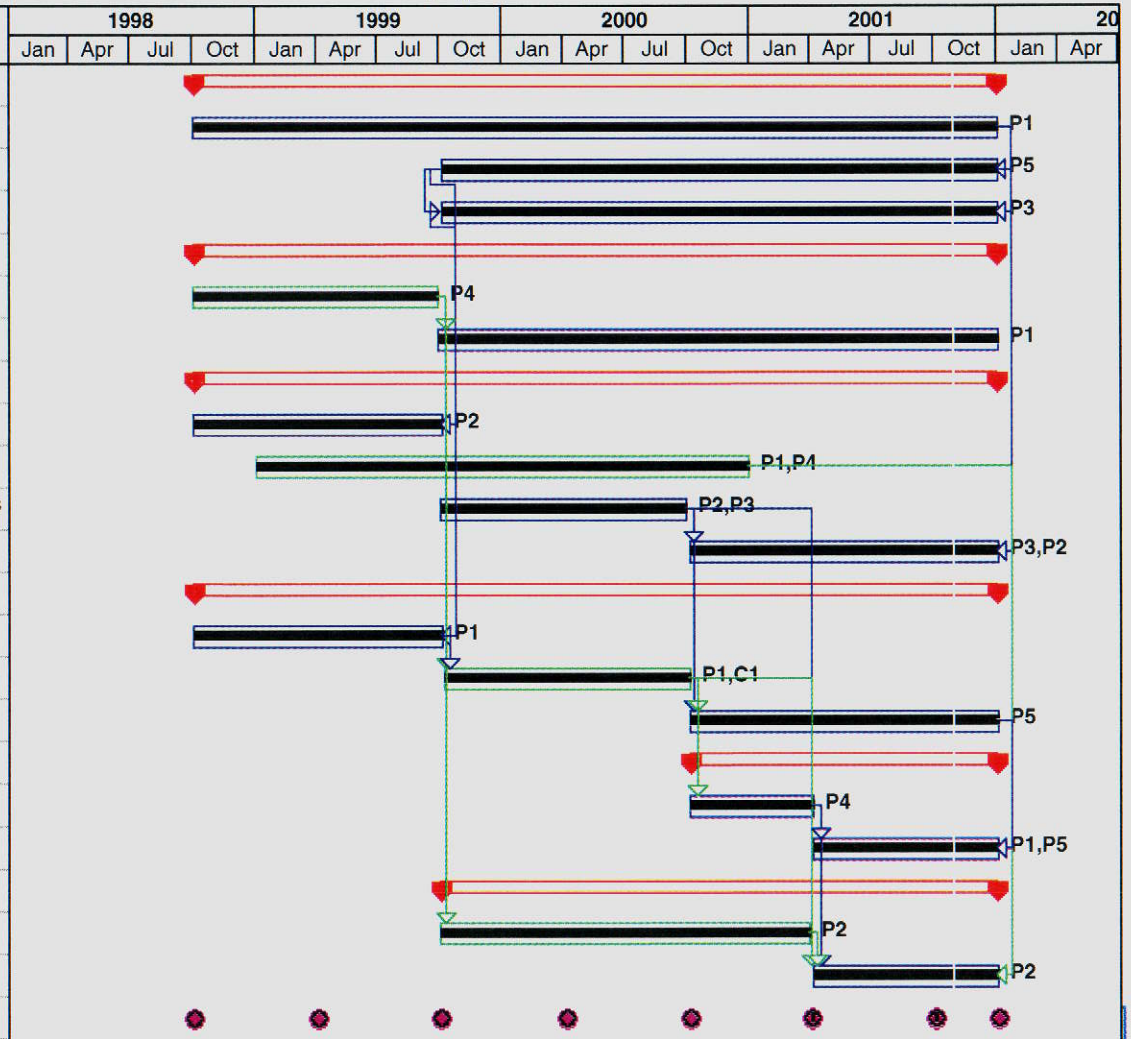
An International Symposium of Remote Sensing in Hydrology was held in Montpellier in October 2-5, 2001. A special session was dedicated to the AIMWATER project including 8 presentations focussing on the main issues of the project (see publications & conferences 2001)

- Reports

1st progress report	Project plan, quality assurance plan, customer requirement template
2nd progress report	State of the art on: EO signal analysis, Catchment models and Customer requirements.
3rd progress report	Selection of the best suited models, EO signal analysis methods
4th progress report	Assimilation methodologies, Derivation of EO hydric indicators
5th progress report	AIMWATER Data Base, AIMWATER Web site, Synthesis of model results regarding user's needs
Final report	Analysis of user's requirements, Validation of operational methodologies for soil moisture monitoring, Comparison of assimilation methodologies, Changes to current reservoir operations and Perspectives on EO data use in reservoir community.

Figure 1: Gantt Chart for the AIMWATER project

ID	Task Name
1	**1000 Management**
2	1100 overall project management
3	1200 Coordination of hydrological models
4	1300 Coordination of EO signal analysis
5	**2000 Customer requirements analysis**
6	2100 State of the art
7	2200 Evolving customer requirement
8	**3000 Derivation of hydric indicators from EO data**
9	3100 State of the art
10	3200 Elaboration of the data base
11	3300 Selection of test catchment and derivation of EO indicators
12	3400 Validation of hydric indicators
13	**4000 Assimilation of EO data into a hydrological model**
14	4100 State of the art
15	4200 Selection of the catchment model for EO data assimilation
16	4300 Comparison of models
17	**5000 Technical feasibility of model integration**
18	5100 Synthesis on model results and requirements
19	5200 Adequacy of the model and implementation
20	**6000 Promotion of results**
21	6100 Set up a server
22	6200 Integration of data into the server
23	**7000 Reports and meetings**
32	



1.2 Technical issues WP 2000, WP 3000, WP 4000, WP 5000

The presentation of these technical issues was provided in different papers presented at the Final Workshop in the 5th International Seminar on Application of Remote Sensing in Hydrology, Montpellier, 2-5/10/01 (see publications 2001). A brief introduction is provided there after:

- **User Needs and Model Suitability**

The first customer is a Portuguese agency, which manages 2 multipurpose reservoirs in South Portugal. Reservoirs must deal with 3 major purposes: irrigation in summer, water supplies for the downstream urban area and they must face sudden floods in mid spring.
The second customer is a French agency managing four large multipurpose reservoirs in the Seine basin. These reservoirs are located in the upper part of the catchment and are responsible for flood protection and water supply of Paris. The tools used by customers to manage their reservoirs are very different:
In the case of the Arade, no model is being used at the present time. Managers need to abide by administrative objectives defining the ideal scheduling of filling and drainage of the reservoir. So in real time, operating their reservoir relies on common sense. Some empirical linear relationships have been set up based on regression on previous streamflow rates and concomitant rainfall values. But all the operation is made based on previous experience with a set of rules predefined in different situations.
In the case of the Seine, customers have also to abide by an administrative rule curve, which translates the dual-purpose policy of low flow augmentation and flood control. To simulate the river system they have a conceptual model to transform rainfall into runoff in upper basins and intermediate tributary basins and they have a propagation function to route the flow throughout the river system. Another model uses the information at the target points in the system to decide on the releases to be made. This model simulates the operation of the reservoir and optimises the releases to achieve the administrative objectives. Therefore, since they don't have the same tools their expectations are also quite different:
The Arade managers expect either a set of improved rules and procedures to perform predefined situations or a very simple model to simulate their river system. In fact any additional information would be an improvement to their management operating system.
The Seine managers would like to reduce the uncertainties remaining in the information available to improve their forecasting capabilities of their river flow simulation rainfall runoff model. They also need to optimise their reservoir releases by developing an integrated model with a stream flow simulation module simultaneously with a reservoir release calculation.
Because the project needs to provide customers with operational tools for water management and reservoir operation, modelling approaches have been selected considering several criteria based on customer requirements but also on models ability to simulate stream flow with a good reliability. The project had also to provide models that are in continuity with the existing customers' modelling tools. So the proposed methodology should appeal to models which are easy to use or implement and don't require much more expertise than they have. The model should be run in operational conditions not too demanding in computer facilities or model input and then should be able to use remotely sensed data.
A selection of the best-suited models was made using quantitative assessment by comparing 38 existing model structures (Perrin *et al*, 2001). Models were tested on a sample of 429 catchments worldwide. This sample includes 1294 calibration periods of three to six calendar years and 3204 validation tests. A first selection showed that conceptual models seem the most suitable among the available modelling tools in the context of this project. Considering the results, several model structures were recommended in the AIMWATER project because of their consistent performance and reliability: GR models: GR4 model (Edijatno *et al.*, 1999) and GRHUM model (Loumagne *et al.*, 1996), IHACRES model (Littlewood *et al.*, 1997) and TOPMODEL (Beven, 1997).
To meet customer requirements this project had to validate the methodology set up to retrieve new and reliable information derived from EO data that can be assimilated in hydrological models. Then the proposed approach had to be assessed in a thorough inter comparison of classical catchment models versus models assimilating EO data.

- **Soil moisture retrieval from EO data and derivation of hydric indicators**

The leading role played by soil moisture as well as in the atmospheric part of the water cycle as in the continental one and the availability of space borne imaging radar (such as the ESA's ERS1-2 and the Canadian RADARSAT), led scientists to analyze the potential use of radar information for the assessment of the surface soil moisture content (Ulaby & El-Rayes, 1987, Evans *et al.*, 1997). In the past 20 years, several ways to derive soil moisture indices from airborne or space borne active measurements have been proposed (Dubois *et al.*, 1995; Wang *et al.*, 1997) and optimized features have been highlighted. But operational radars available now have not such optimized features and new methodologies have been proposed for soil moisture monitoring.
The first proposed methodology is based on the selection of targets where the measured SAR signal depends on the soil water content. The methodology is developed from previous studies within the consortium (Cognard *et al.*, 1995, Taconet *et al.*, 1996, Quesney *et al*, 2000, Le Hégarat-Mascle *et al*, 2000) using airborne and space borne radar campaigns together with terrain data, from the field to the catchment scale. These studies have produced a pre-operational methodology and have shown that a soil moisture indicator can be derived from microwave space borne SAR and reliably provide surface soil moisture information at a watershed scale. Besides soil moisture, the radar signal backscattered from vegetated areas depends also on soil roughness and vegetation effects. Thus in order to cope with these effects this methodology carries out successive steps in signal treatment and catchment analysis over the Seine sub-catchments thoroughly described in Quesney *et al.*, 2000. The methodology was first applied to a homogeneous cover (wheat crop). It was then generalized to multiple cover types, in order to extend its applicability in both time and space.
The second proposed methodology is based on the complementary use of optical and microwave SAR data to derive hydric indicators for the Arade basin in Portugal. SPOT and LANDSAT optical images are used to derive a land surface classification. SAR images acquired by ERS satellites are used to detect changes in surface roughness and soil dielectrical properties related to soil moisture changes. A temporal series of SAR images are used to derive soil moisture maps over the study area in order to obtain soil moisture evolution, while the land classification map is used to account for specific modeling and corrections associated to each land surface type. Besides geometric correction and calibration, SAR images require compensation for vegetation effects. The SAR images with the vegetation effect compensated are used to obtain soil moisture maps. They are calculated by calibrating the SAR image pixel values with soil moisture measurements in the study area.

- **Assimilation methodologies for hydrological models**

The difficulty of predicting floods in a reliable way originates partly from a certain lack of accuracy of hydrological models, particularly during unusual hydrologic events. Because rainfall-runoff models are far from being perfect, hydrologists need to put the model in better compliance with the current observation prior to using it in forecasting mode. Methods have thus been developed to improve hydrologic forecasting. The fundamental idea is that if the model predictions diverge from the observations at a given time, there is little chance that future estimations will approach the correct values. The improvement then, comes from a correction of the trajectory of the model based on observations preceding the forecast. This operation has been termed updating in hydrology and assimilation in meteorology.
O'Connell & Clarke (1981) and, more recently, Refsgaard (1997) reported on four different methodologies used for model updating. These methodologies depend on what is considered to be the main cause of discrepancy between observed and computed stream flow values. Therefore, when using conceptual rainfall-runoff models, which are non linear, the most accurate methodologies are the state and parameter updating. In the AIMWATER project three methodologies have been set up:
The first methodology is derived from previous work carried on by Yang and Michel (2000) in the field of flood forecasting. Part of this research has been used as a basis of the parameter updating assimilation procedure developed within the consortium. The model's inability to produce correct stream flow values generally translates into parameter uncertainty. Parameter calibration is the means by which a model structure adjusts to a given set of data. Therefore, parameter updating seems to be a natural way to amend the current error in stream flow value. Consequently, a specific methodology of parameter updating was chosen as the starting point from which soil moisture assimilation could best be coped with. This methodology consists of adjusting the parameter of the model over a certain number of days preceding the date when a forecast is desired, so that over that period the calculated values (either discharge or discharge and soil moisture) fit better with the observed ones.

The second methodology is based on sequential assimilation algorithm, with state updating (Quesney *et al,* 2001, François *et al,* 2001). The aim of the sequential assimilation is to improve the performances of a hydrological model by controlling its evolution and by limiting the divergence between the model and available observations. The most widespread method of sequential correction is that proposed by Kalman (1960). For about ten years, thanks to numerical calculation advances, the Kalman filter has been subject of applications in environmental sciences. The Kalman filter consists of calculating this correction term taking into account the estimated errors on model and observations, and locally linearising the model: the correction will be done by using linear equations, and the more the observations are accurate, the more the *a posteriori* internal state (after correction) are close to them.
The third methodology describes the estimation of a catchment wetness index from measured or simulated distributed soil moisture data along with the assimilation of remote sensing data into hydrological models. In the implementation set up within the project the estimation of catchment wetness index is calculated from model simulations and is used as an input for the hydrological model in order to substitute soil moisture accounting procedures of the models in forcing mode. The catchment wetness index is estimated by averaging the wetness indices at a catchment scale in a hydrological distributed model. The model can be calibrated against the wetness index obtained from observations such as remote sensing data or TDR data. Then the continuous time series of soil moisture data calculated by the model can be used as input in rainfall-runoff models. The efficiency of the different assimilation procedures set up for the project are discussed in this report focusing on the contribution of soil moisture data.

- **Methodology, adequacy and suggestions for implementation in an operational context**

In order to satisfy user needs the AIMWATER project had to address two kinds of issues:
A better knowledge of the state of the system in order to improve forecast of runoff and consequently improve operation of the reservoirs, and a better choice of releases from reservoirs, especially for flood protection, requires not only a better knowledge of the internal state of the system but also that reservoir operation rules specifically take the forecast of runoff into consideration.
The translation into needed tool development was focused on two areas, the first one was to search for relatively simple and robust models of the rainfall-runoff process that can assimilate soil moisture information into their algorithmic structures and the second one was to retrieve from an EO signal a useful measure of watershed wetness that can be incorporated in a rainfall-runoff model in forecasting mode.
In the project, the adequacy of these methodologies are analysed from a users perspective considering not only the improvements but also the gaps or difficulties of possible implementation of the proposed procedures. The improvement in forecasting mode has been assessed using a deterministic future i.e. a future when rainfall values are known with certainty.
Some suggestions are provided on how these methodologies can be implemented in reservoir operation. It is quite difficult to determine exactly how reservoirs operate on a day-to-day basis because it is almost always done on the basis of the long practical experience of the operators, naturally within a set of regulations, which do leave some room for flexibility. Nevertheless models are used more and more for decision support, and for simulations of various scenarios. From a study of the impact of these various scenarios on the behavior of the system, the operator then decides on a particular release to be made.
Currently, the management procedure includes neither past releases nor future ones; experience in its use has shown that this procedure leads to oscillations in the pattern of the releases. The procedure then can be modified by an updating methodology to avoid these oscillations and by placing limitations on the range of changes in the values of the releases from one day to the next. Since theory suggests that past and future releases do influence the release to be made today, it can be sound to modify the current simulation procedure for operations with the most realistic theoretical approach developed in the context of a deterministic future and then extended to the case of an uncertain one. Unfortunately there is not a single unique objective way to include the uncertainty in the optimisation formulation for the problem. The formulation is subjective depending on the attitude of the manager with respect to risk. Analyses of the different attitudes of the manager with respect to risk leading to different scenarios and rules of operation are provided in the report.

The main results of the issues presented above are displayed in the following chapters, part of them have been presented either in previous progress reports or in the final Workshop dedicated to the project.

References

Beven K.J. (1997). TOPMODEL: a critique. *Hydrological Processes*, **11(9)**, 1069-1085.

Cognard A.L., Loumagne C., Normand M., Olivier Ph., Ottlé C., Vidal-Madjar D., Louahala S., Vidal A. (1995). Evaluation of the ERS 1/synthetic aperture radar capacity to estimate surface soil moisture: two-year results over the Naizin watershed. *Water Resource Research*, **vol.31, n°4**, 975-982.

Dubois P.C., Van Zyl J., Engman T. (1995). Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote sensing*, **vol 33, n°4**, 877-895.

Edijatno, Nascimento N.O., Yang X., Makhlouf Z. and Michel C. (1999). GR3J: a daily watershed model with three free parameters. *Hydrol. Sci. J.*, **44(2)**, 263-278.

Engman E.T. (1990). Use of microwave remotely sensed soil moisture in hydrologic modeling. In: *Application of Remote Sensing in Hydrology*, ed. G.W. Kite & A. Wankiewicz, 279-292. Proc. Symp. No.5, NHRI, Saskatoon, Canada.

Evans D.L., Jeffrey J.P., Stofan E.R. (1997). Overview of the spaceborne Imaging radar C/X band synthetic aperture radar (SIRC/XSAR) missions. *Remote Sens. Environ.*, **vol 59, n°2**, 135-140.

François, C., Quesney, A. and Ottlé, C., (2001b). Sequential assimilation of SAR/ERS1 data into a coupled land surface-hydrological model using an Extended Kalman Filter. *Journal of Hydrometeorology, submitted.*

Kalman R.E. (1960). A new approach to linear filtering and prediction problems. *J.Basic Eng.*, **vol 82 D**, 35-45

Le Hégarat-Mascle, S., Alem, F., Quesney, A., Normand, M., Loumagne, C., (2000). Estimation of watershed soil moisture index from ERS/SAR data, *Proceedings of EUSAR2000, in Munich, Germany, on May 23-25, 2000,* pp.679-682.

Littlewood I.G., Down K., Parker J.R. and Post D.A. (1997). The PC version of IHACRES for catchment-scale rainfall-streamflow modeling. Version 1.0. User Guide. *Institute of Hydrology,* 89p.

Loumagne C., Michel C., Normand M. (1991). Etat hydrique du sol et prévision des débits. *Journal of Hydrology*, **vol. 123, n°1-2**, 1-17.

Loumagne C., Chkir N., Normand M., Ottlé C. and Vidal-Madjar D. (1996). Introduction of soil/vegetation/atmosphere continuum in a conceptual rainfall-runoff model. *Hydrological Sciences Journal*, **41(6)**, 889-902.

O'Connell P.E., Clarke RT. (1981). Adaptive hydrological forecasting a review. *Hydrol.Sci. Bull,* **26(2)**, 179-205.

Perrin C., Michel C., Andréassian V., 2001 : Does a large number of parameters enhance model performance ? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology,* 242, 275-301

Pietroniro A., Soulis E.D., Kouven N., Rotunno O., Mullins D.W. (1993). Using wide swath C-band SAR imagery for basin soil moisture mapping. *Can. J. Remote Sens.*, Special Issue, January, 77-82.

Quesney A., Le Hégarat-Mascle S., Taconet O., Vidal-Madjar D., Wigneron J.P., Normand M., Loumagne C. (2000). Estimation of watershed soil moisture Index from ERS/SAR data. *Remote Sensing of Environment,* **vol 72**, 290-303.

Quesney, A., François, C., Ottlé, C., LeHégarat-Mascle, S., C. Loumagne and Normand, M., (2001). Sequential assimilation of SAR/ERS data in a surface hydric model coupled to a global hydrological model with an Extended Kalman Filter,, Actes du Huitième *Symposium International "Mesures Physiques et Signatures en Télédétection", Aussois, France*, pp. 689-693.

Refsgaard J.C. (1997). Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrol.*, **28**, 65-84.

Taconet O., Vidal-Madjar D., Emblanch C. and Normand M. (1996). Taking into account vegetation effects to estimate soil moisture from C-band radar measurements. *Remote Sensing of Environment*, **vol.56**, 52-56.

Ulaby F.T. and El-Rayes M.A. (1987). Microwave Dielectric Spectrum of Vegetation. Part II: Dual-Dispersion Model. *IEEE Transaction on Geosci. and Remote Sens.*, **vol.25, n°5**, 550-557.

Wang J., Hsu A., Shi JC., O'Neil P., Engman T. (1997). Estimating surface soil moisture from SIR-C measurements over the little Washita River watershed. *Remote Sens. Environ.*, **vol 59, n°2**, 308-320.

Yang X., Michel C. (2000). Flood forecasting with a watershed model: a new method of parameter updating, Hydrol. Sci. J. **45(4)**, 537-546.

2. CONTEXT OF THE PROJECT

This chapter provides a description of the AIMWATER context including the studied area and the database set up for the project. Within this context, user's needs have been analysed along with the suitability of the approaches developed in the project to respond to their expectations.

2.1 The AIMWATER database

The customers involved in the project deal with the management of multi-purpose reservoirs under two contrasting European climates: humid temperate (the Seine basin, France) and semi-arid Mediterranean (the Arade basin, Portugal).
In order to set up the methodologies developed in the project different catchments were selected across the study area and a database was established:

During the AIMWATER project a large amount of existing data has been collected. Some are time-varying data, which include stream flow, rainfall, evaporation and meteorological information; the remainder are spatial data sets and include satellite images, soil and land use maps.
Most time-varying information is drawn from existing sources but some soil moisture field measurements using a TDR system were taken in the study areas.

The data have been collected for the following catchments:
France: *Grand Morin; Petit Morin; Orgeval; Saulx and Serein*
Portugal: *Arade*

The time varying information has been stored within an Access database; all spatial data are stored as various types of image in folders on the CEMAGREF FTP site in Antony, France and on the WEB site specially dedicated to AIMWATER (http://dataserv.cetp.ipsl.fr/AIMWATER/).
A description of the soil moisture field measurements that were collected as part of the project is also presented.

- **Studied Catchments**

- *The French catchments*

Different sub-catchments within the Seine basin have been selected across the modeling domain:
- the Serein catchment (1120 km² at Chablis) ;
- the Saulx catchment (2100 km² at Vitry-en-Perthois) ;
- and the Grand Morin catchment (1190 km² at Montry ; 1070km² at Crécy La Chapelle).

Two further study areas have been added since the project started; these are:
- the Orgeval catchment (104 km²), a sub-catchment of the Grand Morin, and
- the Petit Morin (605 km²) a catchment that lies along the northern boundary of the Grand Morin.

For all sub-catchments, hydrological data at a daily time step have already been collected: rainfall data over the catchment, potential evapotranspiration data at representative stations and runoff data at the catchment outlets. Examples of the distribution of sites are given in Figures 2, 3 and 4 (Grand Morin, Serein and Saulx).
This data have been used for calibrating and testing the hydrological models.
As mentioned above additional data such as catchment characteristics, land use and soil maps are also available from various sources for input requirements.

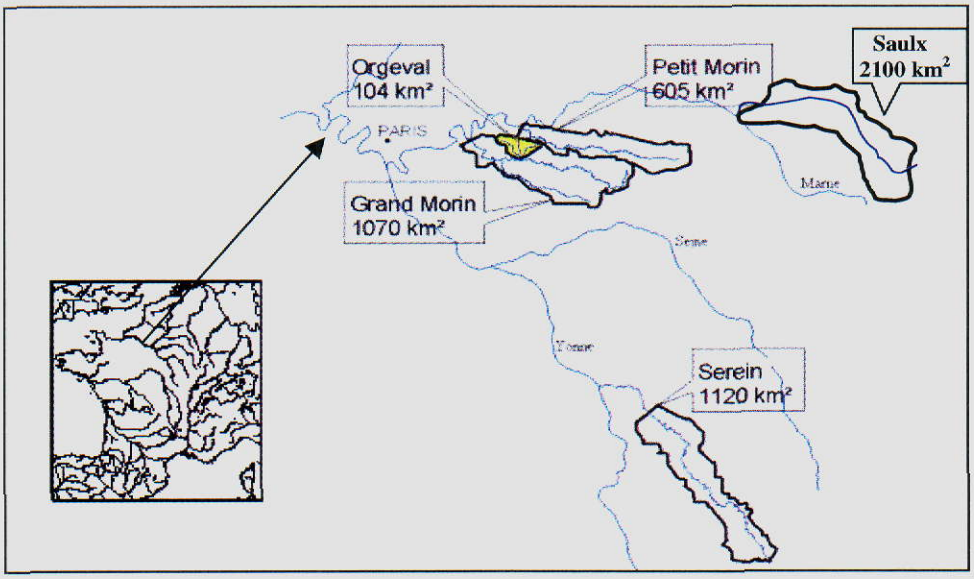


Figure 1: Location of the Seine basin (43800 km²) and the studied sub-catchments

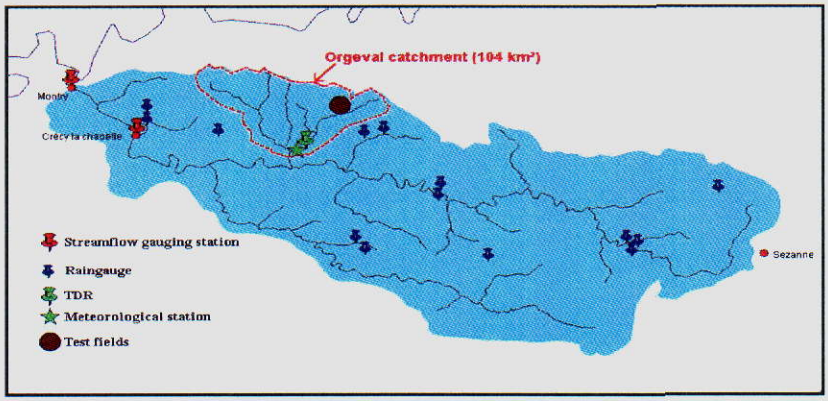
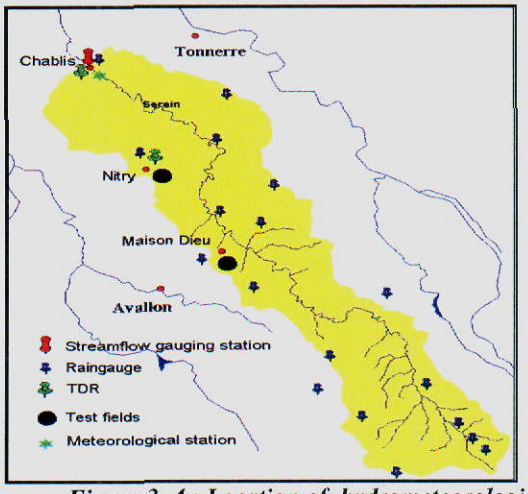


Figure 2 : Location of hydrometeorological statins in the Grand Morin catchment



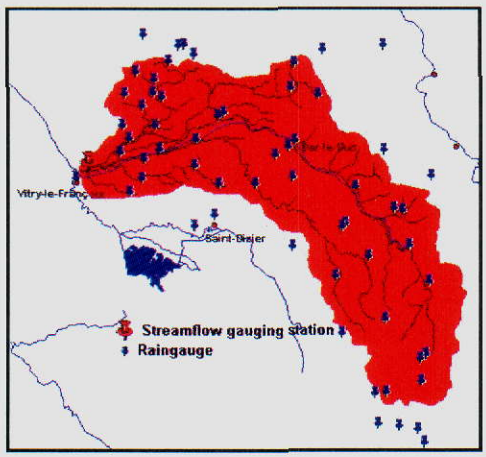


Figure 3, 4 : Location of hydrometeorological stations in the Serein and Saulx catchments

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

- *The Portuguese (Arade) catchment*

The Arade bassin covers an area of 975 km^2 and lies close to the southern coast of Portugal (Figure 5). The studied area is the sub-catchment of Funcho dam, which covers 257.5 km^2.

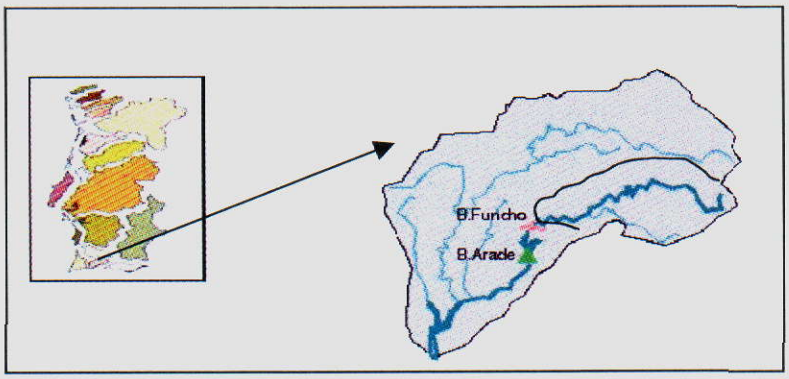


Figure 5: Location of the Arade basin and Funcho catchment

Unlike the French catchments the spatial coverage of rainfall data is limited; only one station, that at Funcho dam, possesses a sufficiently long daily record to be of value to the modelling simulations. This means it has not been possible to spatially distribute rainfall in the models; instead a uniform distribution has been assumed. Streamflow and meteorological data from the Funcho Dam station have also been used for the modelling exercises.

Spatial data relating to topography, land use and soil types are available in Arc View format (Figure 6). These maps were obtained from INAG (web site http://snirh.inag.pt/snirh/). LANDSAT, ERS and SPOT images and have been used to classify and provide a spatial distribution of vegetation types in the catchment to supplement and update the information already available in map form.

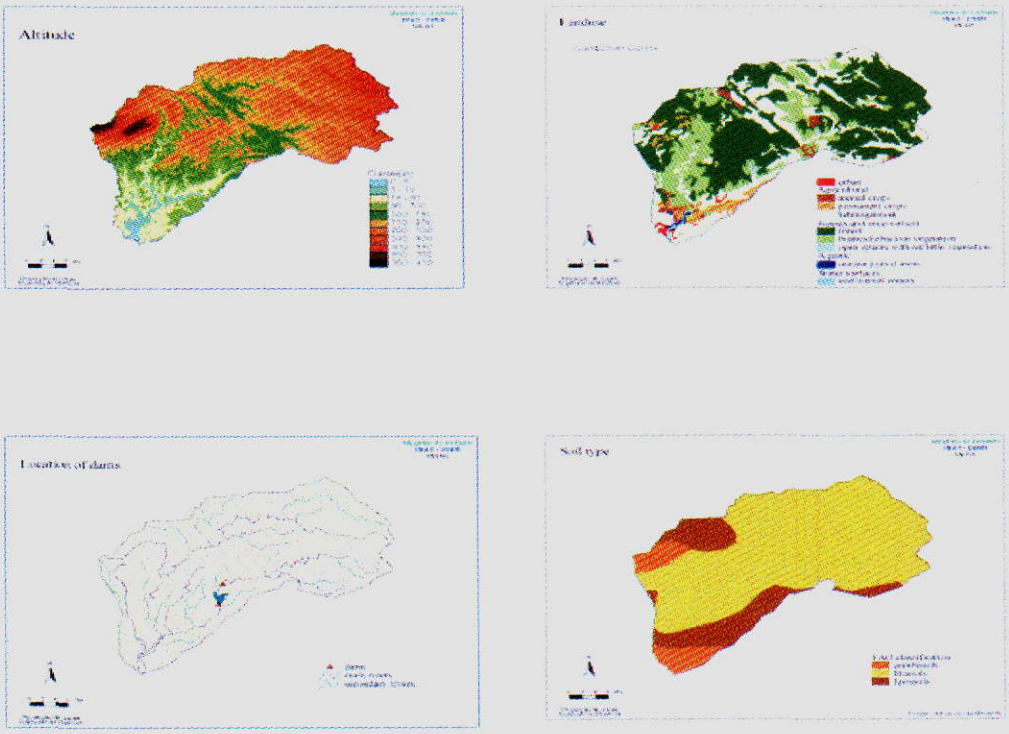


Figure 6: Examples of Arc View maps from INAG

- **The hydrometeorological data: The Access database**

All time-varying data has been entered into a Microsoft Access database. Access is a 32-bit PC-based relational database management system that runs under Windows and forms part of the widely used Microsoft Office package; it is thus widely available.
The strength of the system is that different types of data can be entered as separate tables but these can be linked through common fields in each table. The common fields can be anything such as gauge names, codes or even dates (Table 1).
The system for retrieving data is to set up a 'query' using simple commands to specify the data that is required. The data may exist in more than one table but because they are linked information can be drawn simultaneously from any number of tables. For example, a query may request all rainfall data between 1959 and 1984 for all rain gauges above an altitude of 300m within a specified area defined by grid references. This information would be obtained by combining information from tables containing the rainfall data and those having information relating to the location and elevation of the gauges. Data can easily be imported as spreadsheets or text files and can be exported in a wide range of formats including Excel, Dbase, Html and text formats.
A comprehensive description of the system is given in Jennings (1995).
A separate database has been set up for each of the two study sites within Aimwater (France and Portugal).

Table 1: Section of an Access database table

Catchment	GaugeName	GaugeCode	Date	FlowMMday	Day	Month	Year
Serein	Serein à Chablis	H2342010	21/09/55	0.0071	21	9	1955
Serein	Serein à Chablis	H2342010	22/09/55	0.0071	22	9	1955
Serein	Serein à Chablis	H2342010	23/09/55	0.0071	23	9	1955
Serein	Serein à Chablis	H2342010	24/09/55	0.0071	24	9	1955
Serein	Serein à Chablis	H2342010	25/09/55	0.0071	25	9	1955
Serein	Serein à Chablis	H2342010	26/09/55	0.00964	26	9	1955
Serein	Serein à Chablis	H2342010	27/09/55	0.00964	27	9	1955
Serein	Serein à Chablis	H2342010	28/09/55	0.0071	28	9	1955
Serein	Serein à Chablis	H2342010	29/09/55	0.0071	29	9	1955
Serein	Serein à Chablis	H2342010	30/09/55	0.0071	30	9	1955
Serein	Serein à Chablis	H2342010	01/10/55	0.0071	1	10	1955
Serein	Serein à Chablis	H2342010	02/10/55	0.0071	2	10	1955
Serein	Serein à Chablis	H2342010	03/10/55	0.0071	3	10	1955
Serein	Serein à Chablis	H2342010	04/10/55	0.0071	4	10	1955
Serein	Serein à Chablis	H2342010	05/10/55	0.0071	5	10	1955
Serein	Serein à Chablis	H2342010	06/10/55	0.01273	6	10	1955

- Time-varying data: French Catchments

The time varying data currently stored on the Access database for the French catchments are summarised in Figures 7 to 11. All data has now been input into the Access database along with the soil moisture measurements, which have been carried out.
In each of the French catchments there is good data availability and coverage of all time varying data types. Daily rainfall data from a sufficient number of rainfall stations is available in each of the major catchments to provide a good spatial distribution for the modelling exercises. Numbers of rain gauge per catchment are given in Table 2. Each of the catchments has a daily flow record for at least one stream gauge station; in the case of the Grand Morin there are two. Potential evapotranspiration (PET) for 10 day intervals has been calculated for each catchment using data from the nearest meteorological station. Three stations have been used, Reims, Saint Dizier and Auxerre. The length of records and type of data available for each of the catchments is summarised in Figures 7 to 11.

Table 2 : Hydrometeorological data available for the French catchments

Catchment	Stream gauge at the catchment outlet	Daily Rainfall stations (N°)	PET data
Grand Morin	Montry	16	Reims
	Crécy La Chapelle	16	Reims
Saulx	Vitry en Perthois	31	Saint Dizier
Serein	Chablis	21	Auxerre
Orgeval	L'Orgeval au Theil	5	Reims
Petit Morin	Gauge No H5412020	6	Reims

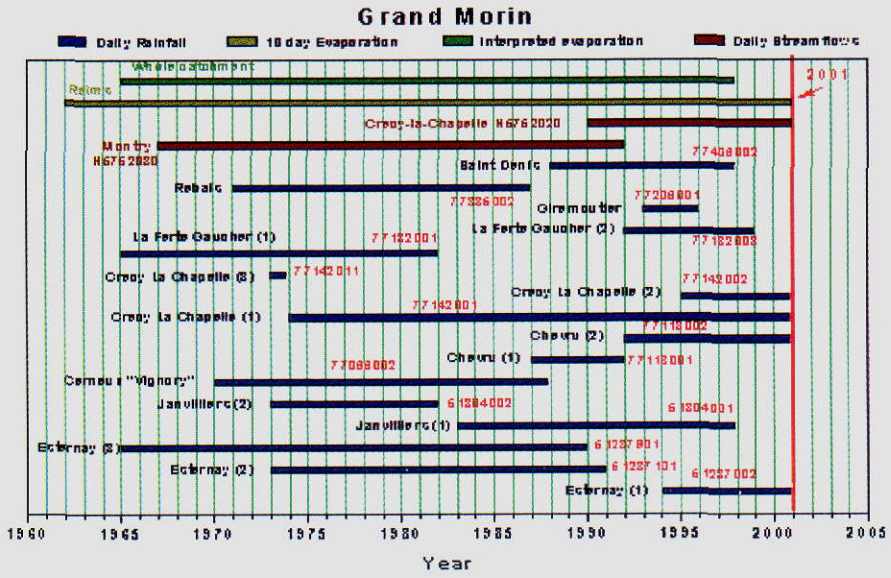


Figure 7

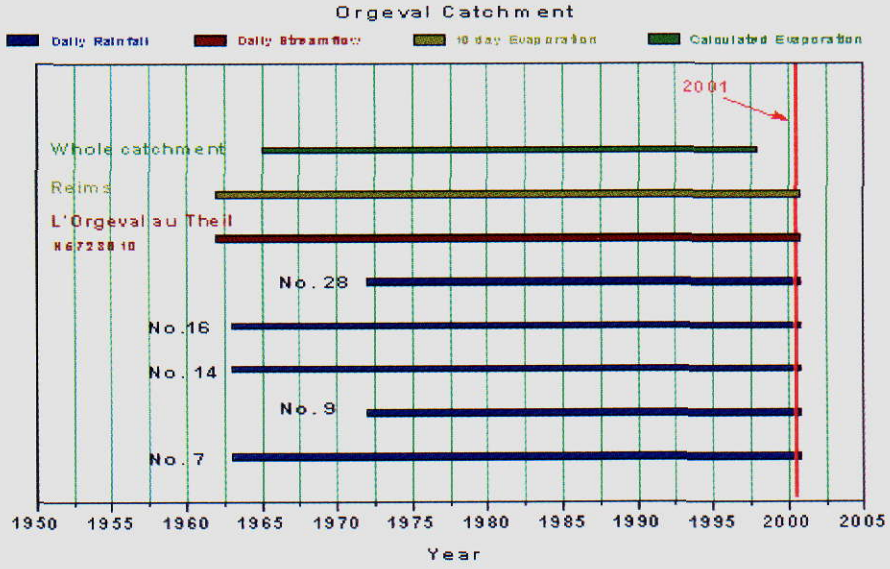


Figure 8

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

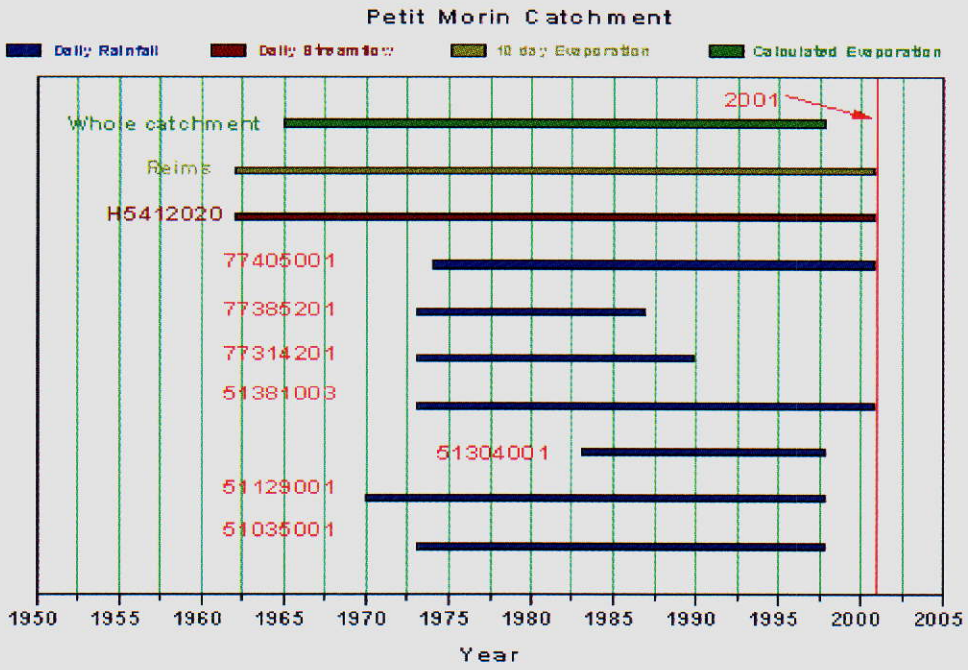


Figure 9

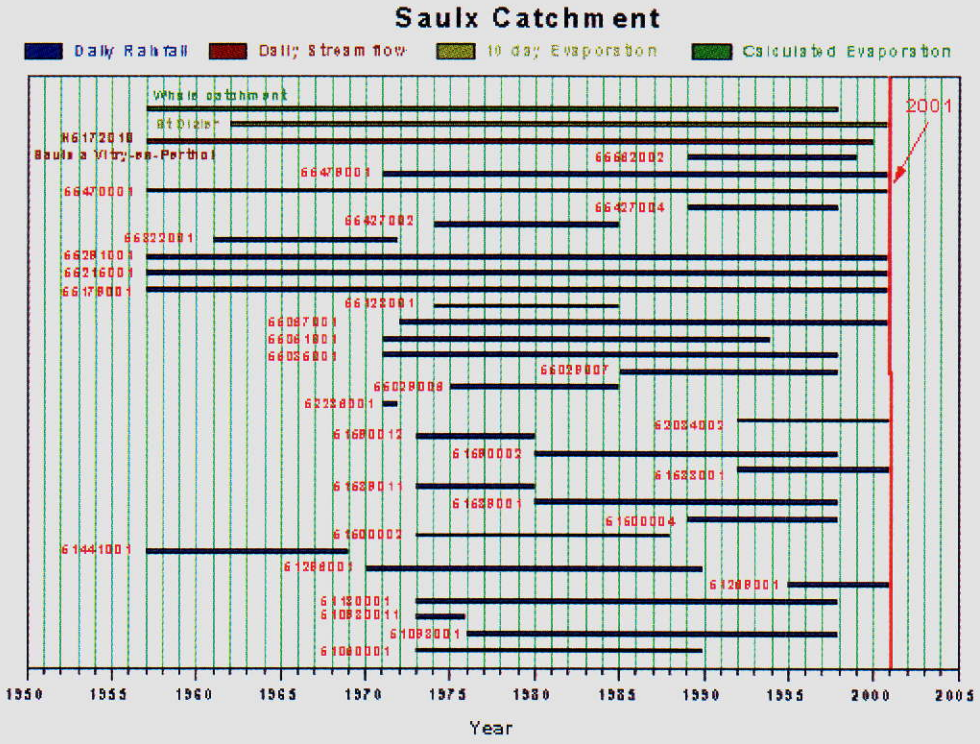


Figure 10

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

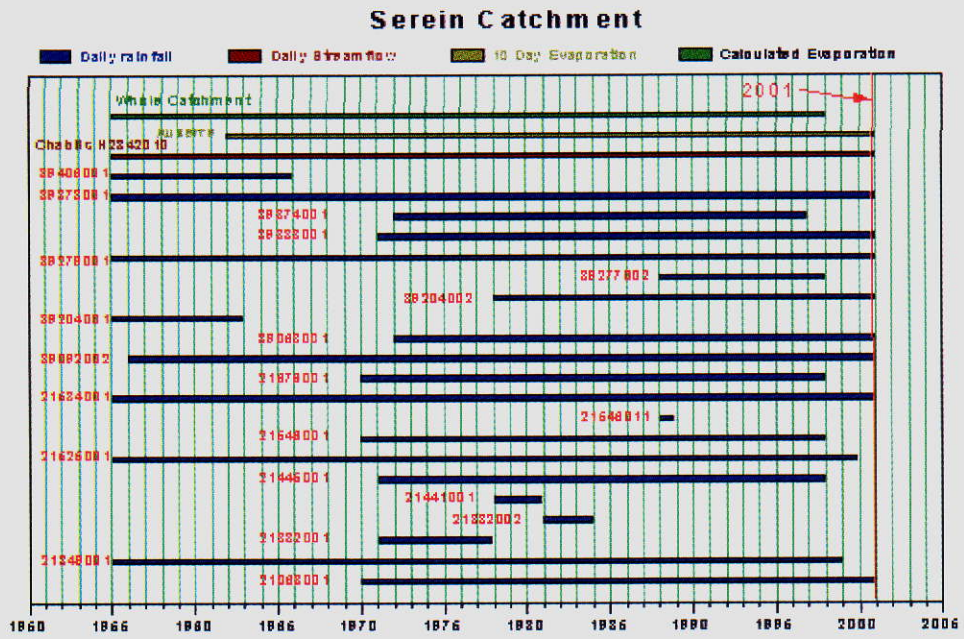


Figure 11

- *Time-varying data: Portuguese (Arade) Catchment*

The data available for the Arade catchment is summarised on Figure 12. Most of the useful daily time varying data is from one site, Funcho dam. Here daily records are available from 1993 to 2000. Long-term daily rainfall records are available for two other sites, Messi and Bartlolomeu, but these were discontinued in 1960 and 1994 respectively. Other daily records are available from 5 other stations but these are restricted to a few months duration.

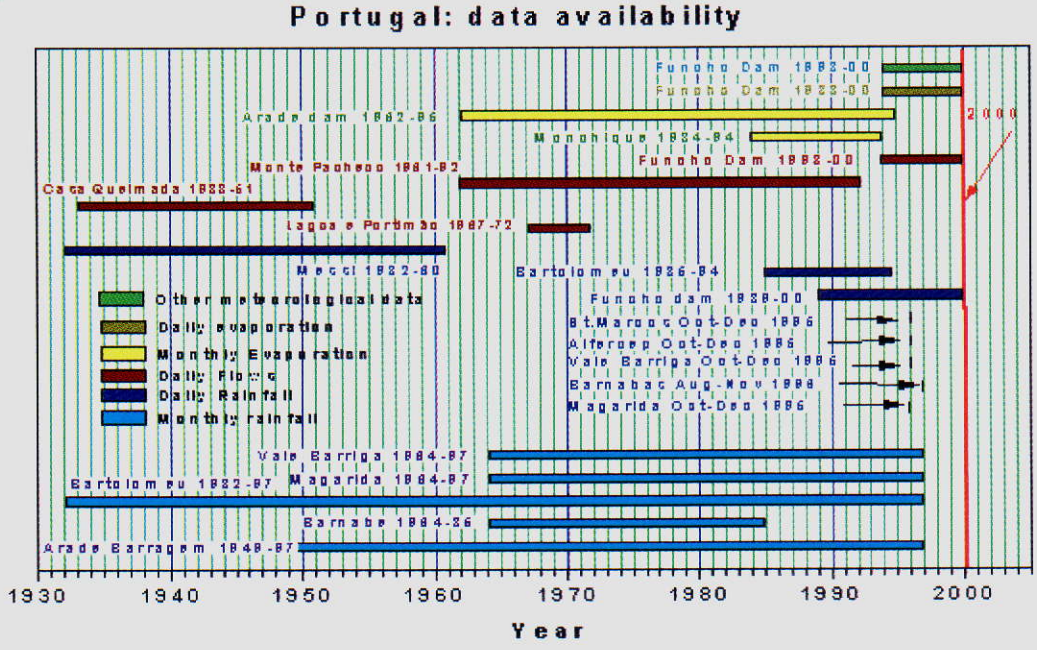


Figure 12

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

Daily stream flow and evaporation is also available for the Funcho site from 1993 to 2000. Other daily stream flow data for Monte Pacheco and Casa Queimada were discontinued in 1992 and 1951 respectively. The modelling input for this catchment has, therefore, been largely restricted to the time-varying data obtained from the Funcho dam site. An updated copy of the Access database is currently lodged on the CEMAGREF FTP site in Antony, France.

- Field Measurements made at the French Catchments

Field measurements to assess the radar calibration/soil moisture relationship in the Grand Morin and Serein catchments began in January 1999. These measurements included:

i. automatic permanent measurements (performed with the Time Domain Reflectometry method) on each sub-catchment ;
ii. extensive soil sampling which coincides with radar satellite images over the test basins ;
iv. establishment of crop maps (beginning of July 1999) over test sites (about 50km² for each catchment).
v. vegetation parameter measurements

i. Automatic permanent point measurements (performed with the Time Domain Reflectometry method) on each sub-catchment

For the Grand Morin, an automatic TDR recorder was already installed at Boissy le Chatel, near the Orgeval sub catchment. This little basin is an experimental and research catchment managed by the CEMAGREF. For the Serein, two permanent measurements (Chablis and Nitry) have been installed in January 1999. The method used to measure volumetric soil moisture is also the Time Domain Reflectometry (TDR). To install the probes at different depths a trench was constructed to the depth of the bedrock or water table. Then the probes were horizontally inserted in the soil profile of the trench at different depths (figure 13) without disturbing the soil layer. The last step consists in refilling the trench with the original soil. For each permanent measurements, volumetric soil moisture is measured every 12 hours and every 10 cm from the top to the bedrock or the water table:

- at Boissy le Chatel (Gd Morin) there are measurements from the top to 105 cm (16 probes, with 3 probes for the first three depths : 5, 15 and 25 cm);
- at Chablis (Serein) from the top to 105 cm (16 probes, with 2 probes for the first 6 depths : 5, 15, 25, 35, 45; 55 and 65 cm);
- at Nitry (Serein) from the top to 65 cm (14 probes, with 3 probes for the first three depths : 5, 15 and 25 cm).

Results of the measurements at several depths at Chablis are shown in figure 14 . The deeper the probes are, the smoother the response is. There is a good reaction to the rainfall at 5, 15, 25, 35 cm. For intermediary depths, the soil moisture does not vary (45, 55, 65, 75 cm). And near the water table, soil moisture is probably under the influence of the water table (85, 95, 105 cm).

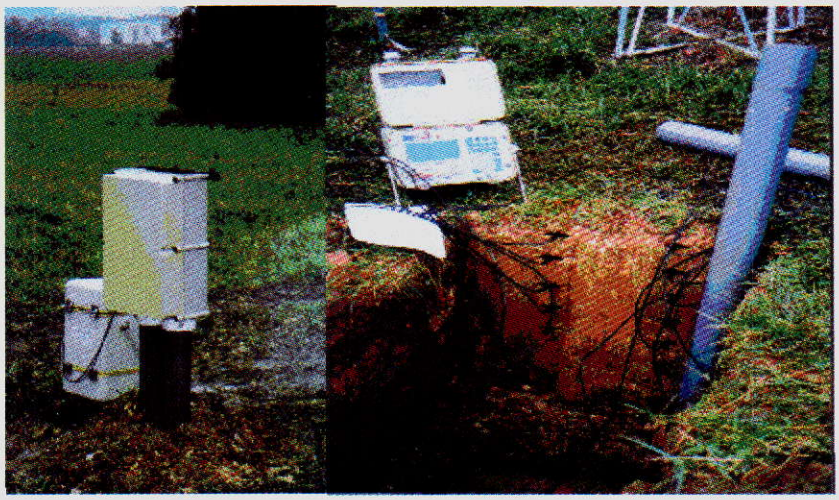


Figure 13: installation of TDR probes (example of the TDR in Chablis)

ii. Extensive soil sampling which coincide with radar satellite images over the test basins

Extensive field campaigns are carried out at each radar satellite passes over the Grand Morin catchment and the Serein catchment, in order to measure volumetric soil moisture (except during frost and snow periods).

For this purpose, test fields have been selected over the catchments. The choice of these test fields encompass the main agricultural practices and vegetation of the site. For example, 10 test fields were selected for the Serein catchment : 3 wheat fields, 5 barley fields and 2 grasslands, these crops are representative of the vegetation cover over this basin (see figures 15 and 17).

For the Grand Morin catchment, the10 test fields are located in the Orgeval basin.

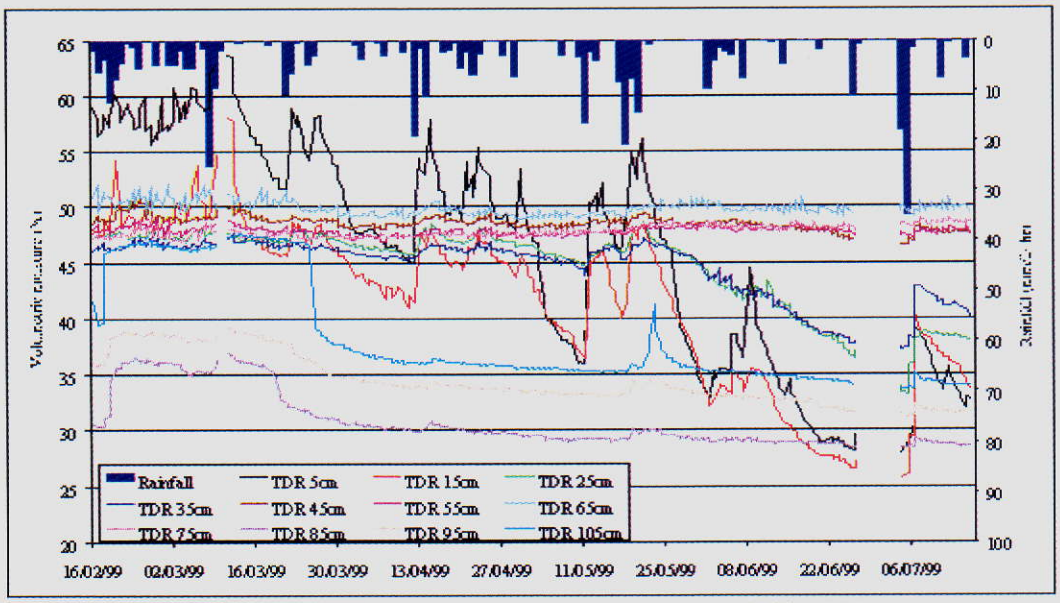
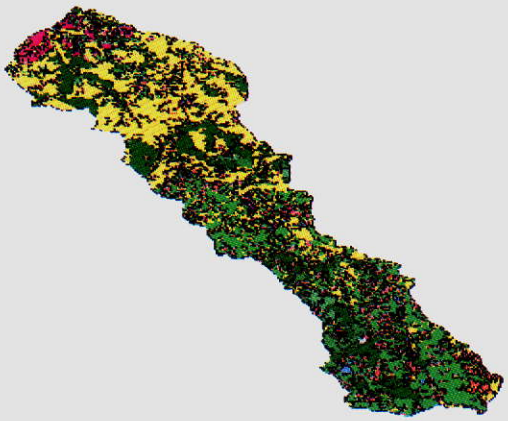


Figure 14: volumetric soil moisture measurements at Chablis (TDR)



34% cultivated areas
25% forest
24% grassland

Figure 15: vegetation cover over the Serein catchment (CORINE land cover data base)

The gravimetric method is used to evaluate the volumetric soil moisture from the test fields.

On each field, samples of soil are collected from the first 0-5 cm to 0-10 cm of the soil surface layer (weight varies between 200 and 300 grams). The number of samples depends on the field size (4-6ha) : 10 to 20 samples are collected for each field (see figure 16). These samples are spatially distributed over the fields.

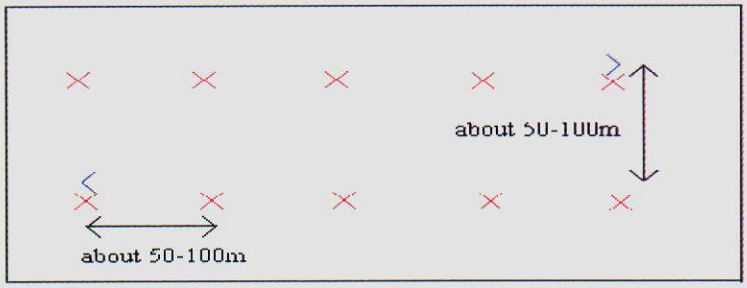


Figure 16: spatial repartition of the soil samples taken over a test field

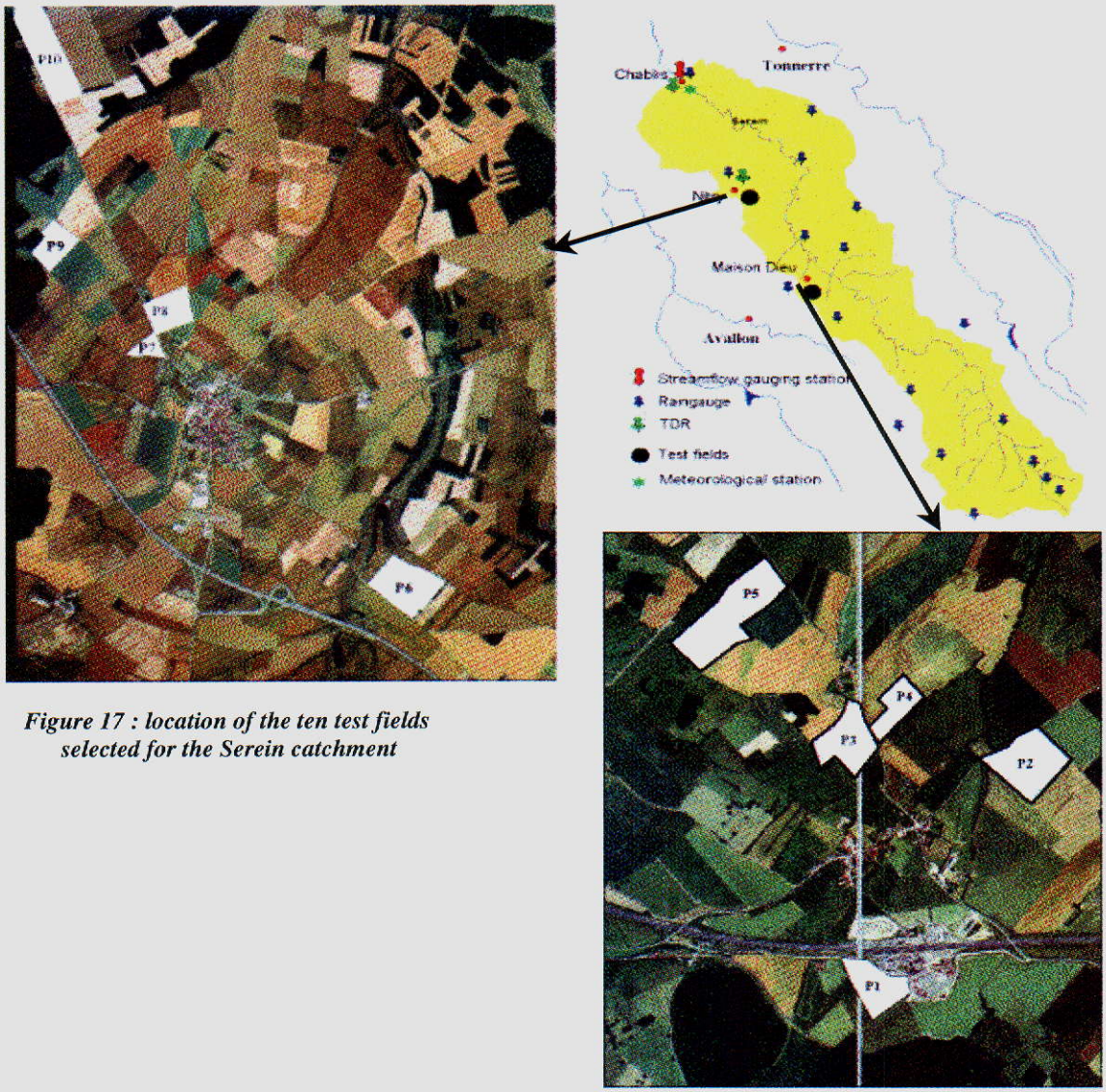


Figure 17 : location of the ten test fields selected for the Serein catchment

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

The samples are weighed before and after being dried in an oven for 48 hours at 105-110 C°. The volumetric soil moisture is given by :

$$H_V = \frac{W_W - W_D}{W_D} \times D_b \times 100 \quad (\%)$$

where W_W is the wet weight, W_D the dry weight and D_b is the bulk density of the concerned field.
For one day and for each field, an average value of the volumetric soil moisture is calculated using the values of volumetric soil moisture for each sample.

iii. Measurements of bulk density
In order to calculate soil moisture on a volumetric basis, it is necessary to know the bulk density of each test field:

$$D_b = \frac{W_D}{V} \quad (\text{g.cm}^{-3})$$

where V is the volume of the soil sample for the dry weight W_D.

Test fields	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
	Grassland		Wheat	Barley	Wheat	Barley	Barley	Barley	Barley	Wheat
Sample 1	1.26	1.41	1.49	1.46	1.29	1.14	1.41	1.37	1.22	0.98
Sample 2	1.22	1.24	1.54	1.43	1.36	1.25	1.40	1.05	1.33	0.93
Sample 3	1.26	1.33	1.46	1.35	1.36	1.27	1.30	1.08	1.20	1.07
Sample 4	1.37	1.33	1.50	1.51	1.28	1.22	1.33	1.11	1.34	0.99
Sample 5	1.12	1.14	1.49	1.46	1.43	1.22	1.32	1.10	1.17	1.00
Sample 6	1.29	1.50	1.48	1.40	1.44	1.10	1.36	1.09	1.26	0.98
Sample 7	1.18	1.31	1.47	1.40	1.48	1.05	1.33	1.04	1.32	0.97
Sample 8	1.36	1.09	1.49	1.43	1.43	0.96	1.12	1.02	1.24	0.98
Sample 9	1.17	1.09	1.51	1.37	1.44	1.21	1.28	1.11	1.42	1.14
Sample 10	1.42	1.39	1.46	1.48	1.41	1.22	1.25	1.17	1.18	1.17
Mean for the test field	**1.27**	**1.28**	**1.49**	**1.43**	**1.39**	**1.16**	**1.31**	**1.11**	**1.27**	**1.02**
Standard deviation for the test field	0.09	0.14	0.03	0.05	0.07	0.10	0.08	0.10	0.08	0.08

Table 3: results of bulk density of the ten test fields of the Serein catchment, July 1999

To measure the bulk density for one test field, we collect 10 undisturbed samples of soil in the test field. The samples are taken with cylinders of known volume. An average value is used for each test field.
The bulk density varies during the year, particularly when there are agricultural changes, like tillage, sowing or harvest. So, measurements need to be made several times during the year.

iv. Establishment of crop maps (began on July 1999) over test sites (about 50 km² for each catchment)
Maps of crop distribution have been established over selected areas of the Serein basin (one test area of about 30 km² around Nitry and one test area of about 20 km² around Maison Dieu) and the Grand Morin basin (one area of about 50 km², in the Orgeval basin). Figure 18 shows an example of the map obtained.

Field contours were first mapped using detailed topographical maps (1/10 000) and aerial photography (1/15 000). Then field campaigns were carried out to check which crops correspond to each field.
These maps are established in order to assess the remotely sensed crop classification that will be established from ERS and LANDSAT images.

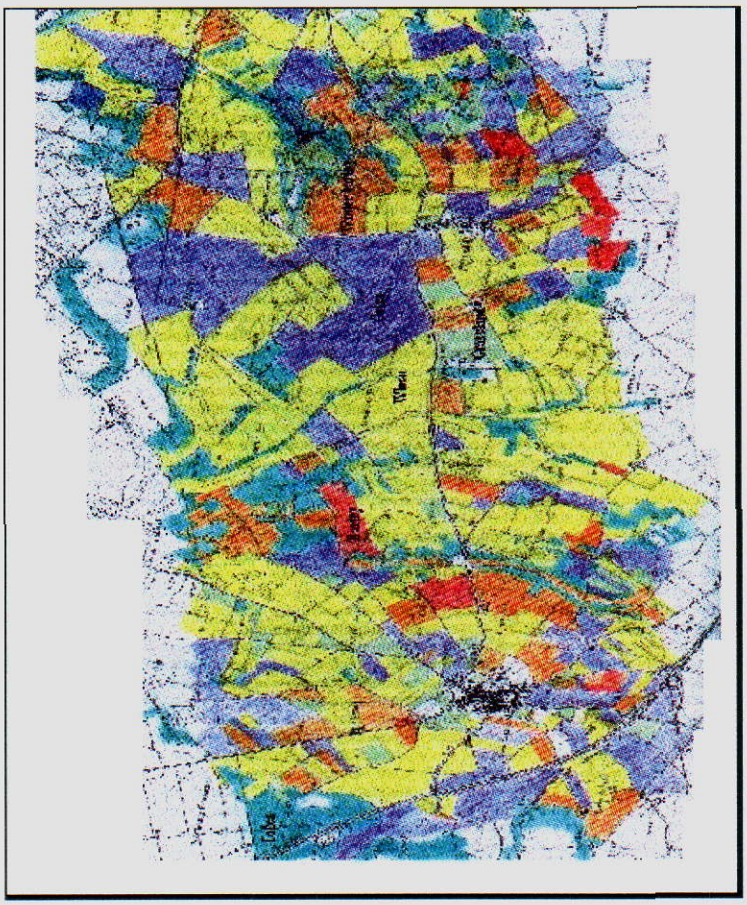


Figure 18: Crop map established over the Serein catchment around Nitry (red: barley, orange: winter barley, yellow: wheat, light green: grasslands, dark green: forest, blue: colza)

v. Vegetation parameter measurements

During the calibration year, the vegetation features have also be measured in order to get accurate estimates of the vegetation effect and free from it, during the calibration step, and to derive the empirical relationship between the crop total height and the other geometrical dimensions of the crop.

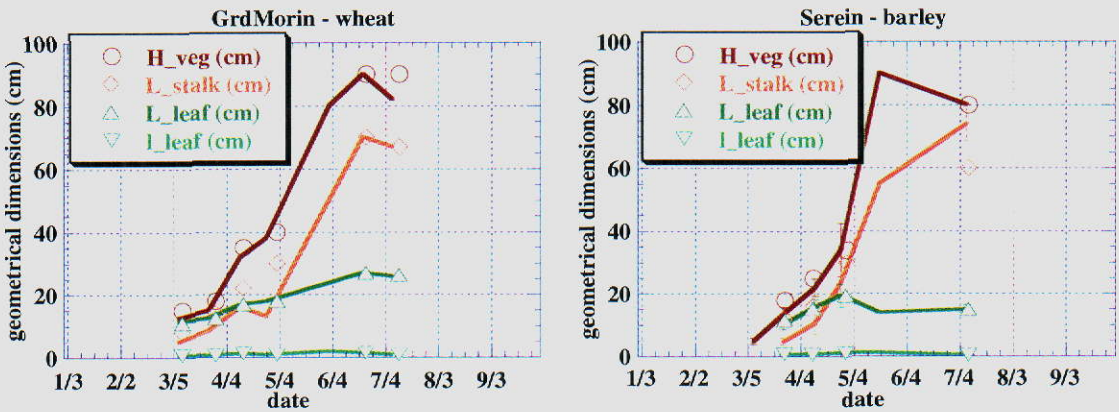


Figure 19: Evolution of some geometrical parameters (a) case of the wheat over the Grand Morin catchment, (b) case of the barley over the Serein catchment.

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

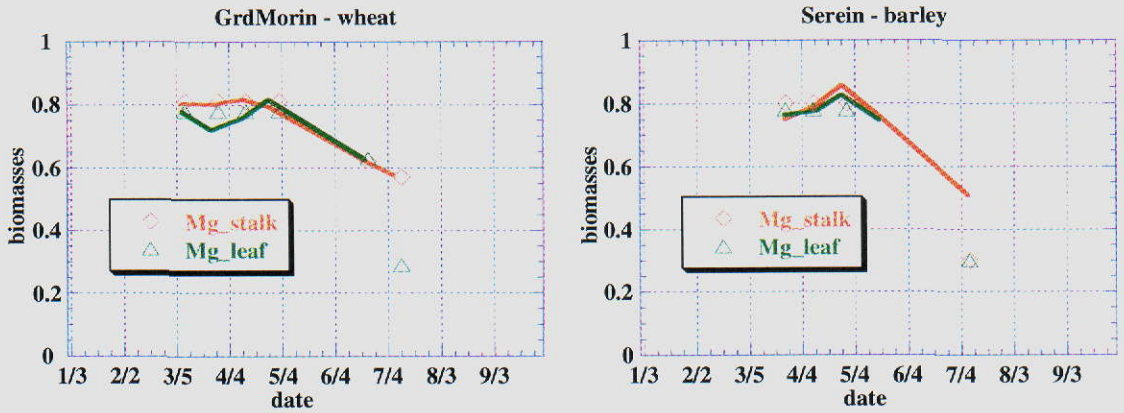


Figure 20: Evolution of the biomasses (a) case of the wheat over the Grand Morin catchment, (b) case of the barley over the Serein catchment.

Figure 19 shows, as an example of the vegetation ground truth measurements, the evolution of some of the vegetation geometrical parameters: the total vegetation layer height, the length of the stalks, the length and the larger of the leaves. Figure 20 shows some examples of the evolution of the leaf and stalk biomasses.

- *Field Measurements made at the Arade Catchment*

For Arade catchment field measurements began in January 2000. These measurements included:
i. automatic moisture measurements (performed with a portable TDR);
ii. Gravimetric moisture sampling;
iii. Measurements of bulk density;

i. Automatic moisture measurements (performed with the ThetaProve soil moisture sensor).
For the Arade catchment, a portable TDR was used to measure volumetric soil moisture content. With this TDR installation is very simple, just push the probe into the soil until the rods are fully covered, and connect the power supply to take readings from the analogue output. Measurements were obtained with this system for a depth of 5 cm. TDR was easily transported to each one of the measurement points due to its small size and weight. Measurements were taken each week in four measurements points distributed over the study area.

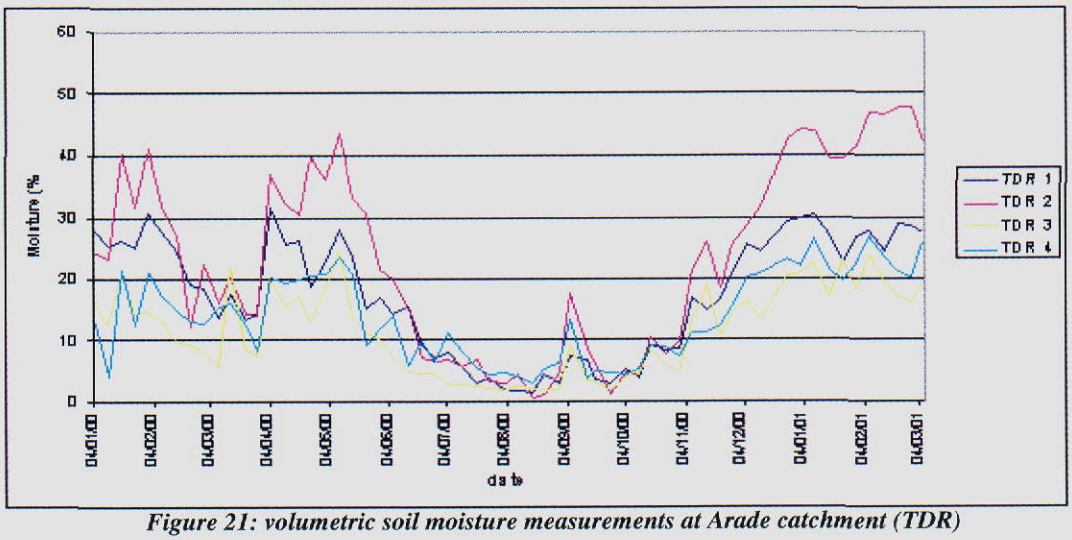


Figure 21: volumetric soil moisture measurements at Arade catchment (TDR)

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

ii. Gravimetric moisture sampling

Gravimetric moisture measurements were taken in the same placements with the same frequency than TDR measurements in order to calibrate TDR values. The process to obtain the volumetric soil moisture is equal to the described for French catchments. Values obtained are shown in figure 22. As we can see gravimetric values follow the same tendency that TDR measurements.

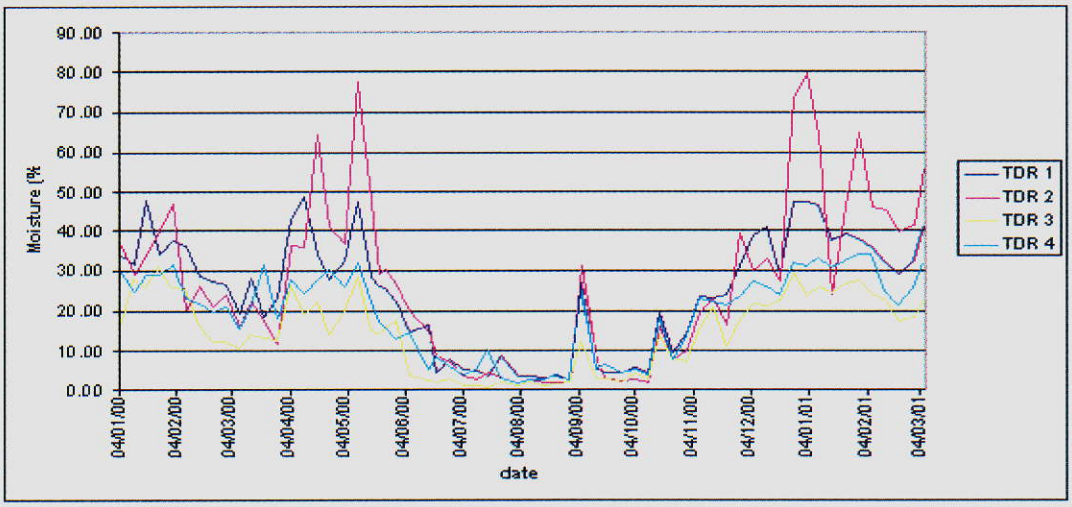


Figure 22: volumetric soil moisture measurements at Arade cacthment (gravimetric)

iii. Measurements of bulk density

Unlike French catchments, bulk density was measured only a time for each one of the measurement points. This is not a critical fact because the areas of measurement points in Arade are no dedicated to agricultural exploitation, so there are no big changes in soil. The measurement procedure is similar to the explained in French catchments.

Sample 1	Sample 2	Sample 3	Sample 4
1.4303	1.0524	1.2639	1.3782

Table 4: bulk density measurements for the Arade catchment (21/11/2000)

• Spatial data and the web site database

A wide range of spatial data in various formats has been accessed during the course of the project. These include ERS/SAR, LANDSAT/TM and SPO/HRVT images; soil maps; and land use maps. Table 5 lists the number and type of satellite images that have been used in the project. Some spatial soil and land use information has been obtained from the CORINE (Co-ORdination of INformation on the Environment) information system, which is available on CD. These are a series of land use maps at a scale of 1:100 000 produced by the EC.

Catchment	ERS/SAR images (N°)	LANDSAT/TM images (N°)	SPOT/HRV images (N°)	Calibration	Validation
Grand Morin	15			x	x
Grand Morin		1		x	
Serein	14			x	x
Serein		1		x	
Petit Morin	14			x	x
Saulx	24				x
Arade	9				
Arade		1			
Arade			1		

Table 5: Type and number of satellite images used in Aimwater

More details of the system may be found on the following Web site [http://etc.satellus.se/the_data/overview.htm]. Soil parameters for the various soil types identified by CORINE were obtained from the FAO Soil Map of the World, Volume 5 (Europe) (FAO-UNESCO, 2001). In addition to land use and soil maps, digital elevation maps (DEM) for each catchment have been obtained from national Ordnance Survey sources.

A more complete description of the EO database is provided for the Seine sub-catchments. The images have been acquired by two satellite instruments: the ERS2/SAR and the LANDSAT/TM. The SAR sensor main features are: wavelength ≈ 5.6 cm (C band: frequency $f \approx 5.3$ GHz), VV polarization system, incidence angle ≈ 23 ° in the middle of the swath, spatial resolution: 25×25 m^2, pixel size 12.5×12.5 m^2. LANDSAT/TM is an multispectral optical sensor: 3 channels in the visible domain, 1 near infrared, 2 medium infrared, and 1 thermal infrared. LANDSAT data is only used for the calibration year (1999) and to improve the classification results by combining its information with the SAR one.

- Remote sensing data over the Seine sub-catchments

Tables 6,7,8 list the acquired images (ERS and LANDSAT). In these tables, "A" denotes an ascending orbit, and "D" a descending one. These images have been superposed and geo-referenced according the process described just below. They have then been filtered and their size reduced by a factor 2 or 3 for limitation of the computer memory necessary for further processes.

♦ Image superposition

All the data images have been superposed such that a pixel pointed out by its line and column numbers represents the same target in all images.

In the case of ERS images corresponding to a same track, this superposition was obtained thanks to a simple image translation. Indeed, the geometry of acquisition of the ERS images corresponding to a same track is about the same, excepting the origin coordinates. Thus, a simple translation is needed for their superposition. In our case, the translation vectors were founded by maximization of the correlation between couples of images belonging to the same track.

For the superposition of the two ERS tracks, we use a second order polynomial transformation. The coefficients of this transformation have been estimated by minimization of the least square error η^2 computed over a selection of reference points located in each track geometry. Then, the pixel values are deduced using an interpolation between the 4 nearest neighbours.

In the same way, the superposition of the ERS and the LANDSAT images was performed by selection of reference point targets, minimization of the quadratic error η^2, and projection in the same geometry using the second order polynomial transformation and an interpolation between the 4 nearest neighbours.

After geo-referencing, the coordinates of the four corners of the images are:

- Grand Morin: (632027.,409956.), (698135.,396140.), (704385.,425500.), (638239.,439311.);
- Serein: (693666.,256170.), (764403.,242245.), (777654.,306587.), (706832.,320391.);
- Petit Morin: (656787.,429333.), (721816.,415988.), (724408.,428219.), (659364.,441560.);
- Saulx: (764499.,385708.), (838165.,372111.), (849118.,428692.), (775339.,442175.).

The size in numbers of lines and columns of the full resolution images (S) are respectively equal to 2120 lines × 5500 columns for the Grand Morin, 5250 lines ×5850 columns for the Serein, 1000 lines ×5400 columns for the Petit Morin, 4600 lines × 6100 columns for the Saulx.

♦ Speckle filtering

We used an averaging window to reduce the speckle: is the averaged value over a 3×3 window. The size of this window has been empirically chosen as a compromise between speckle reduction and preservation of small image features (for image classification).

Moreover, in order to make the image speckle have the characteristics of an additive noise we consider dB values of: $<\sigma> : < \sigma_{db} >$.

date	orbit	frame	track	geo-referenced (S: superposed, Rn: size reduced by factor n)	
1999 / 01 / 18	D-19587	2619	194	GM: S, R2;	PM: S, R2
1999 / 01 / 31	A-19780	0963+09	387	GM: S, R2;	PM: S, R2
1999 / 02 / 03	D-19816	2619	423	GM: S, R2;	PM: S, R2
1999 / 02 / 22	D-20088	2619	194	GM: S, R2;	PM: S, R2
1999 / 03 / 10	D-20317	2619	423	GM: S, R2;	PM: S, R2
1999 / 03 / 29	D-20589	2619	194	GM: S, R2;	PM: S, R2
1999 / 04 / 11	A-20782	0963+09	387	GM: S, R2;	PM: S, R2
1999 / 04 / 14	D-20818	2619	423	GM: S, R2;	PM: S, R2
1999 / 05 / 03	D-21090	2619	194	GM: S, R2;	PM: S, R2
1999 / 05 / 03	LANDSAT/TM			GM: S, R3;	PM: not used
1999 / 05 / 19	D-21319	2619	423	GM: S, R2;	PM: S, R2
1999 / 06 / 07	D-21591	2619	194	GM: S, R2;	PM: S, R2
1999 / 06 / 20	A-21784	0963+09	387	GM: S, R2;	PM: S, R2
1999 / 06 / 23	D-21820	2619	423	GM: S, R2;	PM: S, R2
1999 / 07 / 12	D-22092	2619	194	GM: S, R2;	PM: S, R2
1999 / 07 / 28	D-22321	2619	423	GM: S, R2;	PM: S, R2
1999 / 08 / 16	D-22593	2619	194	GM: S, R2;	PM: S, R2
1999 / 08 / 29	A-22786	0963+09	387	GM: S, R2;	PM: S, R2
1999 / 09 / 20	D-23094	2619	194	GM: S, R2;	PM: S, R2
1999 / 10 / 06	D-23323	2619	423	GM: S, R2;	PM: S, R2
1999 / 10 / 25	D-23595	2619	194	GM: S, R2;	PM: /
1999 / 11 / 07	A-23788	0963+09	387	GM: S, R2;	PM: S, R2
1999 / 11 / 10	D-23824	2619	423	GM: S, R2;	PM: S, R2
1999 / 11 / 29	D-24096	2619	194	GM: S, R2;	PM: S, R2
1999 / 12 / 15	D-24325	2619	423	GM: S, R2;	PM: S, R2
2000 / 01 / 19	D-24826	2619	423	GM: S, R2;	PM: S, R2
2000 / 03 / 13	D-25599	2619	194	GM: S, R2;	PM: S, R2
2000 / 03 / 26	A-25792	0963+09	387	GM: S, R2;	PM: S, R2
2000 / 03 / 29	D-25828	2619	423	GM: S, R2;	PM: S, R2
2000 / 04 / 17	D-26100	2619	194	GM: S, R2;	PM: S, R2
2000 / 05 / 03	D-26329	2619	423	GM: R2;	PM: S, R2
2000 / 05 / 22	D-26601	2619	194	GM: R2;	PM: S, R2
2000 / 06 / 04	A-26794	0963+09	387	not used	
2000 / 06 / 07	D-26830	2619	423	GM: R2;	PM: not used
2000 / 06 / 26	D-27102	2619	194	GM: R2;	PM: S, R2
2000 / 07 / 12	D-27331	2619	423	GM: R2;	PM: S, R2
2000 / 07 / 31	D-27603	2619	194	GM: R2;	PM: S, R2
2000 / 08 / 01	LANDSAT/TM			not used	
2000 / 08 / 16	D-27832	2619	423	GM: R2;	PM: S, R2
2000 / 09 / 04	D-28104	2619	194	GM: R2;	PM: S, R2
2000 / 09 / 20	D-28333	2619	423	GM: R2;	PM: S, R2
2000 / 10 / 25	D-28834	2619	423	GM: R2;	PM: /
2000 / 11 / 13	D-29106	2619	194	GM: R2;	PM: S, R2
2000 / 11 / 29	D-29335	2619	423	GM: R2;	PM: S, R2
2000 / 12 / 18	D-29607	2619	194	GM: R2;	PM: S, R2
2001 / 01 / 22	D-30108	2619		not acquired (satellite problems)	
2001 / 02 / 07	D-30337	2619		not acquired (satellite problems)	
2001 / 02 / 26	D-30609	2619		GM: R2;	PM: S, R2
2001 / 03 / 14	D-30838	2619		GM: R2;	PM: S, R2
2001 / 04 / 02	D-31110	2619		GM: R2;	PM: S, R2

Table 6: ERS acquisitions over the Grand Morin (GM) and Petit Morin (PM) watersheds.

date	orbit	frame	track	geo-referenced (S: superposed, Rn: size reduced by factor n)
1999 / 01 / 12	A-19508	0945	115	R2
1999 / 01 / 15	D-19544	2655-08	151	S, R2
1999 / 01 / 31	D-19773	2655-06	380	S, R2
1999 / 01 / 31	A-19780	0945	387	R2
1999 / 02 / 19	D-20045	2655-08	151	S, R2
1999 / 03 / 26	D-20546	2655-08	151	S, R2
1999 / 04 / 01	LANDSAT/TM			S, R2
1999 / 04 / 11	D-20775	2655-06	380	S, R2
1999 / 04 / 11	A-20782	0945	387	R2
1999 / 04 / 30	D-21047	2655-08	151	S, R2
1999 / 05 / 16	D-21276	2655-06	380	S, R2
1999 / 06 / 01	A-21512	0945	115	R2
1999 / 06 / 04	D-21548	2655-08	151	S, R2
1999 / 06 / 20	D-21777	2655-06	380	S, R2
1999 / 06 / 20	A-21784	0945	387	R2
1999 / 07 / 09	D-22049	2655-08	151	S, R2
1999 / 07 / 25	D-22278	2655-06	380	S, R2
1999 / 08 / 10	A-22514	0945	115	R2
1999 / 08 / 13	D-22550	2655-08	151	S, R2
1999 / 08 / 29	D-22779	2655-06	380	S, R2
1999 / 09 / 17	D-23051	2655-08	151	S, R2
1999 / 10 / 03	D-23280	2655-06	380	S, R2
1999 / 10 / 19	A-23516	0945	115	R2
1999 / 10 / 22	D-23552	2655-08	151	S, R2
1999 / 11 / 07	D-23781	2655-06	380	S, R2, R3
1999 / 11 / 07	A-23788	0945	387	R2, R3
1999 / 11 / 26	D-24053	2655-08	151	S, R2, R3
1999 / 12 / 12	D-24282	2655-06	380	S, R2, R3
1999 / 12 / 28	A-24518	0945	115	R2, R3
2000 / 02 / 04	D-25055	2655-08	151	S, R2, R3
2000 / 02 / 20	D-25284	2655-06	380	S, R2, R3
2000 / 03 / 07	A-25520	0945	115	R2, R3
2000 / 03 / 10	D-25556	2655-08	151	S, R2, R3
2000 / 03 / 26	D-25785	2655-06	380	S, R2, R3
2000 / 04 / 14	D-26057	2655-08	151	S, R2, R3
2000 / 04 / 30	D-26286	2655-06	380	R2, R3
2000 / 05 / 16	A-26522	0945	115	R2
2000 / 05 / 19	D-26558	2655-08	151	R2, R3
2000 / 06 / 04	D-26787	2655-06	380	R2, R3
2000 / 06 / 04	A-26794	0945	387	R2
2000 / 07 / 09	D-27288	2655-06	380	R2, R3
2000 / 07 / 28	D-27560	2655-08	151	R2, R3
2000 / 08 / 13	D-27789	2655-06	380	R2, R3
2000 / 09 / 01	D-28061	2655-08	151	R2, R3
2000 / 09 / 17	D-28290	2655-06	380	R2, R3
2000 / 10 / 06	D-8562	2655-08	151	R2,R3
2000 / 10 / 22	D-28791	2655-06	380	R2,R3
2000 / 11 / 10	D629063	2655-08	151	R2,R3
2000 / 11 / 26	D-29292	2655-06	380	R2,R3
2000 / 12 / 15	D-29564	2655-08	151	R2,R3
2001 / 01 / 19	D-30065	2655-08		not acquired (satellite problems)
2001 / 02 / 04	D-30294	2655-06		not acquired (satellite problems)
2001 / 02 / 23	D-30566	2655-08		R2,R3
2001 / 03 / 11	D-30795	2655-06		R2,R3
2001 / 03 / 30	D-31067	2655-08		R2,R3

Table 7: ERS acquisitions over the Serein basin.

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

date	orbit	frame	track	geo-referenced (S: superposed, Rn: size reduced factor n)
1999 / 01 / 31	D-19773	2625	380	R2
1999 / 02 / 16	D-20002	2627	108	R2
1999 / 03 / 07	D-20274	2625	380	R2
1999 / 03 / 23	D-20503	2627	108	R2
1999 / 04 / 11	D-20775	2625	380	R2
1999 / 04 / 24	A-20968	0970		not processed
1999 / 04 / 27	D-21004	2627	108	R2
1999 / 05 / 13	A-21240	0969		not processed
1999 / 05 / 16	D-21276	2625	380	R2
1999 / 06 / 01	D-21505	2627	108	R2
1999 / 06 / 20	D-21777	2625	380	R2
1999 / 07 / 03	A-21970	0970		not processed
1999 / 07 / 06	D-22006	2627	108	R2
1999 / 07 / 22	A-22242	0969		not processed
1999 / 07 / 25	D-22278	2625	380	R2
1999 / 08 / 10	D-22507	2627	108	R2
1999 / 08 / 29	D-22779	2625	380	R2
1999 / 09 / 14	D-23008	2627	108	R2
1999 / 10 / 03	D-23280	2625	380	R2
1999 / 10 / 19	D-23509	2627	108	R2
1999 / 11 / 07	D-23781	2625	380	R2
1999 / 11 / 20	A-23974	0970		not processed
1999 / 11 / 23	D-24010	2627	108	R2
1999 / 12 / 09	A-24246	0969		not processed
1999 / 12 / 12	D-24282	2625	380	R2
2000 / 01 / 29	A-24976	0970		not processed
2000 / 02 / 01	D-25012	2627	108	R2
2000 / 02 / 17	A-25248	0969		not processed
2000 / 02 / 20	D-25284	2625	380	R2
2000 / 03 / 07	D-25513	2627	108	R2
2000 / 03 / 26	D-25785	2625	380	R2
2000 / 04 / 11	D-26014	2627	108	R2

Table 8: ERS acquisitions over the Saulx basin.

Examples of images

Figure 23 and figure 24 show two examples of superposed images. They represent two colored compositions of ERS images respectively acquired over the Grand Morin basin and the Serein basin. We clearly see the different sub-part of the watersheds covered by the different tracks. In the case of the Serein site, we note the differences, between the ascending and descending modes, in terms of illumination for the areas with relief.

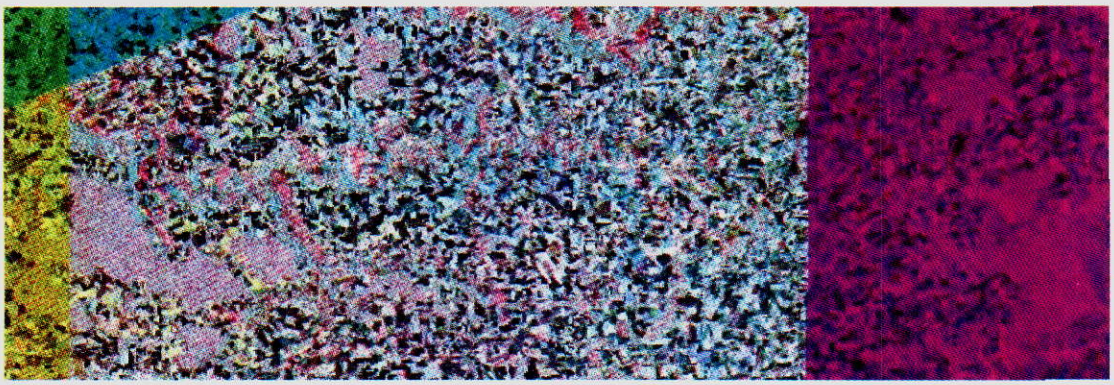


Figure 23: Colored composition of 3 ERS images (Red: 99/05/19 → track 423, Green: 99/06/07 → track 194, Blue: 99/06/20 → track 387) acquired over the Grand Morin basin

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

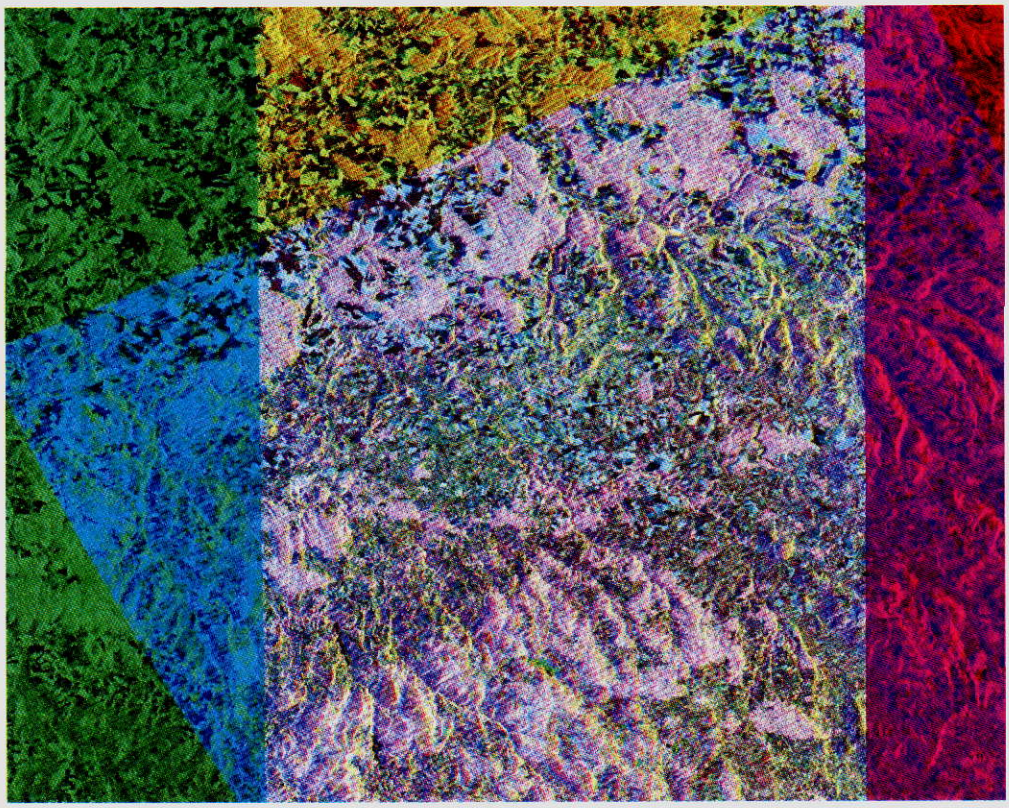


Figure 24: Colored composition of 3 ERS images (Red: 99/05/169 → track 380, Green: 99/06/04 → track 151, Blue: 99/06/01 → track 115) acquired over the Serein basin

- Organization of the Seine database

Figure 25 describes part of the database concerning the Serein watershed.

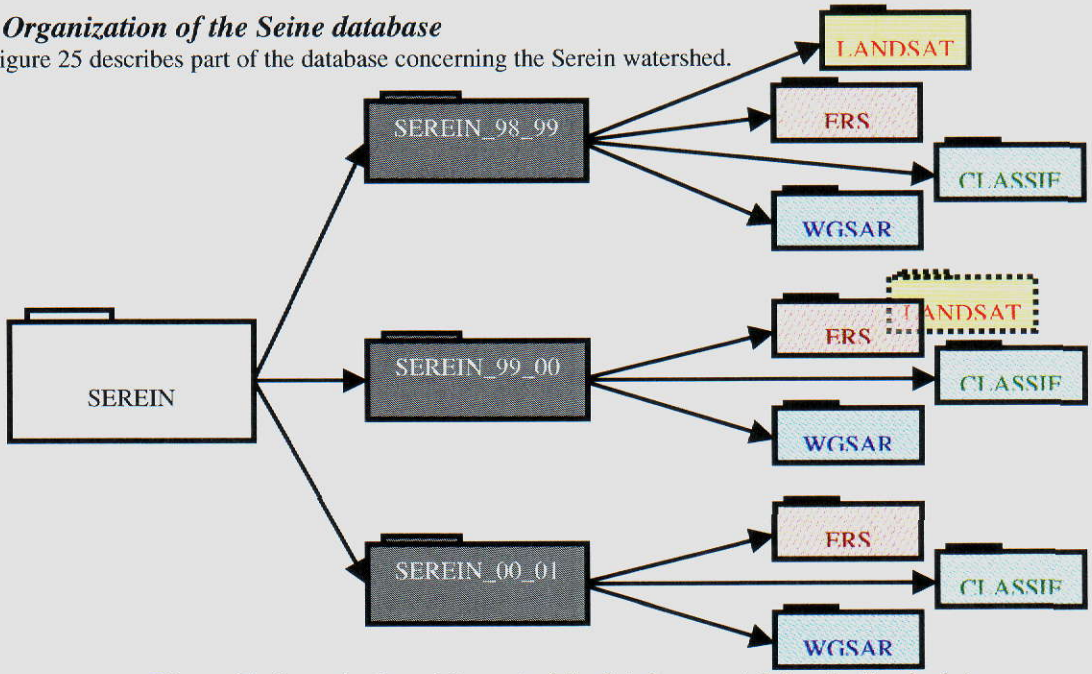


Figure 25: Organization of the part of the data base containing the Serein data.

Origin: H. Morel-Seytoux/Cemagref/U. Independente /
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

The directory labelled "SEREIN" is divided in three sub-directories respectively corresponding to each of the agricultural year considered (within the January 1999 to April 2001 period). Each of these directories per year is itself divided in four or three sub-directories corresponding to the different products offered by the database: geo-referenced images (ERS and LANDSAT when available), land cover maps derived from classifications (directories labelled "CLASSIF"), and soil moisture indix series (directories labelled "WGSAR").
The methodologies and algorithms used to derive land cover maps and soil moisture indices are described in chapter 3.
The same structure is reproduced for the other watersheds: the Grand Morin, the Petit Morin and the Saulx, except that in some cases less data are available.

2.2 User's requirements

This section proceeds, as concretely as possible, with a mixed description of the "needs", as expressed, to some degree by the user but mostly as perceived by the scientists, in general, for better management of the reservoirs, and, in particular, for soil moisture information from EO data, to reach that goal. This description served as a focus point for the direction of the research project.
It is worthwhile, to avoid later confusion, to review the meaning of terms used in this section and which have to be understood solely in the context of the AIMWATER project. It is also necessary to review the roles played in the project by, on one hand, the "scientists" funded by the project and on the other hand by a potential "user" such as the IIBRBS (Institution Interdépartementale de Barrages Réservoirs du Bassin Seine), which was not an actual partner. IIBRBS acted as an observer, albeit an interested one. In general, the IIBRBS, as an operational government agency, is accustomed to assess its needs on its own and, if it cannot satisfy them, it contracts out. In that relationship with a contractor IIBRBS defines, from the start, precisely its **requirements**. There is no ambiguity as to what the requirements are and whether, when the project is completed, they were met or not. This is not at all the contractual relationship that exists between AIMWATER and IIBRBS. Given that IIBRBS did not contribute to the project as a partner, so that here the word "requirement" has to be interpreted as a "wish list" on the part of IIBRBS. Useful outcome from the project is welcome but is not demanded. However if no requirements were placed on the project how would one know that it met its objectives? Fundamentally this means that the project is bound to scientific standards and must be subjected to technical cross-examination but is not contractually bound to meet specific user's requirements.
As indicated previously IIBRBS constantly assesses its needs for technical improvements in its operations. However it is neither a research agency nor a high-tech consulting company. It is quite possible that some outsiders with their specialized expertise could see that improvements could result from a change in their operations. They would see **"needs"** for additional information and tools in order to accomplish that purpose. These are the very specific needs as **perceived** by the scientists not as expressed in very general terms by the potential user. The practical success of AIMWATER, as opposed to its scientific one, will hinge on the capability of the scientists, through dialogue and/or demonstration, to convince the potential user that indeed their suggestions (1) are feasible in an economic and practical sense and (2) really will improve the operations.

- **Needs regarding real-time operations**

- Case of IIBRBS (Institution Interdépartementale des Barrages Réservoirs du Bassin Seine)

Historical background

L'IIBRBS (Institution Interdépartementale des Barrages Réservoirs du Bassin de la Seine) was created in 1969. It is a government organization (its statute is one of "établissement public"). Its mission is to protect the Paris region from floods and to sustain river flows during the dry weather months, tasks made possible thanks to the regulating capacity of the reservoirs. The IIBRBS has agreed to participate in the AIMWATER project and it has repeatedly stated that its concern is not to secure immediate results, hypothetically induced by specific desired outcomes. More basically its goal is to assist in the steps, which can lead to improvement in the knowledge and competence of hydrologists.

Current Infrastructure Situation

The current reservoir capacities (BCEOM, 1994) are for:

"Pannecière" on the Yonne	
(+ a volume in two reservoirs in the Morvan region) :	104 Mm3.
"Seine" off-stream reservoir, near the city of Troyes:	205 Mm3
"Marne" off-stream reservoir, near Saint-Dizier:	350 Mm3

"Aube" off-stream reservoir, near Brienne-le-Château: 170 Mm3

The total reservoir capacity for the system as a whole is thus of the order of 824 Mm3. That value is confirmed as approximately 830 Mm3 in a more recent document (Les Grands Lacs de Seine, 1998) with a slightly lower value being available for low flow augmentation, 800, versus 820 for flood protection.
What are some of the problems faced by the managers of the system?
The treatment capacity for drinking water of the Paris urban area already exceeds the dry weather minimal discharge observed during 10 consecutive days in Paris in 1976. Note also that in particular the Pannecière reservoir controls only 2% of the Yonne basin and the Paris region is not immune from floods generated in that basin.

Current Modelling Capability

Several years ago, the IIBRBS contracted a study, which led to the development of a river basin simulation model of the natural system and of the operations of the reservoirs. That computer model is called PEGASE.
Naturally, the calculated values for the releases by the model only serve as aid to the decision process. The operators do not, by any means, use these values automatically and blindly. In fact, at present, the models are probably not used at all for real-time operations but for simulations to verify that the rules of operation are reasonable and to estimate the level of satisfactory performance of the system over long periods of historical records, given a set of existing or potential infrastructures
The model PEGASE consists of two major simulation modules. One module describes the response of the river system, in terms of calculation of discharges at many points of interest in the basin, to daily rainfall over the entire basin. Runoff is generated in the various watersheds and routed through the main rivers of the system. The other module simulates the operation of the reservoirs and optimizes the releases to achieve selected objectives at various target points. Our discussion will focus on the "rainfall-runoff-routing" module.
The basin is subdivided into upstream watersheds, intermediate (or lateral) watersheds and a network of river segments. The amount of runoff that will result from a given rainfall will depend greatly on the initial degree of saturation of the basin, or in other words the amount of water held in storage in the basin, not necessarily entirely, but predominantly, in the soil. Currently the hydrologic components of PEGASE, available to the Institution, to calculate daily runoff from daily precipitation are:

(1) the GR-daily (or "GR" for short) model developed by the CEMAGREF and
(2) a model referred to as "HMS".

Only the GR component has been integrated in the operational PEGASE version. A model that calculates runoff given rainfall can use information about soil moisture if such data were available in several ways. Whereas the directly useful information is runoff, either model (GR, and HMS) calculates internally variable values that represent (or are supposed to represent) a certain state of moisture in the watershed.

Real-time operations

Especially when it appears that a relatively serious flood may occur, decisions about the timing and the magnitude of the releases must be made on short notice. To do a "perfect" job one would need: (1) a perfect forecast of the forthcoming rains, (2) a model which, given the forecasted rains, can calculate the future river flows with perfect accuracy throughout the system and finally (3) an effective scheme to optimize the releases simultaneously from all the reservoirs. Naturally future rains will not be forecasted with great accuracy for a long time horizon. This point is capital because propagation times from a reservoir to Paris can be of the order of 7 days (BCEOM, 95).
Whatever the quality of the forecast for precipitations, it is still necessary to convert these precipitations into runoff without adding another level of uncertainty about the future behaviour of the system but this point is not the objective of the AIMWATER project.
Models are not perfect as a result of: (1) their conceptual construction depicting in a simplified manner the physical processes taking place in the basin and (2) calibration of the model parameters on data that are sparse and sometime not very accurate themselves. Given all that information (calculated runoffs) one must select an optimum pattern of releases, a task, which can be arduous, as the objectives are not easily formulated in

quantitative terms or rather routine as, imposed by administrative rules ("réglements d'eau"). For better management of the system the Institution wants to improve on all aspects of the decision process, thus with a special regard to (2) model calculation of river flows and (3) conjunctive releases optimization. To accomplish these goals the Institution is receptive to acquisition of additional useful data to augment the capability to model modifications to make fuller use of the new information available and to suggestions for optimally managed reservoir releases. In a very general way these goals define the needs as well. More specifically the Institution has stated in a recent document (Grands Lacs de Seine, 1998) the following needs:

(a) « During the periods of low flow the current rainfall - runoff model does not seem to respond properly. We are looking for tools to improve our forecasting capabilities»

(b) «In order for the Institution to inform (users, authorities and the public at large) in the Paris region regarding the impact of its releases from the reservoirs, it is necessary to know the contributions of the intermediate basins»

(c) «Given current project under consideration, it is important to have forecasts for the floods in order to determine the opportune moment to fill or empty dedicated flood compartments »

- Case of ARBSLP (Associaçao de Regantes e Beneficiarios de Silves Lagoa e Portimao)

Historical Background

The Arade and Funcho dams are located in the Arade river basin, in the southern part of Portugal, near Silves and Portimao.

After the 30's there was a need to develop agriculture in some areas in Portugal where it was only possible by irrigation. A series of dams were designed and built including the Arade dam for this purpose. The Arade dam has been used to irrigate approximately 2000 ha, which produce mainly fruit trees (citrus) and corn and rice as annual crops.

In the winter time the Arade dam serves the purpose of flood control, protecting the village of Silves and the city of Portimao. With the constant increase in the tourist population since the late 70's, a new dam was built with the purpose of urban supply to the west part of the Algarve region, irrigation and flood control.

This dam (Funcho) was built upstream from the Arade dam in the Arade river basin. The Funcho dam was operational in 2000, although the urban supply infrastructures were still not operational at this time.

Current infrastructure situation

The dams are located in the Arade River upstream from the village of Silves (Funcho dam upstream from Arade dam).

What are some of the problems faced by the managers of the system? Shortage of water for irrigation, which in dry years is managed for use only for perennial crops, and the annual crops are eliminated for that year. Another problem relates to flood control. When rainfall events occur in mid-spring, they promote some of the biggest floods in the system.

Current model capability

No models are being used at the time. The system works based on past experience and common sense, with a set of rules for each case scenario.

Real time operations

All the operations made on real time, are based on previous experience, and good common sense. To do a better job one would need to have better rainfall forecasts, an appropriate rainfall-runoff model (which is not being used at the time), and define a set of procedures to be followed in several predefined situations.

- **Needs Regarding Soil Moisture Information**

- Case of IIBRBS

Retrieving soil moisture information using EO data

The Institution will need answers to a variety of questions regarding the information contributed by the EO. What will the radar observation measure exactly? An average value of moisture contained in the topsoil layer with a depth of about 10 cm? Will it provide several values for a catchment or a single value? What is the level of resolution of the observations? In what form will the information become available to the Institution? Which agency or company will be responsible to provide the information to the Institution as a customer of such information? How much will that information have to be processed by the Institution to deduce from it the information relevant to the models needed for the forecasting of river flows? How much of that information process will have to be incorporated in the simulation models?

- Case of ARBSLP

The needs regarding soil moisture information are not different from IIBRBS, since ARBSLP still doesn't have a model. Any additional information would be an improvement.

- **Needs regarding the assimilation of the soil moisture information into the operational practices.**

- Case of IIBRBS

Assimilation capacities of current modules of PEGASE

At present the models available to the Institution consider only lumped values for a given watershed. Either these models will be further distributed in the future or the observed data are averaged in some way to provide a single value for each watershed. Assuming the latter then the question is: what is the relation between the observations and the internally calculated values in the models GR or HMS. In GR there is no variable that represents explicitly an average soil water content in the top of soil, but the value S of reservoir A can be interpreted as a measure of storage in the soil. (A newer version of GR called GRHUM has integrated in the original structure the description of soil moisture in two soil layers). In HMS the variable "average soil water content in the top layer" appears explicitly in the structure of the model. It would then correspond directly to the observation as long as the thickness of the topsoil layer is selected to match the one that is sampled by the radar signals. Given that both models are to some degree conceptual (as are all models), one cannot hope to develop a perfect correspondence between model variables and the observed signal, but ideally a strong correlation between them.

Use of the EO information

PEGASE is run in a continuous simulation mode. On a day-to-day basis it calculates discharges at target points in the system and for all upstream or intermediate basins calculates " internal states" of the system. Let us assume then that a relation has been defined between the observed water content of the basin (OWC) and an internal variable (for either GR4 or HMS) thus a model calculated water content (MWC). Let us assume that an observation is available on a given day. For that day both OWC and MWC are available. If there were no errors in the model (conceptually, and input wise, etc.) and if there were no errors in the conversion from signal to water content and in the correlation between OWC and MWC, the two values would be the same. In practice the two values will differ. The question is thus: which one do we trust? If we trust the observed value without reservation, then one replaces the model value by the observed one and proceeds to make forecasts for the coming days using that updated initial condition in the model. Going one level further and presuming that the model is

"good" but that the inputs are in error (rainfall for example) one updates the past or current rainfall value in such a way that the model value will match perfectly the observed one. In reality we cannot trust either the observation or the model values completely and thus it will be necessary to use a weighted average value between them and use that value as the initial condition for future forecasts or as a basis for updating the "faulty" input. Such a procedure is known as a "filter". Formally it requires an estimation for both the "observation error" and the "model error" to estimate the values of the weights in the averaging process. Thus what the Institution will need is an assessment of both errors. Whereas that can be pursued formally it is more likely that a priori postulated weights will be used and tested on long term simulations with the filter and the weights adjusted based on the results of the simulations.
It is not clear at present whether PEGASE has the capability to operate in a "filter"mode. The existence of that "filter mode" capability in PEGASE needs to be ascertained. Given that, if soil moisture information derived from EO data is to be incorporated in the model for use in estimation of runoff and river flows, such capability is absolutely necessary.

- Case of ARBSLP

In the case of ARBLSP, since no model is being used at the present time, there is a need to choose an appropriate rainfall-runoff model that could easily integrate soil moisture information to be used their operational practices.

• Technical interpretation of user's needs

For better management of its system of reservoirs on the Seine river and its tributaries, the IBRBS wants to improve on all aspects of the process by which it decides on the releases to be made on a given day at a particular reservoir. Thus it wants to improve its tools at its disposal for: (1) forecast of rains, (2) calculation of river flows and (3) conjunctive releases optimization.
The AIMWATER project can provide partial answers to problems or needs defined by the just cited items (2), (3) but (1) is beyond the scope of the project.

The amount of runoff that will result from a given rainfall will depend greatly on the initial degree of saturation of the basin, or in other words, the amount of water is held in storage in the basin, predominantly in the soil. A model that calculates runoff given rainfall can use information about soil moisture in several ways if such data were available. In order to meet the users needs what can the AIMWATER project contribute? To answer that question a successful AIMWATER project would have to show that:

(1) it is possible by remote sensing to measure a signal that can be related to soil moisture in a particular stratum of the soil or to some overall watershed moisture index and
(2) that using a particular model (the one considered best, based on some criteria) it is possible to incorporate EO data in the rainfall-runoff process to improve the prediction of streamflow by deterministic simulation using observed rainfall. Currently the hydrologic component of PEGASE, available operationally to the Institution, to calculate daily runoff from daily precipitation is the GR model developed by the CEMAGREF. This model calculates internally variable values that represent a certain state of moisture in the watershed.

The challenges are thus to: (1) relate the EO signal into a useful soil moisture index for a watershed and (2) to incorporate (assimilate) that information into the GR or an appropriately modified GR model.

- Further dialogue with customer

A question raised by IIBRBS states: "How it is possible to integrate the parameter soils' water content in a rainfall-runoff model focused on real time applications?"
Another is: "Regarding the capabilities of assimilation of the current modules in PEGASE, will it be necessary to plan new developments, allowing these modules to function, optionally, in a distributed or a lumped mode?

Part of these questions was addressed based on the main assumption driven behind the AIMWATER project: a better knowledge of the state of the system (i.e. water storage over the various watersheds that constitute the

basin) will improve forecast of runoff and consequently improve operations of the reservoirs to meet the IIBRBS management objectives. Better choice of releases from reservoirs, specially for flood protection, requires a better knowledge of the state of moisture in the system but it also requires that future rainfall events be forecast with a certain level of accuracy and finally that the reservoir operations rules do specifically take the forecast of runoff into consideration. Recent studies have shown that spatial information derived from Earth Observation (EO) data could reliably provided spatialised and temporal soil moisture information. It will be shown later on how this information was used to answer user needs and expectations.

- Translation into needed tools development

There were two basic areas of investigation that were needed for the project. The first addresses the matter of extracting from an EO signal, that depends among other things, on the state of soil moisture, a useful measure of watershed wetness that can then be incorporated in a rainfall-runoff model. The second consists in searching for a relatively simple and robust model of the rainfall-runoff process that can utilize (assimilate) soil moisture information into its algorithmic structure. Scientifically both areas have been investigated satisfactorily as presented in the section below.

• Adequacy of Results to Customer Needs

This section is based on results presented in chapter 3 and 4, these results are analysed here from a user perspective considering not only the improvements but also the gaps or difficulties of possible implementation of the proposed procedures.

- Positive contributions

Relationship between signal and water content

Results of the study have shown that during the major part of the year, zones (sensitive targets) can be selected where it is possible to estimate soil water content in a top layer in spite of the presence of vegetation and ground roughness (Quesney et al, 2000). Use of such methodology makes it possible to secure a wetness index representative of the water content of the topsoil layer at the scale of the watershed. A linear relation is then empirically determined between the corrected radar signal (i.e. the signal that would have existed had there been no vegetation, a bare soil signal so to speak) and field measurements. The coefficient of correlation is high: 0.95.

It is possible to utilize other several "sensitive" vegetation covers to establish a water content characteristic of the watershed. Here the word "sensitive" refers to types of crops and times of year where the soil moisture contribution to the radar signal is not smothered by the influence of the vegetation.

Universal (quasi) slope

With the multi-selective method the involvement in the process of diverse land uses smoothes the variations in humidity due to soil drying from specific crops. As a result the slope in the linear relation between the radar signal and field measurements is in good agreement with that obtained on the Naizin basin in Britany (Cognard, 1996). The ERS/SAR sensibility, as measured by that slope, appears to be a transferable result to other agricultural basins. Application of this methodology to the other selected Seine river basins has shown also good agreement with that obtained.

However, this methodology has to be applied for every land use and every watershed in order to get a soil moisture indicator that can be assimilated. But considering the stability of the value of the slope in the relationship the assimilation could be implemented on the basis of a standard value without waiting for studies and field measurements in every sub-basin of the region.

Proof of robustness of models

This topic has been addressed in section 4.1. Let us repeat that extensive studies over long historical records and for a variety of basins located all over the world, have demonstrated the robustness of the models GR4, GRHUM and IHACRES that can be compared favourably with other models like TOPMODEL while maintaining a desirable level of simplicity.
When a rainfall-runoff model is coupled with an updating procedure, it is valuable to start with as an efficient and reliable model as possible in terms of streamflow simulation and forecast but also with a structure that allows the introduction of soil moisture information derived from EO data.

Successful modification for assimilation

Some of the selected models can integrate external information about soil moisture without changes in their structure but it is not always the case. Coupling the rainfall-runoff models with the updating procedure requires changing somewhat the structure in order to get a better correspondence between estimated values and observed values of soil moisture. For example when developing the Kalman filtering technique for GRHUM it was necessary to make some changes in the structure of the model. Such changes were implemented and the Kalman filtering technique was successfully tested.
When the parameter updating technique was applied with GR4 and GRHUM no change in structure is necessary. However for GR4 as with GRHUM, it is necessary to incorporate a first step, one that converts soil moisture as deduced from the radar signal to an internal variable of the model that is supposed to represent humidity in the watershed. These relations, external to the model, referred to as "constraint equation" (passage relations) are established prior to the updating process such relations were developed successfully.

Successful development of assimilation algorithms

Several assimilation techniques were tested and all have been successfully implemented, they all showed high persistence criterion, which is considered as the most severe one applied to short term forecasting method assessment. These techniques have been discussed in some details in previous sections. All have in common some filtering and updating capabilities.
Even though they started as separate examples of the techniques of "parameter updating" or "state updating", they tend to be hybrid combinations to some degree when using sequential and variational algorithms. That is especially true of the "parameter updating" technique, as was applied on the Seine basins.
The scaling method consists of using the DICASM model simulations (model calibrated against EO data) as wetness index input for the rainfall-runoff models in a forcing mode.

Assimilation of moisture alone

Certainly one would have to be rather clumsy to manage to do worse with more information! It is true though that if the information is of bad quality its use may lead to worse results. However the radar information is of good quality and the soil moisture information deduced from the radar signal is obtained with reasonable accuracy (4%), where there is relatively little dispersion about the regression line.
This information has to be converted further into the model equivalent internal variable through the proper "constraint equation" and some loss of accuracy does result. Use of this information on face value is unwise as is apparent from the results when used as if totally error free under the technique of blind "forcing".
The sequential procedure provides better improvement than the variational one when using scarce data. Of course the big question is: such improvement worth the extra cost and added complexity? The definite conclusion is that if soil moisture information is to be used it must be used with an updating mechanism.

Assimilation with moisture and discharge

As the aim of the assimilation procedure is to improve the estimation of runoff, one cannot use only soil moisture data: it is necessary to assimilate runoff and soil moisture jointly. In the sequential method, this is particularly true for short-term prediction and for medium to large basins: the influence of soil moisture is sensible in the production scheme, which is implemented before the unit hydrographs of the transfer scheme. The effects of soil moisture assimilation are thus slowed down by the model's time delay. However, the complication of the implementation of the Kalman filter must be kept in mind and can balance the improvements.

- *Gaps and/or Difficulties of Implementation*

Coping with vegetation effects

In order to retrieve the surface soil moisture, the total radar signal has to be corrected from the contribution of the vegetation. This effect can be calculated over crops for which measurements in the field have been secured (such as wheat) for use in a theoretical model that calculates the scattering by that particular vegetation and its attenuation of the ground signal. Tests have shown that the theoretical models provided reliable estimates of the vegetation influence on the measured signal and consequently made it possible to extract the soil moisture component from the overall measured radar signal with good accuracy. For other crops, such theoretical models are not yet available and need to be developed.

Dealing with dense vegetation

When the contribution of vegetation can be corrected, the relation between radar signal and moisture can be established when vegetation is not too dense. No such relation can be established for forests. When the percentage of forest is high, a global indicator of soil moisture can no longer be considered as representative of the whole basin. This implies that two separate moisture states must be calculated one for the forested zone, the other for the rest of the basin where the assimilation procedure can be carried out. To accommodate this situation the rainfall-runoff model should be run in a semi-distributed mode and one would have to test the effectiveness of the procedure for different percentage of forest.
Even on non-forested zones when the vegetation is very dense, it is not possible to extract a meaningful moisture content value from the measured radar signal. When there is no other sensitive crops sufficiently representative over the basin, a solution consists of correlating the time evolution of moisture probes in the field with the estimated moisture based on the signal and use that relation to deduce a value of the wetness index for the watershed during the dense vegetation period. This implies that a minimum number of moisture probes be implanted in the basin and followed through the year through measurements at regular intervals.

Calibrating the relation: radar signal versus water content

The difficulty results from the need to have observed field values of soil moisture to correlate with the corrected radar signal. In principle such measures should be made for each sub-basin in the region where soil moisture is retrieved. However it was shown that the slope of this relation was quite stable. Just as an empirical relation between the estimated catchment wetness index and conceptual moisture states in the model was developed, similarly one can calibrate the parameters in the relation. One advantage based on the results of the study is that the slope of the relation can be constrained within narrow limits since results have shown that it varied in a very narrow range. Similarly the intercept of the regression line can be varied within limits, based on current results. This procedure has the advantage to eliminate the need for numerous field measurements or, at least, to allow use of the approach without waiting for the completion of such an extensive campaign of measurements.

Rainfall forecast

All measures of performance, with or without assimilation, were carried out assuming that rainfall amounts in future days were known with certainty. This was necessary to evaluate the merit of an assimilation technique as compared to the situation that no information is available on the soil moisture state of the basin. This type of comparison is perfectly scientific and justified. Unfortunately in practice rainfall amounts for the future have to be estimated and a great uncertainty clouds this forecast. This was beyond the scope of the project but the full interest of the assimilation procedure will be justified when studies will be done using a rainfall forecasting technique.

Propagation in rivers and impact on reservoir operations

If the technique of assimilation is to be used by IIBRBS, it will have to be tested on an ensemble of basins with long propagation in the rivers of the lateral flow input as predicted by the basin models.
Also, before any such new procedures have a chance to be adopted it will be necessary first to design rules of operations that will explicitly incorporate in them the forecast by the model and then to show that such rules of operations will lead to better reservoir operations. The improved forecasting by a model will not be useful if the rules of reservoir operations do not include the prediction coming from the model. Currently as practiced by IIBRBS such model results are not directly used explicitly in the PEGASE model, which describes the response of the river system and the operation of the reservoir.

• Conclusions

The project has shown that it is possible to derive from EO radar backscattered signals a watershed soil moisture index. However there remains some uncertainty on how to proceed when there is a significant percentage of forests in the basin or when the vegetation is at its peak. Continuous ground soil moisture monitoring may be necessary to fill the gaps when EO data cannot be utilised.

The project has shown also that a filtering procedure to predict runoff will improve the forecast over the direct use of the EO derived moisture index and certainly over simply not using the EO data at all. However, the results of the EO data assimilation technique showed little improvements because of the scarcity of data.

Finally the improvement was assessed using a deterministic future i.e. a future when rainfall values are known with certainty. In real time operations the values of the future rainfalls are very uncertain and it remains an open question whether the improvement of using EO data will not be lost in the noise brought about by the wide uncertainty in estimating future rainfalls.

Some suggestions are provided in Chapter 5 for possible implementation of these procedures in an operational context.

References

BCEOM, 1994. "Etude d'Optimisation de la Gestion Coordonnée des Barrages Réservoirs du Bassin de la Seine", Rapport Final, tome 3, "Module de Simulation Hydraulique-Structure et Bases Théoriques", June 1994, 94 pages.

BCEOM, 1995."Interconnexions Grands Lacs de Seine", Rapport Final, December 1995, Annexes,pages 91-98.

Bezdek, J.C., Ehrlich, R., and Full, W., 1984, "FCM: the fuzzy c-means algorithm", *Computer and Geoscience*, 10(2-3):191_203.

Cognard, A.L., Loumagne, C., Normand, M., Olivier, P., Ottlé, C., Vidal-Madjar, D., Louahala, S., Vidal, A., 1995, "Evaluation of the ERS-1/synthetic aperture radar capacity to estimate surface soil moisture: Two-year results over the Naizin watershed", *Water Resources Research*, 31(4):975-982.

Cognard, AL. (1996), Suivi de l'état hydrique des sols par télédétection spatiale (radar et infrarouge thermique) et modélisation hydrologique à l'échelle du bassin versant. thèse de l'Univ., Paris XI, 150 pg+annexes.

FAO-UNESCO, 2001. Soil Map of the World. Report No. Volume V. UNESCO, Rome, 199 pp.

Geman, S., and Geman, D., 1984, "Stochastic relaxation, Gibbs distribution and Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721-741.

Jennings R, 1995. Special Edition: Using Access 95. Que Corporation, Indianapolis, USA, 1290 pp.

Karam, M.A, Fung, A.K., Lang, R.H., Chauhan, N.S., 1992, "A microwave scattering model for layered vegetation ", *IEEE Transactions on Geosciences and Remote Sensing*, 30(4):767-784.

Les Grands Lacs de Seine, 1998. Presentation de la Gestion des Ouvrages de Grands Lacs de Seine au CEMAGREF , Journée du 17 Novembre 1998, 14 pages.

Le Hégarat-Mascle, S., Bloch, I., Vidal-Madjar, D., 1997, "Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing", *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):1018-1031.

Le Hégarat-Mascle, S., Quesney, A., Vidal-Madjar, D., Taconet, O., Normand, M., Loumagne, C., 2000a, "Land cover discrimination from multitemporal ERS images and multispectral LANDSAT images: a study case in an agricultural area in France", *International Journal of Remote Sensing*, 21(3):435-456.

Le Hégarat-Mascle, S., Alem, F., Quesney, A., Normand, M., Loumagne, C., 2000b, "Estimation of watershed soil moisture index from ERS/SAR data", *Proceedings of EUSAR2000*, in Munich, Germany, on May 23-25, 2000, pp.679-682.

Quesney, A., Le Hégarat-Mascle, S., Taconet, O., Vidal-Madjar, D., Wigneron, J.P., Loumagne, C., Normand, M., 2000, "Estimation of watershed soil moisture index from ERS/SAR data", *Remote Sensing of Environment*, 72(3):290-303.

Taconet, O., Vidal-Madjar, D., Emblanch, C., Normand, M., 1996, "Taking into account vegetation effects to estimate soil moisture from C-band radar measurements", *Remote Sensing of Environment*, 56:52-56.

3. DERIVATION OF HYDRIC INDICATORS FROM EO DATA

Two main remote sensing techniques are described in this chapter, the two more commonly being used in hydrological applications: microwave SAR data and optical data (visible/infrared). Other types of data (like active optical sensors, LIDAR) are not described here because there are not yet operational systems in space making use of such capabilities.
The applicability of remote sensing techniques in the context of the AIMWATER project is conditioned by the specific characteristics of the two main study areas of the project: the Seine basin with a humid temperate climate and the Arade basin with a semi-arid climate. Over these two basins, two methodologies have been set up and will be described thoroughly along with the results obtained in validation.

3.1 EO signal analysis methods

The aim of this section is to introduce the potential capabilities and applications of satellite remote sensing data in hydrological applications, especially those dealing with the topics more relevant for the AIMWATER project, soil moisture being the driver parameter. A review of the existing remote sensing data along with the retrieval techniques for soil moisture will be provided with special attention to the techniques used in the context of the project.

• Requirements for hydrological models

The requirements for hydrological models in terms of remote sensing data depend on the category of models and on their application domain. For hydrological applications, Kite and Pietroniro (1996) show how the development of physically based distributed models has increased the demand for spatial data. An improved understanding of the hydrological cycle requires measurements of time series of data at a point (a), time series which vary over an area (b), and data which do not change over the time scale of modelling period (c). The same authors give a list of data types requirements for hydrological models and classify them by dimension and time-dependency: Meteorological/climatological data (a,b,c), vegetation/land cover (a,b,c), physiography (a), soils (a, b). They provide also some examples of parameters used in hydrological modelling which have been derived from satellite data and provide sample references. Some details are given about estimation or measurement of precipitation using visible or infrared satellites, of soil moisture using SAR data, of snow cover extent with sensors in the visible and the infrared and snow water equivalent using passive or active microwaves, of leaf area index and vegetation index using visible and infrared sensors, and of evapotranspiration using infrared sensors. Kite and Pietroniro (1996), in their conclusion, notice however that, in spite of the spatially distributed character of remote sensing data, they are difficult to use in hydrological models, and that research is needed in the development of generalized algorithms and into the design of hydrological models more suited to the routine use of remotely sensed data.
Engman (1996), making the distinction between « applied hydrology » and « scientific hydrology », states that the latest needs better and different data to answer the detailed questions asked about the intermediate stages of the hydrological cycle. Two issues that traditional hydrological instrumentation cannot address are the spatial variability of hydrological processes and the wide disparity of space and time scale that scientific hydrology must address. Remote sensing can address the spatial heterogeneity and the scale disparity problems. The same author discusses about four characteristics of remote sensing data that make them a very powerful tool for advancing hydrological sciences. The first one is the possibility of measuring system states using thermal infrared and microwave remote sensing to infer surface properties important to hydrology such as surface temperature, soil moisture and snow water content. Secondly, remote sensing provides area data (instead of point data in the case of in situ measurements) and may help to understand scaling and scale interdependence in hydrological systems. The third one is the possibility of acquiring temporal series data used to monitor various hydrological states over very large areas as well as monitor the dynamic properties in hydrology. The last one is related to the new data forms provided by remote sensing, that can be combined with other spatial data (such as soil maps) and even point data

through a data assimilation scheme or Geographic Information System (GIS). Engman (1996) states that remote sensing data can be very useful for hydrological modelling, but underlines also that there is a need to develop new concepts and to change the historical way of conceptualising hydrological processes.

• Existing remote sensing data

- *Microwaves*

Passive microwaves

Multifrequency and multipolarization radiometers have operated in space for many years. The Scanning multichannel Microwave radiometer (SMMR) and the Special Sensor Microwave Imager (SSMI) were designed for ocean-ice-atmosphere sensing, and operate at high frequencies: 6.6 and 37 GHz for SMMR; 19.3 and 85.5 GHz for SSMI. Moreover they have very low ground spatial resolutions: 27 and 150 km for SMMR, 15 and 70 km for SSMI.

Considering the question of soil moisture sensing, the use of a low frequency, e.g. 1.4 GHz (L band), is unanimously considered as the most suitable, both because of the reduced atmospheric attenuation and because of the greater penetration inside the vegetation cover. Indeed, roughness and vegetation may be considered as perturbing effects on soil moisture retrieval from passive microwave measurements. However, for wavelengths superior to 10 cm, the effects of vegetation and roughness are much reduced.

Until now, there has not been any low-frequency radiometer system operating continuously in space able to perform large area repetitive measurements at seasonal or annual time scale. Recently, renewed efforts have been made to gain supports for the development of a low-frequency space borne passive microwave soil moisture sensor. The studies have focused mainly on single frequency systems operating at 1.4 GHz, horizontal polarization, and providing approximately a 10 km spatial resolution (such good resolution was achievable thanks to the development of interferometric techniques) with a global mapping capability every 3 days. This is, for example, the configuration of the Electronically Scanned Thinned Array Radiometer (ESTAR) (le Vine *et al.*, 1989; Swift, 1993). ESTAR is the aircraft radiometer simulator of the HYDROSTAR project, which was submitted to the NASA. In the same way, the development of a new sensor, RAMSES- Radiométrie Appliquée à la Mesure de la Salinité et de l'Eau du Sol (Kerr *et al.*, 1997) renamed SMOS- Soil Moisture and Ocean Salinity (Kerr *et al.*, 1998), has been proposed by the European scientific community to the A.O. ESA project. It was accepted for phase A studies, and the simulator (MIRAS) has already flown. The main advantage of SMOS is that it will be a multipolarization (H and V), multiangle, and multifrequency (L and C band) instrument. The interest of such multichannel information is that it enables the retrieval of the vegetation and/or roughness effect, and therefore the correction the signal (Njoku and Entekhabi, 1996).

Active microwaves

In the past 10 to 20 years, several ways to derive soil moisture indices from airborne or spaceborne active measurements have been proposed (see for example Dubois *et al.*, 1995; Wang *et al.*, 1997). In the early years, Ulaby *et al.* (1986) have demonstrated that an optimized instrument could be defined for soil moisture estimation: operating in C band and at about 10° incident angle. At such incidence angles the soil backscatter signal does not depend on soil roughness state, but only on soil moisture. If the C band frequency requirement is easy to fulfill, the low incident angle is not possible to perform if high ground resolution has to be achieved at the same time. The two spaceborne SAR systems, which are currently in operation, ERS/SAR and RadarSAT, are not able to acquire images at smaller incidence angles than 20° (at least in standard mode). This will be also the case for the EnviSAT SAR, scheduled to be launched in 2002.

The characteristics of the two operational systems ERS and RadarSAT are the following:

The ERS (European Remote Sensing Satellites) Synthetic Aperture Radar is a C band (wavelength ≈ 5.6 cm) and VV polarization system. Its incidence angle is equal to 23° in the middle of the swath. In the case of 4-look images, the pixel resolutions in distance and in azimuth are both equal to about 12.5 m. ERS whole orbital cycle is equal to 35 days. However, combining different tracks and ascending and descending modes, some regions can be imaged up to 4 times within 35 days.

The RadarSAT (The Canadian) Synthetic Aperture Radar is also a C band system. The images are acquired in the HH polarization, and the satellite has an orbital cycle equal to 24 days. The main advantage of RadarSAT is that its incidence angle may vary. In the standard mode, RadarSAT incidence angle range is 20-49°. The 4-look pixel

resolution is about 25 m. RadarSAT may also operate at other modes corresponding either to different pixel resolutions (from about 10 m to 100 m) and swath widths (hence temporal resolution), or to an extension of the incidence angle range (extended modes propose either 50-60° or 10-20° for incidence angle).

How to make a choice?

A fundamental difference between active and passive microwave systems lies in their spatial and temporal resolutions. In the case of passive sensors, the spatial resolutions are about 100 km, and temporal resolution achieves only few days. Interferometric systems achieve spatial resolutions in the 10-20 km range. In the case of active sensors, the spatial resolution is much better: about 10-20 m for SAR systems, but temporal resolution is very poor: e.g. 35 days for ERS. Clearly, a compromise has to be found between spatial and temporal resolutions. This is the reason why the recent and future systems, such as RadarSAT and EnviSAT, are defined to operate in several modes with different resolutions.
Another noticeable difference, until there is no L band operational SAR (at the present time, we do not consider JERS as an operational system since data are still not available), is the sensing depth of the wave in the surface. In the case of bare soil, it remains inferior to few centimeters in both cases (L band passive or C band active). However, in the case of a vegetated area, a higher wavelength such as L band allows to get information about soil surface whereas C band systems do not.
Finally, even if their wavelength domain is the same (microwave), passive and active sensors have highly different characteristics, which makes, as we will see in the next sections, the soil moisture retrieval techniques and performance also highly different, and therefore synergism between these two techniques appears very promising.

- Optical

Currently operational optical sensors that can be actually used in the context of the AIMWATER project mainly include those on board the LANDSAT and SPOT satellites. Although there are other systems in operation, for some reasons it is not fully feasible to use these data in an operational way, so that to avoid any potential conflict we will only focus on LANDSAT and SPOT data. Sensors having very low resolution but frequent (daily) coverage, such as the Advanced Very High Resolution Radiometers (AVHRR) on board the NOAA series of satellites, quite useful for other applications, are not described here because they are not useful in the context of AIMWATER.

Optical sensors in currently operational satellites

The SPOT satellites describe a near-circular, near-polar, sun-synchronous orbit, having a temporal resolution of 26 days. The most relevant characteristic of the SPOT family is its high spatial resolution: 20 m in the multispectral mode and 10 in the panchromatic.
The first three SPOT satellites payload includes two HRV (High Resolution Visible) sensors whereas SPOT-4 carries two HRVIR (Haute Resolution Visible InfraRouge). Operating independently of each other, the two sensors acquire images in either mustispectral and/or panchromatic modes at any viewing angle within plus or minus 27 degrees. This off-nadir viewing enables the acquisition of stereoscopic imagery.
The SPOT-5 launch is planned for 2000, equipped with the HRG High Resolution Geometry instrument. Spectral bands will be the same as those for SPOT-4, but the panchromatic band will return to the values used for SPOT 1-3. The resolution will be improved: in the panchromatic mode it will go from 10 m to 5 m /3 m and in the multispectral mode, B4 (intermediate infrared) remains identical, but in the three spectral bands in the visible and near infrared it will go from 20m to 10 m.

SPOT HRVIR/HRV Characteristics

Mode	Multispectral	Panchromatic	Multispectral	Panchromatic
Resolution	20 m at nadir	10 m at nadir	20 m at nadir	10 m at nadir
Swath	60 Km at nadir	60 Km at nadir	60 Km at nadir	60 Km at nadir
Bands (μm)	0.50 - 0.59 0.61 - 0.68 0.79 - 0.89 1.58 - 1.75	0.61 - 0.68	0.50 - 0.59 0.61 - 0.68 0.79 - 0.89	0.51 – 0.73

The Landsat satellites series started in 1972, with the launch of Landsat-1, originally called ERTS1. All Landsat satellites describe a polar, sun-synchronous orbit. The Landsat satellites payload consisted on the RBV (Return Beam Vidicon) and the MSS (Multispectral Scanner) sensors till Landsat-4, which were the first ones equipped with the TM (Thematic Mapper) sensor. The latest Landsat satellite, Landsat 7 has being launched in 1999, carrying the ETM+ (Enhanced Thematic Mapper) sensor. The temporal resolution of every actual Landsat satellite is 16 days, but working together they can offer repeat coverage at any ground location every eight days.

TM / ETM+ Characteristics

Temporal Resolution	16 days	16 days
Spectral range	0.45 - 12.5 μm	0.45 - 12.5 μm
Number of Bands	7	8
Swath	185 Km at nadir	183 Km at nadir
Spatial Resolution	Band 6: 120 m at nadir Other Bands:30 m at nadir	Band 6: 60 m at nadir Band Pan: 15 m at nadir Other Bands: 30 m at nadir
Size of Image	185 Km x 172 Km	183 Km x 172 Km

TM / ETM+ Bands description

Band	Bandwidth (μm)		Application
	TM	ETM+	
1	0.45 – 0.53	0.45 – 0.515	Coastal water mapping, soil/vegetation differentiation, decicuous/coniferous differentiation, chlorophyll absorption
2	0.52 – 0.60	0.525 – 0.605	Green reflectance, peak of healthy vegetation, plant vigor
3	0.63 – 0.69	0.63 – 0.690	Clorophyll absorption, plant type discrimination
4	0.79 – 0.90	0.75 – 0.90	Biomass surveys, water body delineation
5	1.55 – 1.75	1.55 – 1.75	Vegetation moisture measurement, snow/cloud differentiation
6	10.40 – 12.50	10.40 – 12.5	Plant heat stress, thermal mapping, soil mapping
7	2.08 – 2.35	2.09 – 2.35	Hydrothermal mapping, geology
Panchromatic		0.52 – 0.90	Mapping, area identification

• Retrieval techniques

The main aspects of the retrieval of surface parameters needed in the context of AIMWATER are described here. Only those aspects relevant to AIMWATER will be considered, although the description is kept general enough to serve as a tutorial on the applicability of remote sensing techniques in such applications.

- Microwaves

Microwave remote sensing techniques (passive or active) provide a privileged way to monitor soil moisture from space, since:

- The measured signal is sensitive to the dielectric constant of the scattering or backscattering targets of the observed surface, and, in the case of the emerged surfaces, this dielectric constant is mainly dependent of the water content (Ulaby *et al.*, 1986);
- The measurements are mainly unaffected by cloud cover and variable surface solar illumination (which is not the case with optical systems).

In soil moisture retrieval, the methodology used is fundamentally different in the case of passive and active microwave: it is based on the signal modeling in the first case, and on an empirical (or semi-empirical) approach in the second one. This difference in strategy is highly due to our capacity to simply model the scattering or backscattering process, and to the availability of numerous data.

Passive microwaves

In low frequency passive microwave remote sensing, the measured brightness temperature T_B is mainly a function of soil moisture (Wang and Choudhury, 1995). However, T_B is also affected by soil surface roughness (Choudhury *et al.*, 1979; Tsang and Newton, 1982; Mo *et al.*, 1987), attenuation and emission by vegetation cover (Jackson *et al.*, 1982; Ulaby *et al.*, 1983; Pampaloni and Paloscia, 1986; Jackson and Schmugge, 1991). Then, in order to maximize the sensitivity of the microwave measure to soil moisture, these perturbing effects have to be modelled, and the signal has to be corrected from their contribution (Njoku and Entekhabi, 1996).

♦ Case of homogeneous area

The emission of thermal microwave radiation from soils is dependent on the soil water content through the effect of the dielectric constant ε on the measured signal. The brightness temperature T_B is related to emissivity e ($T_B = e.T$, where T is the true temperature), e is related to the reflectivity r ($e = 1-r$), and r depends on the dielectric constant ε and the incidence θ. Ignoring the dependence on the temperature (provided that water remains in liquid phase), ε is mainly a function of the soil type, of its volumetric water content, and of the frequency which defines the "penetration depth" in the observed medium (at 1.5 GHz, the penetration depth varies between 10 cm and 1 m from saturated to dry soil).
In the case of rough surfaces, semi-empirical expressions have also been proposed (Wang *et al.*, 1983) for the expression of reflectivity, such as

$$r_p' = [Q.r_p + (1-Q).r_q] \times \exp\{-h'\}$$

where the subscripts p and q stand for H and V, and r_H and r_V are the Fresnel theoretical reflection factors. The parameters Q and h' depend on the frequency, the height standard deviation, and the viewing angle (e.g. h' may be written $h' = h.\cos^2\theta$), and have to be determined experimentally.
In simple radiative transfer models, the vegetation is modeled as a single homogeneous layer above soil, and interactions with the atmosphere are neglected. Then, the observed brightness temperature T_B depends on the soil reflectivity r_p', the vegetation transmissivity γ, the simple diffusion albedo of vegetation ω, the vegetation temperature T_V, the soil equivalent temperature T_e, and the descending atmosphere brightness temperature T_a. For more simplification, at low frequencies, the atmosphere contribution may be neglected ($\omega \approx 0$), and it is assumed $T_V = T_e$. The vegetation transmissivity can be written: $\gamma = \exp\{-\tau/\cos(\theta)\}$ where θ is the incidence angle and τ is the vegetation opacity. τ depends on the vegetation type (geometrical structure of leaves and woody components), the water content of the vegetation (which affects its dielectric properties), and the frequency ν. In (Jackson and Schmugge, 1991), a linear relation between τ and W is fitted experimentally: $\tau = b.W$, where b is an empirical parameter, whose dependence on the kind of vegetation regarding frequency has been often studied (Mo *et al.*, 1982, Wigneron *et al.*, 1995, Haboudane *et al.*, 1996).
Therefore, the measured brightness temperature is not only a function of the soil moisture but also of the soil texture, surface roughness, soil and vegetation temperature, and vegetation type and water content. Thus, in the process of retrieving soil moisture from passive microwave remote sensing data, corrections are needed when these additional factors introduce significant errors. The basic approach is to make corrections sequentially for each of the mentioned effects (e.g., Wang *et al.*, 1989; Jackson, 1993; Jackson and le Vine, 1996). Several ancillary data are needed:

- A land cover database or other *a priori* information is used to determined the broad classification of the surface type, e.g. forest, grassland, wheat crop… b parameter is deduced;
- The water content W needed to compute the vegetation opacity τ is estimated either from optical or infrared remotely sensed index, or from multipolarization microwave index (Pampaloni and Paloscia, 1986; Wigneron, 1993) if the microwave sensor is multipolarized;
- The surface temperature is estimated either from a simultaneous thermal infrared radiometer measurement or by extrapolation from a local surface air temperature measurement. If T_e and T_V are not assumed to be equal, the knowledge of their difference is required;
- If Q and h' parameters have to be estimated, additional ground truth measurements are needed to calibrate roughness effect; otherwise, in the absence of information, and for data at 1.4 GHz, Q may be set equal to zero, and h' may be chosen between 0 and 0.3 (Jackson, 1993);

Knowing all the parameters involved, we can deduce the set of the values of the soil dielectric constant ε. It is then possible to simulate the observed brightness temperature T_B. Finally, soil moistures corresponding to these dielectric constants are the solutions.
This approach is only feasible if the opacity τ is sufficiently small so that the equation can be inverted without unacceptable error amplification. In other words, assuming a fixed error bar on brightness temperature data, if τ is too large, the set of solutions for ε or soil moisture will be equal to its whole range of values. Thus, in such an approach, the quantification of the influence of parameter estimation errors as well as measured temperature imprecision is a fundamental step for a valid interpretation of the results.

♦ Soil and vegetation heterogeneity effects

The poor resolution (several kilometers) of the measured signal raises the problem of surface heterogeneity and in particular of the variability of moisture and temperature of the soil and vegetation. As the combination of these parameters in the computation of the emitted microwave radiation is non-linear, it is not correct to use averaged parameters in the previous equations.
In the direct modeling, the solution is rather to consider sub-regions A_i with homogeneous soil and vegetation parameters, to compute (according to the previous equations) the corresponding brightness temperature T_{Bi}, and to estimate the observed brightness temperature as a weighted sum of all the: $T_B = \sum \alpha_i T_{Bi}$. In the simplest approach, the weight coefficients α_i may correspond to the fractional spatial extension of the areas within the pixel; they may also take into account the effect of antenna pattern.
In the soil moisture inversion, the strategy is basically unchanged, except that now, in order to simulate the brightness temperature, the model parameters (b, α, T_e, T_v Q and h') have to be estimated for each of these homogeneous sub-regions.

Active microwaves

Contrary to the case of passive microwave techniques (to retrieve soil moisture), which were based on theory but, nowadays, are not really operational because of the lack of spatial data, active microwave techniques are mainly empirical and their performance when using satellite data have been recently shown (e.g. Quesney *et al.*, 2000).
At incidence angles superior to the optimal 10° (i.e. in the case of the operational SAR systems), the radar signal backscattered from vegetated areas can be divided into 3 components: (i) a signal from the soil surface volume, which is mainly dependent on the soil moisture content, (ii) a signal from the soil surface which is driven by the soil roughness, and (iii) a signal from the vegetation canopy overlying the soil. Moreover, the presence of a vegetation cover induces an additional attenuation of the soil-backscattered signal. Then, to estimate soil moisture, it is necessary to find a way to cope with soil roughness and vegetation effects (as in the case of the passive microwave approach). In the following sections, we first briefly recall the main effects of soil and vegetation features on the measured SAR signal, and then we present some of the empirical but operational SAR signal processing to retrieve soil moisture in agricultural watershed.
A major advantage of techniques using active rather than passive microwave is the high resolution of the SAR systems. Indeed, it allows to consider homogeneous targets and then to solve the problem of soil and vegetation heterogeneity. Most of the approaches are empirical in the sense that some of the parameters used in the algorithms are calibrated against ground data. However, they are generally well supported by theoretical model simulations.

♦ Retrieval of soil moisture from multichannel systems

Several ways to retrieve soil moisture from SAR data have been proposed and tested. They are based on the use of multiparameter radar: (i) two or more incidence angles (e.g. Autret *et al.*, 1989, studied the possibility to derive soil roughness in the case of bare soils), (ii) diversity in frequency and/or in polarization (Oh *et al.*, 1994). Indeed, the use of multichannel data allows the separation of the vegetation and soil signal contributions, and the different effects of roughness and moisture in the soil signal component. For example, Oh *et al.* (1994) show that the rms height and the dielectric constant of soil surface can be retrieved from polarimetric data using the co-and cross-polarized ratios, since they both have an empirical formulation which depends only on these two parameters (e.g. they are about independent of the correlation length).
Most of these methods have been implemented and tested using the SIR-C/X-SAR mission data (Evans *et al.*, 1997; Kasischke *et al.*, 1997), which offered the possibility to deal with a full multiparameter SAR. Although some success has been obtained (Schmullius and Evans, 1997), the main drawback of the proposed methods is

that they require multiple imagery. At the present time, all these images cannot be obtained simultaneously with a single (operational) satellite, and delays between two data acquisitions are generally not compatible with soil moisture or plant density evolution. Therefore, it clearly appears that methods allowing the derivation of soil moisture content using simple radar such as the one on board of the ERS 1&2 satellites still present an high interest.

♦ Retrieval of soil moisture from SAR/ERS system

Numerous studies showed that soil moisture index could be retrieved from ERS 1&2 SAR measurements (Beaudoin *et al.*, 1990; Cognard *et al.*, 1995; Taconet *et al.*, 1996; Griffiths and Wooding, 1996; Moran *et al.*, 1998; Quesney *et al.*, 2000). The main problems encountered for a robust relation between SAR signal and soil moisture is to overcome vegetation and roughness effects.
Regarding vegetation, the simplest approach is to only consider bare soil pixels in the ERS image. This is possible thanks to the high resolution of this sensor, which allows to select sub-fields areas. Some most sophisticated techniques propose to consider targets such as the wheat fields (in the case of agricultural watershed) where vegetation effect modeling is possible either using empirical rules (e.g. Taconet *et al.*, 1996), or a simple radiative transfer model. In this last case, the evolution of the plant features (geometry and biomass) during the vegetation cycle requires *a priori* regular ground truth measurements. Fortunately, even if in most studies ground truth measurements on plant parameters are used, the information which is really necessary to run the microwave/plant interaction model is sufficiently crude (mainly crop density and height) to be accessible from *a priori* knowledge on plant phenology and local agricultural practices and from visible space remote sensing throw the use of classical vegetation indices. The canopy contribution being quantified, the global radar response can be corrected, at least in the case of sparse vegetation, to retrieve the "bare" soil contribution.
Regarding soil roughness, in the case of a single SAR measurement, a solution is to assume constant roughness effect. However, due to agricultural practices, it is not possible at field scale. In (Quesney *et al.*, 2000), the proposed approach was to consider sufficiently large scale so that this assumption is true. The efficiency of this approach was verified empirically for small agricultural watersheds (about 100 km^2) presenting a sufficient number of different culture types with different soil practice periods.

- *Optical*

Retrieval techniques based on optical data can be summarized in three different categories:

Those based on classification techniques:

The basis are merely statistical, and there are no essential distinctions between optical and microwave data. Apart from the classical supervised and unsupervised methods (maximum likelihood, clustering techniques, etc.) well known in statistical data analysis in other fields, and apart from the advances in non-parametric methods with increased computer capabilities, new techniques (like neural networks, fuzzy analysis, intelligent data fusion, etc.) are being used for the same classification purpose. The degree of success depends essentially on the characteristics of the study area, being the spatial resolution a critical aspect to avoid confusion in the resulting classification. High-resolution systems (like SPOT or LANDSAT) provide the best results.

Although in the case of SAR data, the single channel/single polarization available in current space systems makes necessary the use of temporal series of data as input "channels" in the classification procedure (one "channel" per each available date), in the case of optical data the different channels used in the classification procedure are the multi-spectral information (several spectral channels in each single image). In current system such as Landsat (7 channels) or SPOT (4 channels) such number is still limited, but airborne sensors already existing provide up to more than 200 spectral channel per single image, thus increasing tremendously the classification possibilities.
Special mention requires the potential use of spatial information (texture, context) in the classification procedure, a key aspect in taking advantage of using remote sensing data.
Methods exist to handle the spatial data analysis, but the techniques are still mainly in the research domain, and the operational application is still not feasible.
In the particular context of AIMWATER, classification techniques are not expected to play a central role, but will constitute essential help in interpretation of SAR data and in the stratification of optical images previous to surface parameters retrievals.

Those based on empirical relationships ("indices")

Vegetation indices are quantitative measurements related to the vigor of vegetation (Bannari, 1995). The indices theory is based on the combination of a small number of bands and their main purpose is to enhance the information contained in remote sensing data that is related to amount health and extent of vegetation. Although several bands have been used, the most common bands used are the Red (R) (630 -690 nm) due to the strong chlorophyll absorption and the high near infrared (NIR) reflectivity (760 - 900nm) reflected by leaf cellular structure. The ultimate goal of the vegetation indices is to enhance the information contained in the data and extract the variability related to vegetation properties, minimizing the soil, atmosphere and sun and view angle effects.

♦The soil line and the vegetation Indices
The soil line is the soil brightness vector. It represents the soils in the spectral space. Some of the factors that affect the soil line are: shadow, organic matter, and moisture. The soil line is used as a base of reference line in studies of vegetated areas (Rondeaux et al 1996). It can be affected by atmosphere and sensor properties. It is soil specific and scene-dependent.
The slope of the line is mainly expressing soil brightness. The width is expressing other properties more inherent to the soil. Some vegetation indices are based on the soil line concept. If we consider a group of pixels with variable amount of vegetation and soil we are interested in maximizing the variability between soil and vegetation. The information about the amount of vegetation is given by the isolines of equal greenness. They measure the distance of a pixel to the soil line.
Examples:
Normalized Difference Vegetation Index, Rouse et al (1974) $NDVI = \frac{NIR - R}{NIR + R}$

Soil Adjusted Vegetation Index, Huete (1988) $SAVI = \frac{NIR - R}{NIR + R + L}(1 + L)$
L ≡ soil adjustment factor = 0.5
Assumption: Linear relationships between NIR and R reflectances from bare soils. Most soil spectra follow the same soil line. N-dimensional orthogonal distance between a pixel and the soil line. Greenness isolines are parallel to the soil line. Better that the soil line at low vegetation cover.

♦ Atmospherically corrected Indices
Intrinsic or soil line related indices. They introduce the atmospheric information contained in the blue channel to reduce aerosol contamination. Interesting to minimize residual atmospheric variations on a pixel-by-pixel basis. Example:

Soil and Atmospherically resistant vegetation index Kauffman and Tanre 1992 $SARVI = \frac{NIR - Rb}{NIR + Rb + L}(1 + L)$
Where: Rb= R - f (B-R) and f depends on the aerosol type.
Minimizes both: soil and atmospheric effects, not adequate for vegetation at low amounts.

Those based on numerical scattering model inversion

The essential idea is to have a physical model describing the transport of radiation through the atmosphere/canopy/soil media, as a function of driving parameters. In our case, since water is the major parameter of interest in the context of AIMWATER, the critical modeling consists in describing water absorption along the spectral interval of interest. This spectral absorption behavior of water is described in figure1

Known the structural geometry of the soil/canopy (that can be described as turbium media), the standard radiative transfer theory can by used to model the signal that would be measured by an optical sensor in a given satellite (direct modeling). The inverse modeling consists in solving the inverse problem; that is, given the signal measured by the satellite to determine the soil/canopy/vegetation parameters.
The following figure shows an example of the overall procedure. The black line correspond to the remote sensing data (radiance values measured by the sensor) and the blue line correspond to the "best-fit" (numerical optimization) of a multiple scattering radiative transfer code describing the atmosphere/canopy/soil coupled system behavior.

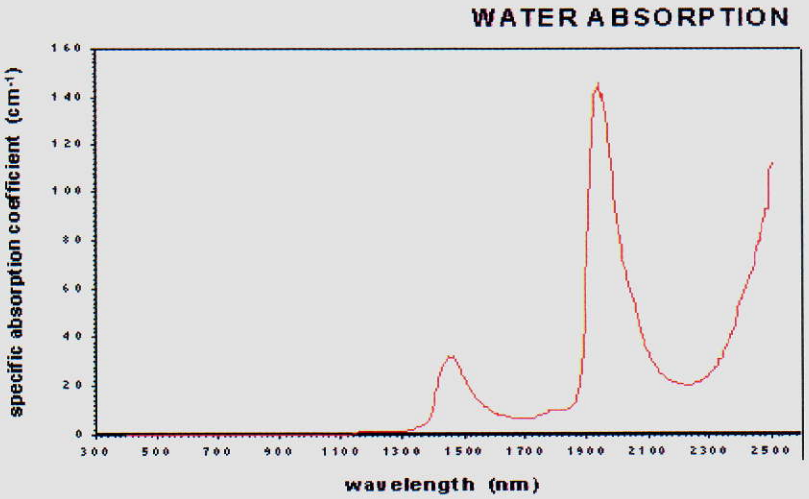


Figure 1: spectral absorption behavior of water

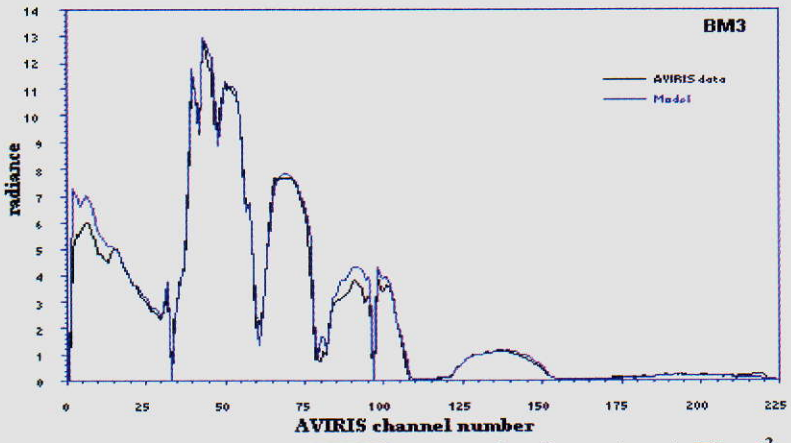


Figure 2: Comparison between AVIRIS measured radiance data (mW cm^{-2} μm^{-1} sr^{-1}) and model fit, for the cornfield BM3 in the area of Barrax, Spain.

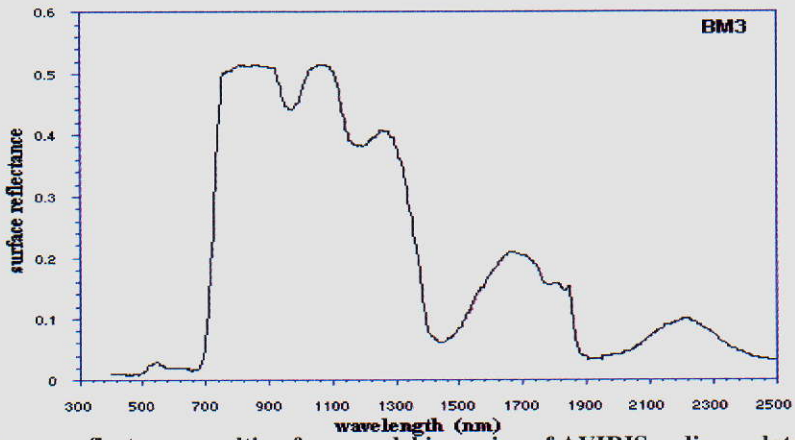


Figure 3: Canopy reflectance resulting from model inversion of AVIRIS radiance data after compensation from atmospheric effects

The inversion of the theoretical model against the measured data allows the derivation of all the relevant surface parameters, by means of the reconstruction of the actual surface reflectance at each wavelength (figure 3).
In this way, all the driving parameters in the model can be derived from the data, provided that the number of measurements (spectral channels) is larger than the number of parameters to determine, and also provide some estimates about the type of surface and overall geometric characteristics than can be determined by ground sampling.
This technique is of course the most sophisticate and the most computational time consuming, but it is also the one providing the best results. Note, however, that not always the most sophisticate approach works better than simple regression techniques, being two the limiting factors:
- the adequacy of the theoretical model to describe real situations, and
- the type of numerical inversion technique applied to retrieve the surface parameters by inverting actual data against the theoretical model.
The problems are not so different in the optical domain than in the case of SAR. However, there are two factors appearing when working with optical data, not present in the case of SAR data:

♦ The presence of the atmosphere
The presence of the atmosphere introduces changes in the apparent reflectance of the targets, due to both the absorption and the scattering of radiation due to atmospheric components, as illustrated in the following figure:

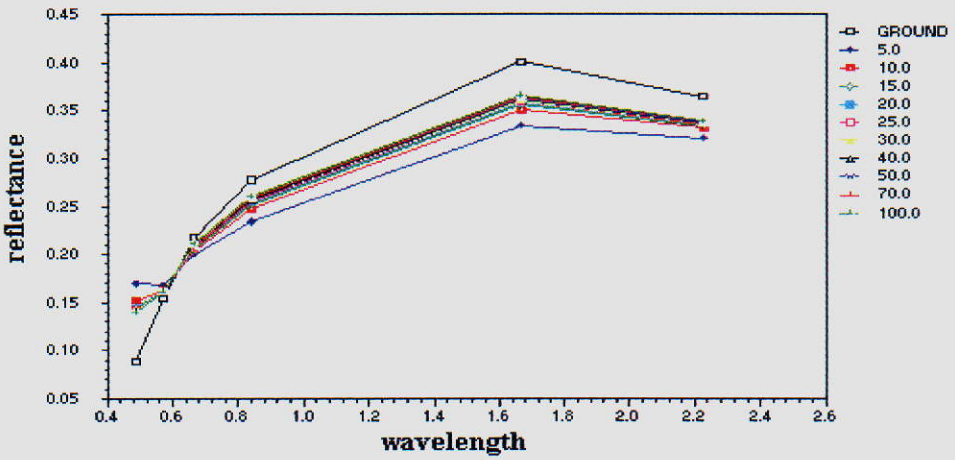


Figure 4: Visibility change in reflectance at different altitudes

This figure shows the effect of visibility changes over the apparent reflectance at the top of the atmosphere, within the spectral channels of the Landsat TM sensor, as compared to the actual ground reflectance (labelled as "ground"), for a typical soil surface, and for a fixed atmospheric water vapor content. Numbers in the legend correspond to different visibility values, in km.
Correction of the data for the perturbing atmospheric effects is now perfectly possible due to the good knowledge of atmospheric effects and the existence of radiative transfer codes to model such effects. The accuracy of the atmospheric corrections depends, however, on the knowledge of atmospheric status at the time of image acquisition. Since this knowledge is in most cases only approximate, this is the major source of uncertainty in the processing of temporal series of optical data.

♦ The diffuse illumination component
In the case of SAR data, the artificial illumination provided by the SAR system is essentially the only source of illumination. In the case of optical data, the source of illumination is the Sun, but, due to the atmospheric scattering, there is a strong component of "diffuse" illumination (coming from all angular directions) as opposed to the directional illumination coming directly from the angular position of the Sun over the target. Since both direct and diffuse illumination have different effects over the apparent reflectance of natural targets under varying illumination conditions (the ratio of diffuse to direct irradiance is highly depending on the particular atmospheric status), such effects must be properly compensated to interpret correctly temporal changes in series of remote sensing data.

Finally, although this also applies to SAR data, it must be emphasized that topographic effects play a significant role in data interpretation. The availability of high-quality Digital Elevation Models (DEM) for the study areas is a requisite to obtain quantitatively good results.

- Synergy use of microwave and optical data

Although recognizing that both data and microwave data, used independently, both provide useful information, it is clear that the simultaneous combined use of both optical and microwave information add many new potential capabilities.
The additional information content in optical-microwave synergy is clearly illustrated in the following figure, corresponding to the study area of Albacete, in Spain, an agricultural area with similar climatic conditions as the Arade area:

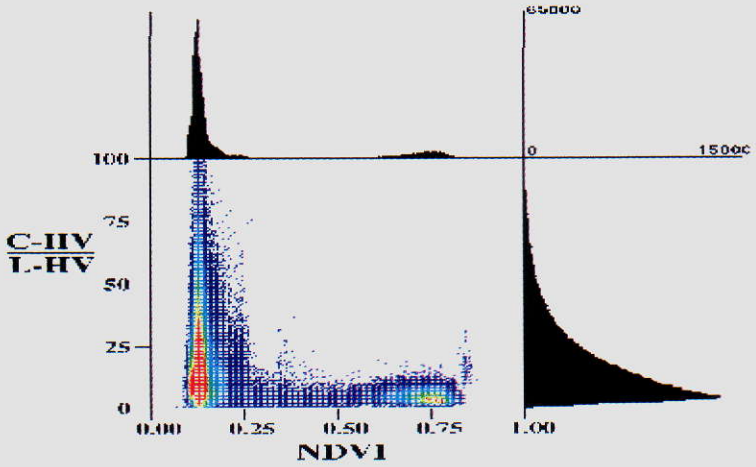


Figure 5: Comparison of the information content for optical data (represented by the vegetation index NDVI) and microwave data (represented by another "vegetation index", CHV/LHV). The lack of correlation is an indicator of the complementarity of the information provided by the two different types of data.

Optical-microwave synergy is typically applied to retrieve soil/canopy structural parameters and soil/canopy water content. If water content is low, the structural effect (macro-roughness) becomes the dominant information both in optical and microwave data, and the canopy structure is the main retrievable parameter. In the more common cases, where water content is driving the overall behavior of the optical and microwave signals, the synergy is used to better retrieve the water content by making use of the specific spectral absorption features in optical data and the sensitivity to changes in dielectric contact with moisture in the microwave domain.
In the case of AIMWATER, SAR data available will be limited to C-band and only one polarisation (VV), so that we will not have the same capabilities as when using full multi-frequency/multi-polarization SAR data. However, the same principle can be applied to ERS/SAR and Landsat/SPOT, so that the potentials of optical/microwave data synergy can be exploited.

• Application to the Seine and the Arade basins

The interest of the remote sensing techniques described above has been applied over the study area of the AIMWATER project: the Seine basin in France, and the Arade basin in Portugal.
For the Seine study area in France, microwave SAR data is playing the leading role, due to the lack of dense vegetation that can mask the soil moisture effects in the signal and the not so strong topographic effects, as well as the difficulty in getting series of optical data due to the high probability of cloud cover, being SAR systems the only way of getting full series of multitemporal data to monitor surface changes and soil moisture evolution. Optical data (Landsat) will play here the role of supporting retrievals based on SAR data.
For the Arade study area in Portugal, the situation is quite different. The abundance of vegetation, that is, the dense vegetation cover in almost the entire area, together with the important topographic effects due to the complex structure, makes very difficult to retrieve soil moisture from SAR data, because the vegetation effects will be dominant. On the other hand, the better weather conditions allow obtaining more easily temporal series of

optical data free of cloud cover. Over the flat central area of the basin, SAR data will provide relevant soil moisture data, while for the rest of the basin optical data will provide estimates of surface (canopy) wetness, still useful to monitor changes in overall moisture conditions.

- *Seine basin (France)*

Strategy

The two main indicators, which can be derived from space and be useful for hydrology models are: a vegetation index and a soil moisture index.
In the case of distributed hydrological models, we need to know these indices for each resolution mesh of the model. Conversely, in the case of the lumped hydrological models selected for the project, these indices should be representative of the whole watershed.
The methodology developed (Quesney *et al,* 2000) concerns the derivation of the vegetation and the soil moisture indices from SAR/ERS data. First, a vegetation map at the SAR/ERS spatial resolution (25×25 m^2) is obtained from SAR image multitemporal classification. Second, the land cover map is used to select the watershed areas for soil moisture index derivation: over these areas, an "equivalent bare soil SAR signal at watershed scale", $\tilde{\sigma}_0$, which is strongly correlated with surface soil moisture, is estimated. Finally, $\tilde{\sigma}_0$ is converted into soil moisture index at watershed scale through an empirical relationship. According to the SAR wavelength: $\lambda \approx 5.6$ cm (C band: frequency $f \approx 5.3$ GHz), the obtained index is representative of the soil moisture of the "superficial" layer, i.e. corresponding to the few first centimeters.
For each of the studied watersheds: the "Grand Morin", the "Serein", and the "Petit Morin", the first year of data acquisition is a calibration year. The following years of data are then used for the validation (see chapter 3.2). The processing strategy is different for the calibration year and for the validation year:
For the land cover map derivation in the first case (calibration year), the used method combines the multitemporal ERS/SAR series with the multispectral LandSAT image. In the second case (validation year), we adopt an operational perspective, which consists in processing each image once it is acquired. In each case, the classification is performed in an unsupervised way, and then, the unsupervised classes are interpreted *a posteriori* in terms of land cover types. This interpretation is done thanks to test areas selected on parts of the basins where the ground truth is known (field measurements).
For soil moisture index derivation, the first year of data acquisition is used as a calibration year during which the empirical relationship between $\tilde{\sigma}_0$ (processed SAR/ERS signal, in dB) and W_s (superficial soil moisture at watershed scale, in cm^3/cm^3) is established. Then, the data acquired during the following year(s) are directly inverted using the previously established relationship. The ERS/SAR signal is processed once the image is available, using the classification result of the considered date. When needed the vegetation corrections are performed using some vegetation parameters derived from a simple semi-empirical vegetation-growing model, also calibrated during the calibration period (first year of data).

Land cover map derivation

Remote sensing images have been used to derive the land cover maps of the studied watersheds. The aim of classification is to obtain a land cover map over the whole site (whereas, from direct measurements, it is known only on a sub-region). Different classification methods: monosource or multisource (i.e. combining the multitemporal ERS/SAR series with the multispectral LANDSAT/TM image) have been compared.

♦ Monosource classification

The monosource classification algorithm is applied either to the multitemporal ERS data or to the multispectral LANDSAT data. It is unsupervised. The classification process is composed of the three following steps: cluster characteristics estimation; image pixel classification; classified image interpretation.
In each pixel, the considered feature vector $\mathbf{x_r}$ for the ERS/SAR multitemporal classification is composed of the ERS/SAR backscattering coefficients σ_{db} measured at the different acquisition dates.
Cluster characteristics are estimated by the fuzzy c-means algorithm (Bezdek *et al.*, 1984). This method is particularly appropriate when there are mixed pixels (i.e. pixels belonging to several classes) on the images. In our case, after application of the averaging window on the ERS/SAR image, mixed pixels may occur specially in the neighborhood of field borders. In the fuzzy c-means algorithm, the number of classes c is an *a priori* (supervised)

parameter of the algorithm. In our case, the algorithm was run for different values of c, and an *a posteriori* value is derived from the best result according to the identification rate criterion (Le Hégarat-Mascle *et al.*, 1997). Moreover, as the fuzzy c-means algorithm is only locally optimal, we propose to initialize the clustering with c+1 classes from the result of the clustering with c classes (Le Hégarat-Mascle *et al.*, 2000).
Once the cluster statistical features have been estimated, classification is performed over the entire image. For classification, we assume that the distribution of the multitemporal feature vector $\mathbf{x}_r$ conditionally to a class i is a multi-variate Gaussian. Such a distribution takes into account the differences in the cluster sizes (in the feature space) through the covariance matrices. Then, we introduce a Markovian Random Field assumption for the label (or class) image to take into account spatial information. This was not done previously, during the cluster characteristic estimation, because of excessive computation time. *Maximum A Posteriori* classification is finally achieved according to (Geman and Geman, 1984) algorithm.
For the visible/infrared LANDSAT/TM data, the components of the considered feature vector $\mathbf{x}_0$ are the radiometric values measured in the different channels excepting the thermal infrared channel. As in the case of ERS/SAR multitemporal classification, we use the fuzzy c-means algorithm for the cluster characteristic estimation, and the *Maximum A Posteriori* criterion for the image classification (we also assume a multi-variate Gaussian distribution conditionally to the clusters for $\mathbf{x}_0$ and use Markov Random Field assumption to introduce neighborhood information).
Having performed the unsupervised classification of the image, the obtained classes are interpreted in terms of land cover types. This interpretation is done thanks to the test areas selected on subregions of the basins where direct measurements have been done and ground truth is known.

♦ Multisource classification

We choose a class subdivision approach. The basic idea of class subdivision is the following:
When two (or more) land cover types L_1 and L_2 are not discriminated by a sensor, they will be mixed in a same class C_1 (from monosource classification). Now, if another sensor can distinguish L_1 from L_2, but not L_1 from L_3, L_1 will be mixed with L_3 in a class C_2 of this second sensor monosource classification. Combining the 2-sensor information, we will able to discriminate L_1 as $C_1 \cap C_2$.
Such an approach is possible only if monosource classifications are good enough, since classification errors create "illegitimous" class subdivisions. However, as we will see in the results (section 3.2), our monosource classification results are sufficiently good to allow such an approach.

Derivation of the $\tilde{\sigma}_0$ - W_s relationship at a watershed scale

♦ Methodology

The methodology used was developed by (Quesney *et al.*, 2000) and validated by (Le Hégarat-Mascle *et al.*, 2000b). The principle is the following: since the measured signal is also a function of vegetation cover and soil roughness (besides soil moisture), we must cope from these two effects before relating it to superficial soil water content. Thus, the following steps have to be performed:

First, the SAR signal is "corrected" from the vegetation effect when this correction is necessary (presence of a vegetation cover) and possible. This second condition excludes in particular the cases of dense vegetation covers (forest areas, fully developed crops). In the case of sparse vegetation, the correction of the SAR signal should be possible for most of the vegetation types. However, nowadays, the radiative transfer model we use has been validated only in the wheat and barley cases, and thus, in this study, we limit our corrections to these crops. This processing provides then, for each date of data acquisition, an image where the pixel values correspond to the σ_0 values which would have been backscattered by the bare soil (suppression of the vegetation effect and obtaining of an "equivalent bare soil" signal). Of course, the obtained "equivalent bare soil" images have some missing pixels, corresponding to more or less large areas where the signal SAR was not corrected from the vegetation effect: forests, and a variable number of fields depending on the acquisition date.
Second, to free from the soil roughness effect, the chosen approach consists in averaging the signal over a sufficiently large region so that we can assume the roughness effect is approximately constant all along the year. In the case of an agricultural site, if the averaging process is done over a large number of crops, since the soil works do not occur at the same period of the year, this hypothesis is empirically verified.
Finally, the methodology consists in the following steps:

1. Selection, for each of the acquisition date, of the "sensitive" targets (land cover types) over which the "equivalent bare soil" SAR signal could be estimated;
2. Correction, for these targets and at the considered date, of the vegetation effect;
3. Filtering of the roughness change effect by performing an average over the greatest number of these targets;
4. Retrieval of the volumetric superficial (~5 cm) soil moisture, W_v, by inversion of the processed SAR signal $\tilde{\sigma}_0$ (vegetation corrected, roughness change filtered). In a first approximation, we use a linear relation between W_s in % and $\tilde{\sigma}_0$ in dB (Cognard *et al.*, 1995; Taconet *et al.*, 1996; Quesney *et al.*, 2000).

The first step is performed using the land cover map previously derived from ERS multitemporal /LANDSAT multispectral classification, and according the vegetation state. Table 1 shows, for the 4 kinds of vegetation considered: wheat, barley, corn and peas, the vegetation state month by month during the agricultural year, and, for each month, the selected crops.

	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
wheat	•	•	•	•	•	•			•	•	•	•
barley	•	•	•	•	•	•			•	•	•	•
corn		•	•	•	•							
peas		•	•	•						•	•	•

Table 1: Vegetation states of the 4 main Orgeval crops; white = bare soil, gray = sparse vegetation, green = dense vegetation; for each month, the red points indicate the selected crops for soil moisture derivation.

The second step is performed using vegetation model. According to a radiative transfer model derived from those of (Karam *et al.*, 1992), the measured signal σ_{meas} is the sum of the signal backscattered by the vegetation layer itself σ_{veg} and the signal backscattered by the soil σ_{soil}, this latter being attenuated when traversing the vegetation layer (before and after reflection):

$$\sigma_{meas} = \sigma_{veg} + \sigma_{soil}.e^{-2\tau} \quad (1)$$

where τ is the optical thickness of the vegetation cover. The effect of the vegetation is thus double: attenuation and contribution. This model has been validated in the case of the wheat crop, and can widen to crops whose features are very close to wheat ones, such as barley (but unlike corn and peas crops).

The correction of the vegetation effect is possible if we know the values of both τ and σ_{veg}. These two parameters can be estimated by the model knowing the geometrical and dielectrical features of the cover (wheat and barley crops). In the case of the calibration year, we use the direct measurements performed over the test fields during the ground truth campaigns concomitant to the ERS acquisitions. In the case of the validation year, the vegetation corrections are performed only during the March-April period (sparse vegetation period for the winter cereals), excluding July. For this period we assume a constant water content, and a linear relationship between the different geometrical dimensions of the leaves (or stalks) and the crop height. This assumption allows the deduction of all the needed (for τ and σ_{veg} estimation) geometrical and dielectrical crop features from its total height, which can be easily given either by the farmer, or (in a next future) by interferometric techniques.

Finally, having estimated τ and σ_{veg}, the correction, $\Delta\sigma$ in dB, to apply to the SAR measurements to retrieve σ_{soil} writes according to equation 1:

$$\Delta\sigma\,(\mathrm{dB}) = \sigma_{soil}\,(\mathrm{dB}) - \sigma_{meas}\,(\mathrm{dB})$$

$$= \frac{20.\tau}{\ln 10} + 10.\log_{10}\left[1 - \frac{\sigma_{veg}}{\sigma_{meas}}\right] \quad (2)$$

The third step is performed simply averaging the "equivalent bare soil" images (excluding uncorrected pixels). This process has two aims: first, this is the empirical way we find to free from the local changes of roughness during the agricultural year; second, since the soil under the different crops may present different states of draining (depending on the characteristics of the crops), the greater is the number of fields and crops considered, the more representative of the whole watershed will be the derived index.

For the fourth step, we distinguish between the calibration year where the $\tilde{\sigma}_0$ - W_s relationship is established for each watershed, and the following years where the data inversion is performed according the relationship corresponding to the considered watershed.

- *Arade basin (Portugal)*

The objective is to use both optical and microwave data in the study of vegetation and soil moisture evolution over the Arade basin. The first ones have been used to correct the vegetation effects and the second ones to obtain moisture levels at the basin scale. For this study, we used a high quality Digital Elevation Model (DEM), along with SPOT (optical) and ERS-2 (microwave) satellite images. The image pre-processing along with the derivation of the NDVI index and soil moisture index are described there after, while the operational methodology to retrieve the surface parameters is presented in full in the next chapter.

Image pre-processing

♦ DEM

We used the DEM shaded image as reference to do geometric and topographic corrections of satellite images. A DEM is a file that contains a matrix with elevation values reproducing the topography of the zone. After this first step, we were able to perform the classification of the area in the classes with hydrological meaning, then allowing further studies and data analysis.

The complete DEM has a size of 6001 x 2501 pixels with a resolution of 8 meters per pixel in the x and y-axis, and 1 meter in the z-axis. Figure 6 represents the elevation of the complete DEM. The minimum elevation value is 22 meters and maximum is 581 meters. The white outline defines the region of interest, the Funcho sub-catchment. The dam is located at the left down in the study area; this is the site with less altitude.

The first step was to study the surface topographic characteristics by analysing the DEM using IDL Rivertools software. This software generates an image, where the values decrease in the direction where water will nominally move starting from a nominal rainfall point. We can use this information for obtain a mask who delimitates the study catchment, comparing this image with an outline of study area, which was drawn in a map. We also use this image in order to define the main rivers.

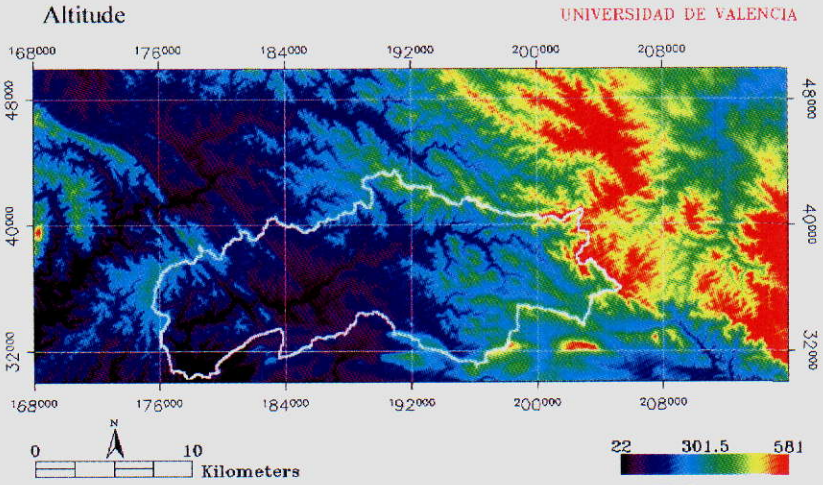


Figure 6: Elevation representation of complete DEM.

♦ SPOT and LANDSAT images

There are two main effects that must be corrected in optical images:

-Geometric distortions for satellite motions during acquisition, terrain curvature, Earth rotation, orbit inclination, etc..., that make necessary a geometric correction.

-Influence of topography and sun position on the irradiance over each pixel. In the SPOT image we can observe some shaded areas and others brighter areas, but this does not mean that these areas are of a different class of terrain than the terrain surrounding them. This is a topographic effect and its correction is called slope-aspect

correction. For the geometric correction and for the slope-aspect correction it is necessary to have a very good digital elevation model (DEM). In our case we have a DEM with a resolution of 8x8 m^2 and 1 meter in elevation. Besides, the elevation model allows us delineate exactly the study area.

Two Spot images should have been acquired in order to follow the seasonal changes: one corresponding to winter conditions and another corresponding to summer conditions. In fact for summer conditions the SPOT image was not available and a LANDSAT image with 30 meters resolution was acquired instead.

Figure 7 shows the full SPOT scene of 2 January 2000 that we have acquired for our study. It has 3000 x 3000 pixels, with a resolution of 20 meters in x and y. It has 3 bands corresponding to green, red and infrared. In this figure we have made a composition with the tree bands in order to obtain a false colour image. We have used the band 3 to red, band 2 to green and band 1 to blue. The processing level is 1A, which means that no corrections have been made aside the detector response normalization and that the image values are the digital counts that the sensor took (SPOT Users Handbook). We have ordered the image with this processing level because it allows us to control the quality that we will obtain in the next processing steps until the final results are obtained.

Due to satellite motions during image acquisition, geometric distortions appear in the resulting images. On the other hand, the terrain curvature, the Earth rotation, the orbit inclination and sometimes the not square pixel due to off-nadir viewing conditions, make necessary the geometric correction when we want to compensate the image deformations.

The procedure is first to obtain a image of the illuminated DEM. The illumination is defined as the cosine of the incident solar angle in relation to the normal on a pixel, thus representing the proportion of the direct solar radiation hitting a pixel. To calculate the illumination of the DEM we used the solar azimuth and elevation angles at the time when the SPOT image was taken. This information is given with the satellite image. With the image of the illuminated model we can take pairs of ground control points in the digital elevation model and in the SPOT image. This consists in locate the same position in the model and in the satellite image. When we have enough ground control points we can apply the program that performs the geometric correction. We have used the ENVI software in order to carry out all this part of the processing. This procedure has been applied only for the study area.

Finally we obtain a corrected satellite image that covers exactly the same area as the DEM and that we can superimpose to it. To obtain a good result in geometric correction depends directly of the quality and resolution of DEM. Ideally we need that its spatial resolution to be better than spatial resolution of SPOT image. This is our case, we have 8 meters of resolution in x and y on our DEM and 20 meters on satellite image. That is the reason for the good results we have obtained.

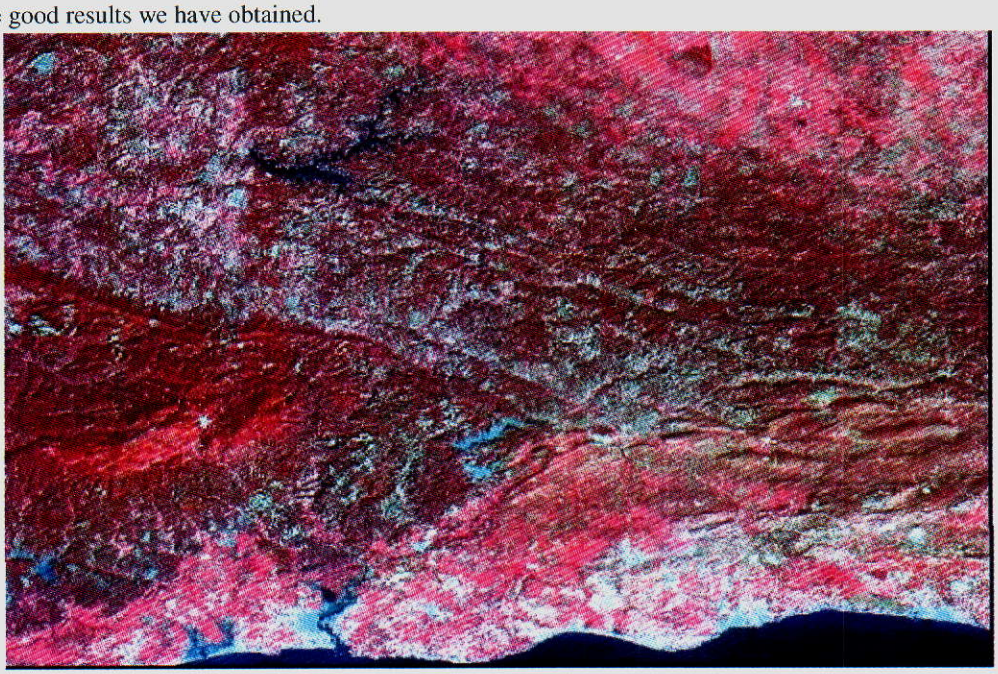


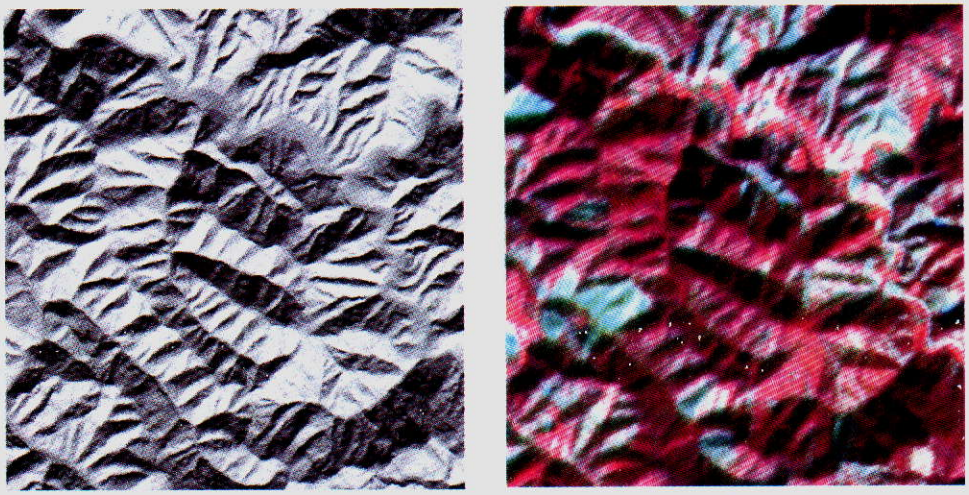
Figure 7: Full SPOT Scene

The geometrical correction is not the final step in the processing of the SPOT image. We must compensate the influence that topography and sun position has on the irradiance over each pixel. In the SPOT image we can observe some shaded areas and others areas more bright, but this does not mean that this areas are of a different class of terrain than the terrain surrounding them. This is a topographic effect and its correction is called slope-aspect correction.

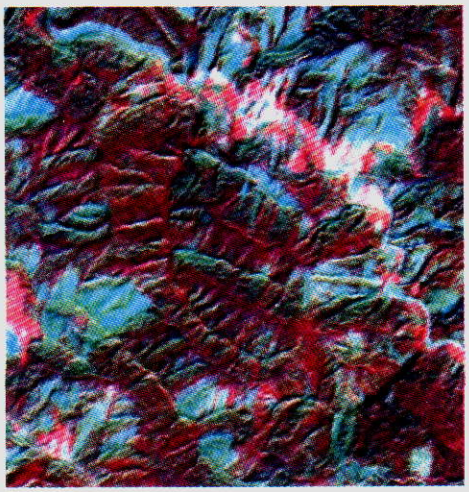
Remember that the processing level of the SPOT image is 1A. We have worked with digital counts. Now, for topographic correction we want to work with reflectance values, so we have to pass the digital counts to reflectance. To do this we apply standard algorithms and finally we have the geometrical corrected SPOT image in reflectance values.

To do the topographic correction we need the *DEM illuminated* that we have used previously in the geometric correction and the geometrical corrected SPOT image. To obtain a good topographic correction we want that the illuminated DEM and the geometrical corrected image coincide so, when better is the geometrical correction of SPOT image better is the topographic correction.

The statistic-empirical correction is a purely statistical approach based on a linear relationship between the original band of the geometrical corrected SPOT image and the illuminated DEM image. Geometrically the correction rotates the regression line to the horizontal to remove the illumination dependence (Meyer, 1993).



(a) (b)



(c)

Figure 8.- (a) DEM Shaded Image; (b) SPOT image before topographic correction; (c) SPOT image after topographic correction.

Figure 8 is an example of the procedure applied. In (a) we can see a zone of the illuminated model and (b) shows the same area in the SPOT image after applying the geometric correction. We can see that geometric correction is very good because the two images are very similar. In (b) we can observe the effects that topography induces (shaded and bright areas). The coincidence between the SPOT image and the illuminated DEM allows us to compensate the topographic effects with the slope-aspect correction. Figure 8 (c) shows the result of applying the topographic correction to the SPOT image. We can see that the topographic effects have been almost completely removed.

Finally we have applied the mask obtained with Rivertools to the SPOT image fully corrected and the result is showed in the Figure 9. With this information the terrain classification can be performed and the vegetation zones can be located. The same procedure is applied to LANDSAT image with corresponding parameters.

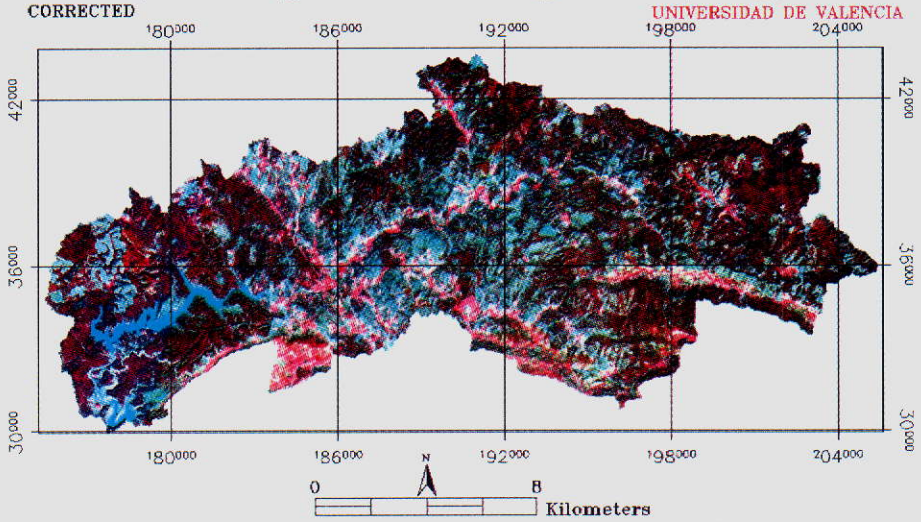


Figure 9: SPOT image of the Arade study area after all applied corrections.

♦ ERS2/SAR

As a complement to the SPOT satellite images ERS SAR data were used. Imaging radars are side-looking rather than nadir-looking instruments, and their geometry is complicated by foreshortening (the top of a mountain appearing closer to the sensor than does the foot of the mountain) and shadow, caused by the 'far side' of a mountain or hill being invisible to the side-looking radar sensor. Furthermore, the interaction between microwave radiation and the ground surface generates a phenomenon called speckle, which is the result of interference arising from the coherent integration of the contributions of all the scatters in the pixel area. Speckle magnitude is proportional to the magnitude of the backscattered signal, and is rather more difficult to remove from the image than the additive noise (Mather, P.M., 1999).

ORBIT	ROW	CENTER LATITUDE			CENTER LONGITUDE			DATE
		N/S	Grad.	Min.	E/W	Grad.	Min.	
25084	180	N	37	23	W	8	26	06/02/2000
25585	180	N	37	23	W	8	26	12/03/2000
26086	180	N	37	23	W	8	26	16/04/2000
26587	180	N	37	23	W	8	26	21/05/2000
27088	180	N	37	23	W	8	26	25/06/2000
27589	180	N	37	23	W	8	26	30/07/2000
28090	180	N	37	23	W	8	26	03/09/2000
29092	180	N	37	23	W	8	26	12/11/2000
29593	180	N	37	23	W	8	26	17/12/2000
31096	180	N	37	23	W	8	26	01/04/2001
31597	180	N	37	23	W	8	26	06/05/2001
32098	180	N	37	23	W	8	26	10/06/2001
33100	180	N	37	23	W	8	26	19/08/2001
33601	180	N	37	23	W	8	26	23/09/2001

Table 2 ERS-2 / SAR satellite images.

Although a monofrequency radar signal does not detect colour information (which is gained from optical wavelength sensors as SPOT) or temperature information, it detects surface roughness and electrical conductivity information (which can be related to soil moisture conditions).
ERS-2 / SAR image is a single channel image (C-VV). We needed time series of images in order to know the evolution of soil moisture and vegetation. The main advantage of SAR is that it is an all-weather, day-night, high-spatial-resolution instrument, which can operate independently of weather conditions or solar illumination. This is an advantage over SPOT, which availability is limited to clear skies.
The first part of the pre-processing of ERS-SAR images is similar to the one for the optical images: geometric correction, calibration of digital counts in a physical variable and topographic correction.
The procedure for the geometric correction is the same used for optical images, but we did not use in this case the DEM illuminated as the reference image for the correction, instead of it we used the corrected SPOT image, in this way the radar image is registered directly over the optical data, previously registered to the DEM. This is necessary for the correction of the vegetation effects on the backscattering signal. The rest of the radar images in the time series were geometrical corrected to the first radar image, in this way we ensure that all the images are in geometrical correspondence, which is important for the spatial and temporal study that we had developed.
The case of the topographic correction is more complex. We tried to apply the same method than for the optical data, but the results were not the expected ones. So we decided not to apply such topographic correction to radar data and to assume the effect of topography in the radar signal. Since the final moisture map will have a resolution of 1 km, we can make some approximations by neglecting some topographic effects.

Derivation of the NDVI index

The idea is to use the NDVI index to determine the areas over the study site with low vegetation. In this way we know that the values of soil moisture that we would obtain with the ERS/SAR images in these areas have no influence of vegetation. In figure 10 we see the NDVI index for the study area.

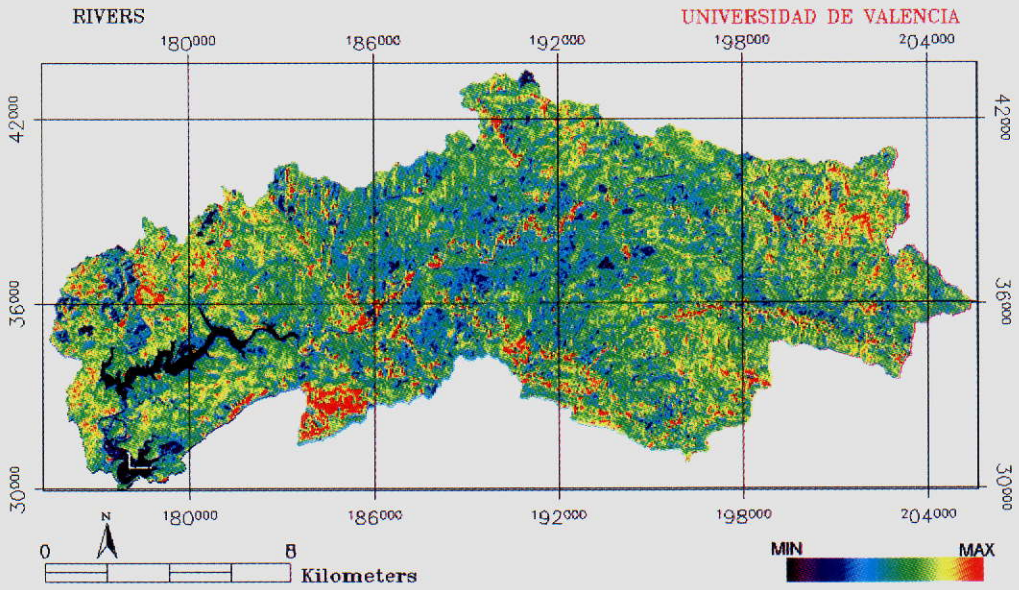


Figure 10: Colour representation of the NDVI over the Arade study area.

The maximum value that we have obtained for the NDVI is around 0.7, so we have considered that areas with a NDVI greater than 0.4 have too much vegetation to be taken into account when deriving soil moisture values from the SAR/ERS images without vegetation effect compensation in backscattering signal. Figure 11 shows the mask that we have obtained with this criterion.

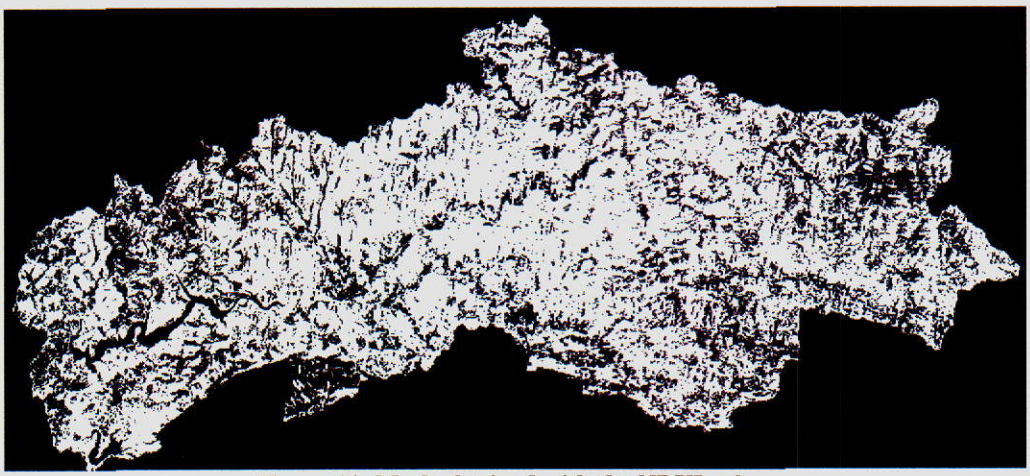


Figure 11: Mask obtained with the NDVI value.

Derivation of soil moisture from SAR/ERS images and ground measurements

For the derivation of soil moisture values we must find a relationship between the backscattering of the SAR/ERS multitemporal series of images and the ground moisture values measured simultaneously to the satellite overpass. We have a total of fourteen images and around two years of ground measurements, TDR and gravimetric, taken weekly (beginning in January).

Each one of these images has been geometrically corrected in order to match exactly with the SPOT image. No more corrections have been applied to the images; tophographic effects in SAR/ERS images can not be removed with the method applied to optical and we have limited to elaborate a mask for each one of the images marking the areas where topography has a big influence in order to not consider them in the soil moisture derivation. We have observed that critical topographic conditions prevent quantitative derivation of soil moisture values, and topographic corrections cannot alleviate this problem. Then we have restricted to areas with no critical topographic conditions by masking the ones with slopes larger than a critical value.

Finally, to obtain the moisture map for each one of the SAR/ERS images we used the relationship between the backscattering of the SAR/ERS multi-temporal series of images and the ground moisture values (see chapter 3.2).

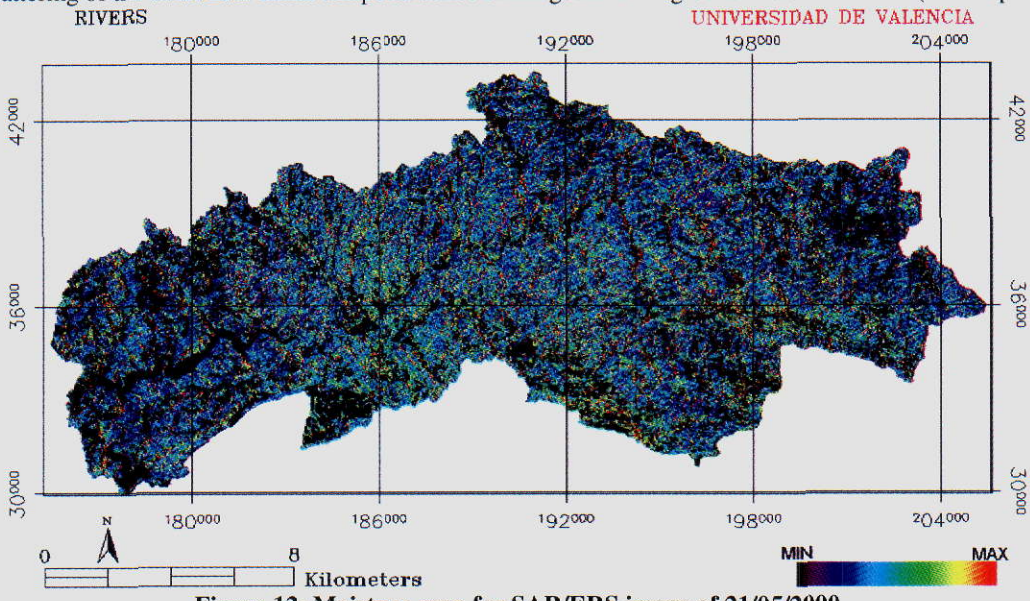


Figure 12: Moisture map for SAR/ERS image of 21/05/2000

Origin: U. Valencia/CETP/Cemagref
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

• Potential and limitations

- *Microwaves*

Limitations

♦ Vegetation

The first and main limitation to the use of microwave remote sensing data to monitor soil moisture from space is the presence of vegetation. We saw that, in the case of dense vegetation cover, no information about soil features can be deduced from the measured signal both in the case of active and passive microwave techniques. This means in particular that such data are useless in the case of watersheds with dense vegetation cover such as those covered by forests or bushes. In the intermediate case where the basin contains a majority of forested areas but also minority areas with bare soil or sparse vegetation (where vegetation corrections are possible), the soil moisture can be estimated on these minority-unforested areas. However, some care must be granted before using the obtained value as representative of the whole watershed soil moisture.

♦ Relief

The case of mountain, small valley or hilly watersheds may also raise some problems. In the case of SAR data, they are due to the necessity to correct the remotely sensed signal from the effect of the relief (inducing variable local incidence angles). Nowadays, some software has been developed to correct the SAR signal from this effect. However, they require DEM (Digital Elevation Model), and even after these corrections, the problem of areas viewed with variable local incidence angles remains. In the case of passive data, the solution would perhaps be to proceed as in the case of heterogeneous soil or vegetation: simulate the remotely sensed brightness temperature as a weighed sum of brightness temperatures computed on homogeneous areas regarding the soil, the vegetation, and the local incidence angle (of course, DEM data are also needed inputs).

♦ Ancillary data

The main problem which could be encountered in the application of soil moisture retrieval techniques, from passive microwave data, over large regions is the acquisition of the needed ancillary data: land cover and associated b parameter values, vegetation water content, surface temperature, Q and h' parameters. If some of these data or parameters may be estimated either from database or multichannel microwave data, it seems difficult to get an estimation of the surface temperature, except using thermal infrared data or meteorological data. The first solution using IRT data is not very satisfying since it cannot be used in the presence of cloud cover, and the second one assume either the presence of one or several meteorological stations within the studied area, or the sufficient precision of some meteorological model simulations. Clearly, for spaceborne operational applications, over large and heterogeneous regions, the acquisition of all these ancillary data appears as a real challenge. In the case of active microwave, the problem is not so acute since, once the empirical linear relation has been calibrated only ERS data are needed. Even the optical image used to get land cover map is not really necessary since it can be obtained (with slight decrease of the performance, Le Hégarat-Mascle *et al.*, 2000a) from multitemporal ERS data.

Potential

♦ In the field of passive microwave, during the past two decades, experiments carried out using ground based and aircraft radiometers have demonstrated the basic principles and feasibility of soil moisture estimation (e.g. Wang *et al.*, 1989; Schmugge *et al.*, 1992; Jackson and Le Vine, 1996). However, large-scale demonstrations should still be done as soon as an operational satellite sensor will be available. In particular, research to infer ancillary data in an operational way is still needed.

♦ In the field of active microwave, previous studies lead to the development of a semi-empirical methodology to derive soil moisture in the case of agricultural fairly flat regions (Quesney *et al.*, 2000). The obtained relation is empirical. The validation of the methodology and the robustness of the relation over a large number of watersheds need to be demonstrated. The case of hilly areas has not yet been considered. Besides processing taking into account variable local incidence angles, the interpretation of the derived soil moisture and its assimilation into hydrological models is not as trivial as in the case of basins exhibiting homogeneous soil moisture.

♦ Finally, the main challenge for remotely sensed soil moisture use is its assimilation in models (soil-vegetation-atmosphere models, hydrological models, as well as meteorological ones). In particular, we saw that the sensed depth is one of the limitations of the microwave data: soil moisture estimations are superficial values (few centimeters), which does not always correspond to the quantities used in the models: soil water content in the different layers, in the root zone, or water stock. The development of assimilation techniques (e.g. Entekhabi *et al*, 1994) should enable to link all these quantities, and to deduce the deep layer moisture from the superficial one.

- Optical

Limitations

There are two major limitations of optical data for hydrological applications:

♦ Cloud cover

The monitoring of surface moisture requires temporal series of data. Since orbital constraints make possible to obtain images only with given periods of time, if some of these potentially possible images are lost due to cloud cover at the time of the satellite overpass still the series can be useful if enough images are acquired, but in some areas with quite frequent cloud cover the number of images lost due to cloudiness can be so high that makes impossible the monitoring applications. This makes optical data more commonly used in Mediterranean countries (like the Arade area) while SAR data are more commonly used in Northern European areas (like the Seine area). Apart from other reasons, cloud cover is in most cases the dominant factor.

♦ Penetration depth

The second main problem with optical data is the penetration depth that is typically proportional to the wavelength of the signal. In the case of SAR data wavelengths are of the order of centimeters (around 5 cm for C band) so that the wave penetrates into the soil (or the vegetation) to get information from below the most superficial layer. In the case of optical data the wavelengths are of the order of micrometers, so that the signal is only sensitive to the most superficial layer, almost fully independent of the status of deeper layers in the soil. If the most superficial layer is completely dry, optical data is fully insensitive to the moisture contents of deeper layers, making these data useless to monitor soil moisture effects. In some cases, the superficial layer is still indicating some moisture content that can be related to the water content of deeper layers by assuming some modeling of the soil water content profile. In any case, the limited penetration aspects of optical data must be considered when interpreting retrievals in actual applications. Synergy with SAR data plays here an important role. In any case, the mayor limitation in the application of optical data for hydrological applications is given by the constraint imposed by topographic effects. Since most of the study areas with hydrological interest (like the case of the Arade study area) are located in mountainous areas with complex topographic structures, proper correction of the data from the problems introduced by topographic distortions are required.

Potential

The perspectives for future improvements on the capabilities to retrieve information about surface parameters from optical data are very promising, especially since there is many new satellite systems that will become available very soon that will provide extremely better capabilities that those provided by current systems such as Landsat or SPOT, previously described. Other systems currently planned but not yet ready for launch also include much more increased capabilities. The main new improvements come from the increase in the number of spectral channels (going up to hundreds of channels in some of the systems) and also due to the incorporation of multiangular capabilities. A particularly interesting future system is the Land Surface Processes and Interaction Earth Explorer Core Mission within the ESA Explorer Program. This mission will represent a quite significant increase in the capability of optical remote sensing to monitor land surface processes.

In any case, most probably none the SAR-based systems or the optical-based systems will be able to provide a real solution to the problems, and it is the synergy among all the different observation systems (especially the integration of optical and microwave SAR data) the solution that appears more promising for future use of remote sensing data in hydrological applications.

3.2 Validation of operational methodologies for soil moisture monitoring

The methodologies described above have been validated over the Seine basin and the Arade basin and operational processes have been proposed.

- **Seine basin (France)**

- Data base

♦ Remote sensing images

Two kinds of satellite data have been acquired:ERS2/ SAR images and the LANDSAT/TM data (visible/infrared optical data) (see § "Spatial data and the web site database" in chapter 2.1). Table 3 summarizes the remote sensing images acquired until April 2001) for the AIMWATER project.

watershed	Grand Morin	Serein	Petit Morin
calibration	20 ERS images 1 LANDSAT image	23 ERS images 1 LANDSAT image	20 ERS images
validation	25 ERS images	29 ERS images	15 ERS images

Table 3: Remote sensing images processed in the frame of the AIMWATER project

♦ Ground truth measurements

Simultaneously to the SAR/ERS acquisitions, ground truth campaigns have been performed. On one hand gravimetric soil moisture measurements have been performed first to calibrate the $\tilde{\sigma}_0$ -W_s relationship, and then to validate it. On the other hand, during the calibration year, the vegetation features have also be measured in order to get accurate estimates of the vegetation effect and free from it during the calibration step, and to derive the empirical relationship between the crop total height and the other geometrical dimensions of the crop. These field measurements were presented in chapter 2.1.

- Calibration year results

Classification results

♦ Monosource classification

In the case of the Grand Morin, for the ERS multitemporal classification, the used images are those acquired respectively on 99/02/03, 99/02/22, 99/03/10, 99/03/29, 99/05/03, 99/05/19, 99/06/07, 99/06/23, and 99/07/28. For the LANDSAT multispectral classification, image acquisition date is 99/05/03. In the case of the Serein, for the ERS multitemporal classification, the used images are those acquired respectively on 99/01/15, 99/01/31, 99/02/19, 99/04/11, 99/04/30, 99/06/04, 99/06/20, 99/07/253, and 99/08/13. For the LANDSAT multispectral classification, image acquisition date is 99/04/01.
Table 3 shows, in the case of the Grand Morin, the obtained identification rates and estimated percentages over the watershed of every land cover types present on the basin. Table 4 is similar to table 3 in the case of the Serein. The identification rate of a given land cover κ represents the confidence we may have in κ land cover interpretation after κ class merging (in other words, the "identification rate" represents the "degree of confidence in the land cover identification").
The percentages giving the significance of the considered land cover type over the basin have been estimated from the classification result. Their sum is different from 100 % since there are some "mixed" classes, i.e. classes, which correspond to several land cover types and cannot be interpreted. Therefore, in many cases, the given values are only an inferior limit of the actual value. Moreover, since the computed significance values depend on the classification result (and its interpretation), they may differ in accordance with the used data. Also, when an identification rate is too bad, we consider the estimated significance is not valid.

Land cover type	Identification rate		Significance over the basin	
Forest	99.7 %	94.4 %	26.9 %	27.3 %
Wheat	98.7 %	89.1 %	29.4 %	28.5 %
Peas	88.2 %	88.7 %	8.2 %	4.6 %
Corn	85.4 %	90.1 %	4.6 %	2.7 %
Colza	93.2 %	95.2 %	9.1 %	5.2 %
Flax	74.7 %	13.0 %	1.8 %	/
Barley	81.0 %	4.8 %	3.9 %	/

Table 4: Grand Morin classification results from: ERS data (grey cells) and LANDSAT data

Land cover type	Identification rate		Significance over the basin	
Forest	93.6 %	99.6 %	20.4 %	32.6 %
Wheat	90.6 %	58.7 %	11.5 %	/
Barley	80.0 %	18.0 %	5.6 %	/
Colza	84.1 %	85.1 %	8.4 %	6.2 %
Grass land	95.0 %	79.3 %	35.8 %	15.7 %

Table 5: Serein classification results from: ERS data (grey cells) and LANDSAT data

From table 4 and table 5, we note that majority land cover types are very well identified. This is not the case for most of the minority land covers types in the optical classification, and for some of them in the SAR classification. The fact that the detection of the land cover types, which are less represented, is penalized, can be easily explained by the optimization criterion choice: the least square error minimization. Indeed, since such a land cover type only contains few samples, having a cluster, which represents it, may less reduce the least square error function than subdividing a land cover type with numerous samples. Moreover, minority land cover types are also penalized in the identification rate computation by the fact that the number of test fields representing them is low.

In the case of the Grand Morin watershed, we note that the ERS/SAR multitemporal classification leads to very good results: only very minority crops, such as sugar beet fields and grass land areas are not identified (they have not been reported in table 2). However, they represent less than 5 % of the basin. In this result, colza fields are slightly overestimated. The LANDSAT/TM multispectral classification seems less performing: it does not allow the identification of flax and barley fields. It underestimates the cornfields. However, its estimation of the colza fields is probably better, even if that does not appear clearly on the identification rate values (because of the limited number of test fields we have).

In the case of the Serein watershed, the performance of the multitemporal ERS/SAR classification is verified once again: it allows the identification of all the land covers, whereas with the LANDSAT/TM classification the distinction between wheat and barley crops is not possible. This is due to the acquisition date of the LANDSAT/TM image: April, when these two crops are very similar (actually their harvest date is a determining feature for their separation, that explaining the success of the multitemporal approach). When we merge wheat and barley in a same land cover type, called "winter crops", the identification rates become respectively equal to 95.6 % for ERS/SAR classification and 69.8 % for LANDSAT/TM one. The second main difference between radar and optical classifications is the estimated percentage of the land cover types over the basin, in particular forest areas and grass land areas: these latter are overestimated with the radar data, while they are underestimated using the optical data. The main undetermined areas are located in the South part of the watershed.

In summary, monosource classifications, such as LANDSAT/TM multispectral classification or ERS/SAR multitemporal classification, allow the identification of the most significant land cover types. The identification of minority land cover types is more difficult: some of them may be identified or not depending of the considered sensor and the acquisition dates. However, since their misdetection may be due to confusion with different land cover types versus the considered sensor (ERS/SAR or LANDSAT/TM), classification based on data fusion may allow their discrimination. Multisource combination may also allow the further improvement of the majority land cover type identification by correction of the slight misclassifications affecting them. The study of this possibility is the purpose of the next section.

♦ *Multisource classification*

Subdivision method is applied from previous monosource classification results. Tables 6 and 7 give the obtained identification rates and significance (in percentages) over the whole basin, in the case of the Grand Morin and the Serein watersheds.

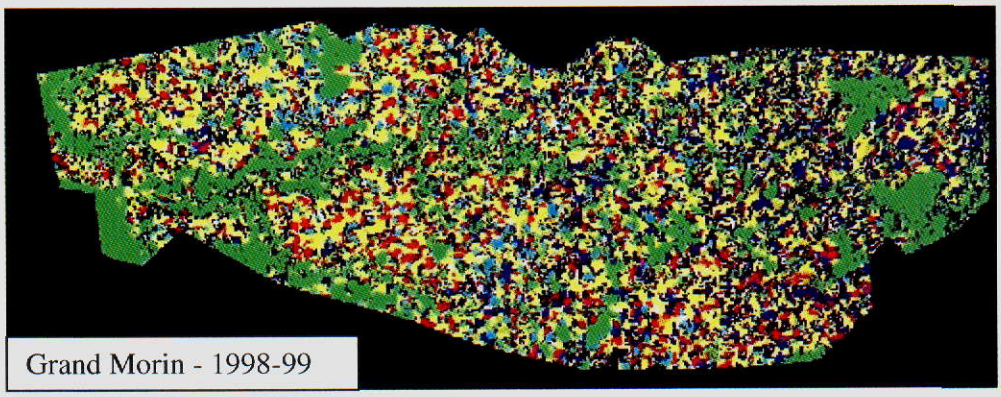
Land cover type	Identification rate	Significance over the basin
Forest	99.7 %	27.3 %
Wheat	99.5 %	28.7 %
Peas	90.4 %	8.1 %
Corn	85.8 %	4.4 %
Colza	97.4 %	4.8 %
Flax	98.8 %	1.4 %
Barley	88.4 %	3.7 %

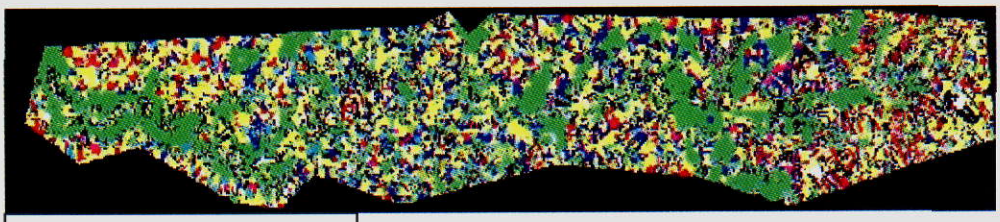
Table 6: Grand Morin multisource classification result

Land cover type	Identification rate	Significance over the basin
Forest	99.9 %	25.6 %
Wheat	94.4 %	10.8 %
Barley	86.9 %	6.0 %
Colza	99.0 %	11.0 %
Grass land	97.8 %	28.3 %

Table 7: Serein *multisource* classification result

In the case of the Grand Morin, we note that all land cover types are now identified with a confidence degree greater than 85 % (the lowest value corresponding to the corn), and that the "greatest" improvements correspond to the colza, flax and barley crops which are minority crops (indeed monosource classification already allows the identification of the majority land cover types).
In the case of the Serein, we note that more than 80 % of the watershed is now interpreted, forest and grass land areas ERS/SAR bias in their estimation having been corrected. Among the uninterpreted areas, in the North part of the basin, we find some vineyards, and some small villages, and in the South part, vegetation areas that are intermediate between forest and grass land in terms of vegetation density and height.
Figure 13a shows the result of ERS-LANDSAT classification combination in the case of the Grand Morin watershed, Figure 13b in the case of the Petit Morin watershed, and Figure 13c in the case of the Serein watershed. In the legend, the label "mixed" gathers the pixels belonging to a mixed class with those belonging to a not interpreted class (i.e. having no intersection with the test fields).


Grand Morin - 1998-99


Petit Morin - 1998-99

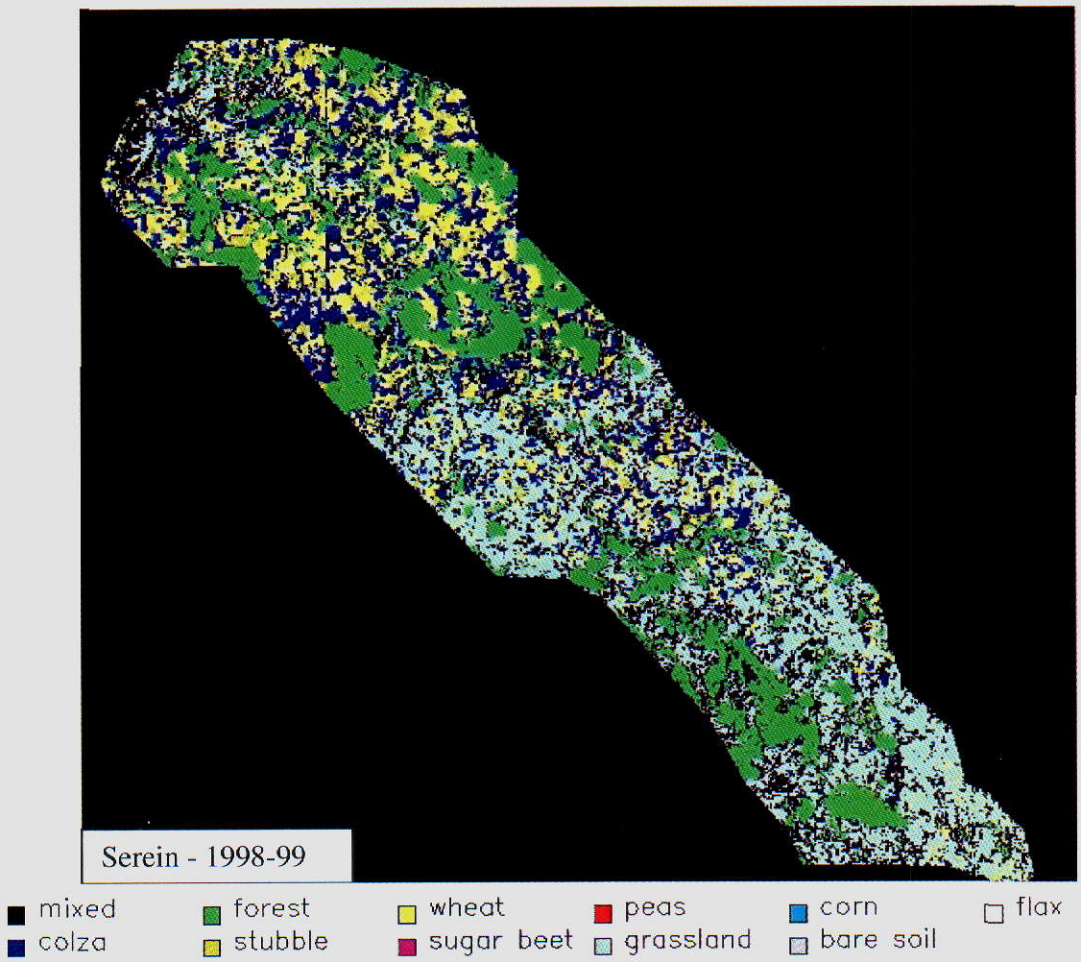


Figure 13: Results of ERS-LANDSAT classification combination after interpretation; (a:) case of the Grand Morin watershed; (b): case of the Petit Morin watershed (c): case of the Serein watershed;

Empirical vegetation growing model

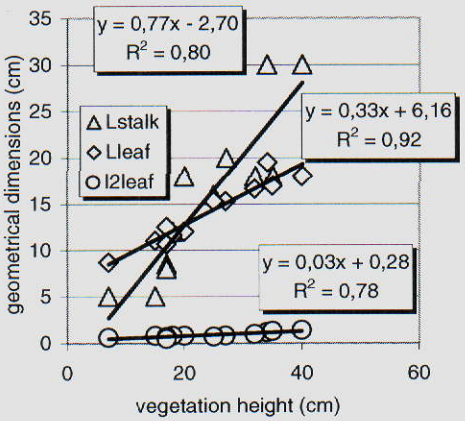


Figure 14: Winter cereal geometric parameter measurements and empirical growth model relating stalk

Origin: U. Valencia/CETP/Cemagref
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

length (Lstalk), leaf length (Lleaf), and leaf width (l2leaf) to the crop total height.

Figure 14 shows the measured vegetation parameters and the derived empirical model. It is valid only for the green sparse vegetation (February to end of April) period and winter cereals. According to it, stalk length, leaf length and width are linearly related to the crop total height. The other geometrical dimensions (stalk diameter and leaf thickness) are considered as constant. There is no ear during the considered period.
We also check the consistency of vegetation water content during this period.

Watershed relationships $\tilde{\sigma}_0$ - W_s

For soil moisture index derivation, the first year of data acquisition is used as a calibration year during which the empirical relationship between $\tilde{\sigma}_0$ (processed SAR/ERS signal, in dB) and W_s (superficial soil moisture at watershed scale, in cm^3/cm^3) is established (see chapter 3.1).

Figure 15 shows the processed SAR/ERS signal $\tilde{\sigma}_0$ versus the measured surface (0-5 cm) soil moisture W_s, in the case of a) the Grand Morin watershed, b) the Serein watershed, and c) the Petit Morin watershed. Correlation coefficient values are rather satisfactory: R^2 ranges from 0.85 to 0.93 depending on the considered watershed. The slope of the linear relationship seems to be very consistent from one watershed to another (0.33 or 0.34). It is also consistent with previous results obtained by (Quesney *et al.*, 2000) or (Le Hégarat-Mascle *et al.*, 2000b). Conversely, the offset may varies with respect to the watershed, because of a different roughness mean value.

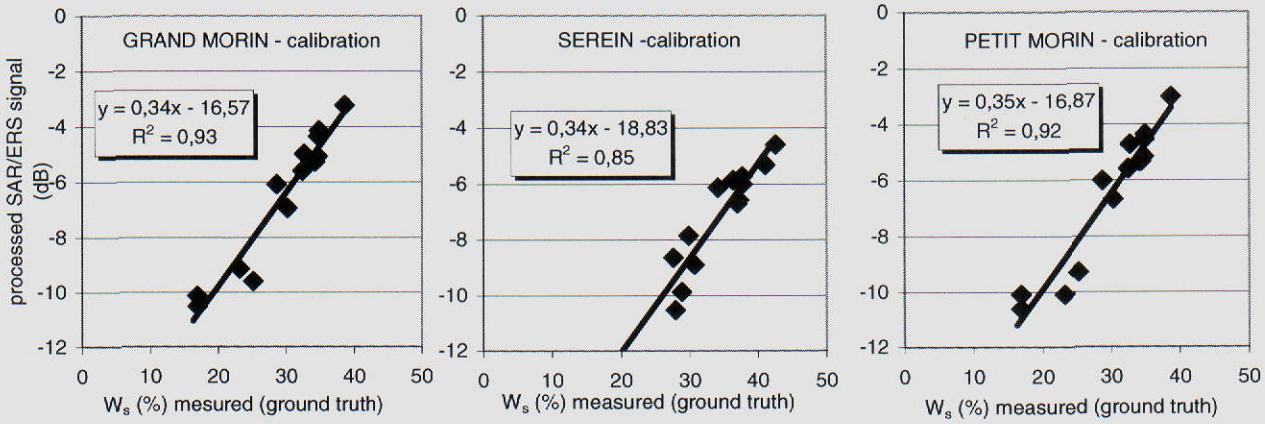


Figure 15: Processed ERS/SAR signal (in dB) versus measured volumetric soil moisture (in %), respectively in the case of a) Grand Morin, b) Serein, and c) Petit Morin watershed. Each point corresponds to a different date of the calibration year (January 1999 to October 1999).

	R^2	number of points	$\tilde{\sigma}_0$ (dB) = f(W_s(%))	W_s (%) = g($\tilde{\sigma}_0$ (dB))
Grand Morin 1999/01 → 1999/10	> 0.9	13	$\tilde{\sigma}_0$ (dB) ≈ 0.34×W_s(%)-16.6	**2.74×$\tilde{\sigma}_0$ (dB)+47.5**
Serein 1999/01 → 1999/10	> 0.8	13	$\tilde{\sigma}_0$ (dB) = 0.34×W_s(%)-18.8	**2.52×$\tilde{\sigma}_0$ (dB)+52.5**
Petit Morin 1999/01 → 1999/10	> 0.9	13	$\tilde{\sigma}_0$ (dB) = 0.35×W_s(%)-16.9	**2.65×$\tilde{\sigma}_0$ (dB)+47.0**
Orgeval 1999/01 → 1999/10	> 0.9	13	$\tilde{\sigma}_0$ (dB) = 0.32×W_s(%)-16.0	**2.85×$\tilde{\sigma}_0$ (dB)+48.2**

Table 8: Relationships between the superficial soil moisture at watershed scale and the processed SAR signal in the cases of the Grand Morin, Serein, Petit Mori, and Orgeval basins.

Then, the data acquired during the following year(s) will be directly inverted using the calibrated relationship. Table 8 summarises, for the different watersheds, the relationships between the processed SAR signal values $\tilde{\sigma}_0$ and the superficial soil moisture W_s obtained at the end of the calibration year.

- *Operational year results*

Classification results

During the operational period, each image is processed once it is acquired. Therefore, the number of ERS images available for the classification increases with time: one at the very the beginning of the (agricultural) year, then two at the date of the second ERS acquisition, etc... In summary, with this approach, the classification is brought up to date and improved all along the year, at each reception of a new ERS data.

From the results obtained during the calibration year for both the Grand Morin and the Serein sites, monosource classifications, LANDSAT multispectral classification or ERS multitemporal classification, allows the identification of the most significant land cover types. Minority land cover type identification is more difficult.

However, for the purpose of our hydrological study, the identification of the main land cover types, first in terms of vegetation density and then in terms of vegetation types, is sufficient. Thus, during winter months, we only need to distinguish between the forest areas, the grasslands (if there are), and the bare soil fields (which will be used for soil moisture index derivation).
From the beginning of March, the winter cereals (such as wheat and barley) fields have to be identified (since some vegetation effect corrections should be applied to these pixels).
Then, as the other crops appear and grow, they have to be identified and their fields removed from the soil moisture index computation.
Finally, in May and June, the only fields, which remain at bare soil state, are the cornfields.

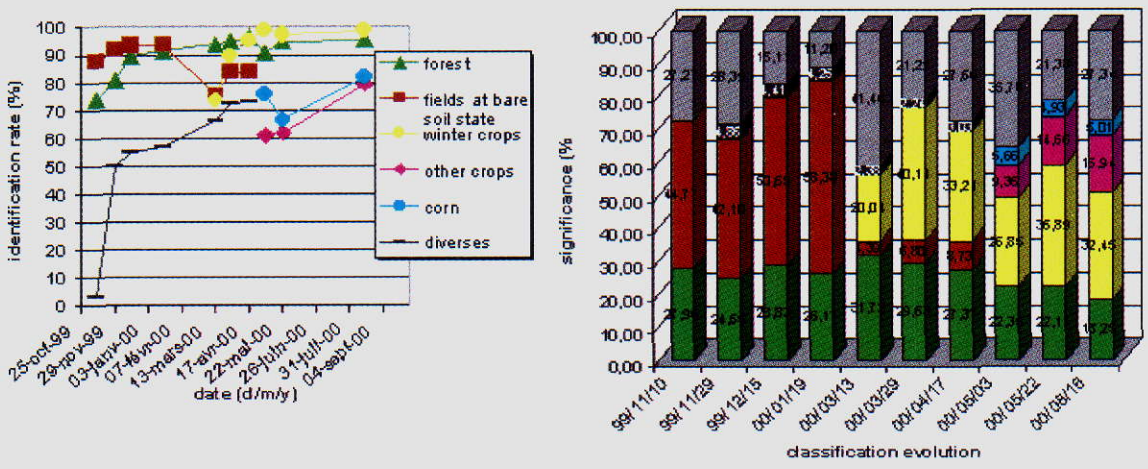


Figure 16: Evolution of the ERS classification result during the 1999-2000 agricultural year in case of the Grand Morin.

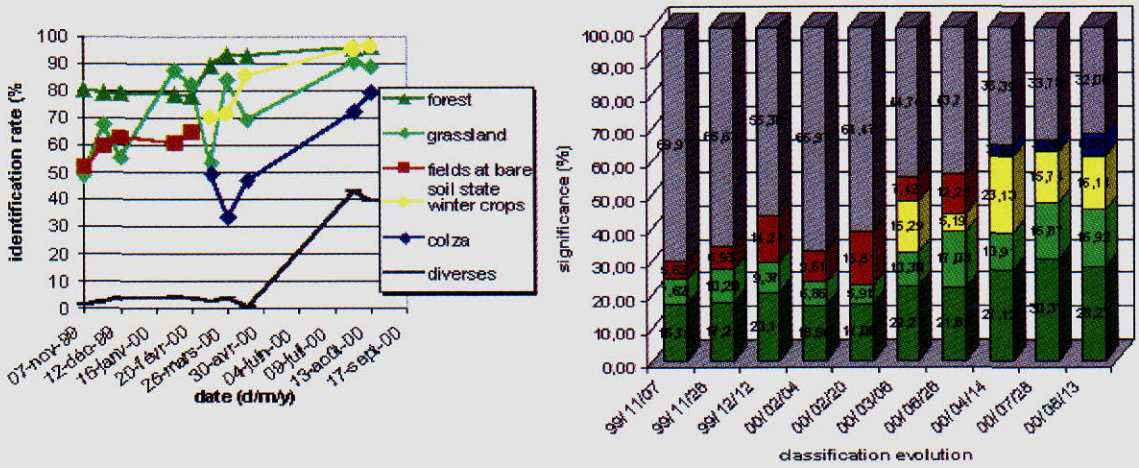


Figure 17: Evolution of the ERS classification result during the 1999-2000 agricultural year ,for the Serein.

Figures 16 and 17 show the evolution of the land cover type identified and their identification rates during the agricultural year, respectively in the case of the Grand Morin and the Serein watersheds. For both watersheds, we note the increase of the identification rates during the year. Until February, we only distinguish between forest, bare soil (fields), and grassland (present only in the case of the Serein). Then, the bare soil class is divided between a winter cereal class and a class corresponding to the remaining bare soil areas. Thus, winter crops are only identified from March as well as colza in the case of the Serein, and for the Grand Morin corn is separated from the other crops only from May. For every land cover types (excepting "diverses" which is a mixing class by definition), identification rates greater than 80 % are achieved during summer. About the land cover significance; we note the decrease along the year of the "unidentified class" (in gray). The bare soil class present during the first months then subdivided between winter crops and remaining bare soil fields, itself then sub-dividing between different crops

Soil moisture indices

In the case of the AIMWATER database, the operational period lasts one year and five months (from November 1999 to April 2001). Figure 18 shows the soil moisture index estimated from SAR/ERS data versus the soil moisture ground truth measurements, for the four studied watersheds: the Grand Morin, the Serein, the Petit Morin, and the Orgeval. In all cases, the points (or ERS/SAR acquisition dates) corresponding to the calibration year, and those corresponding to the operational period have been plotted separately. The corresponding linear regressions and their equations are shown.

Focusing on the operational period, we see that, in the case of the Grand Morin watershed, the results are very satisfactory since the slope is equal to 1, with nearly zero offset, and this, with a good correlation, namely R^2 greater than 0.85. In the case of the Serein, the correlation remains high: R^2 equal to 0.85, but the slope is slightly less satisfactory: 0.85, and there is a 10 % bias (offset). It should be noted that both of these effects were already observable (in a lower proportion for the offset) during the calibration period. Now, if we consider only relative values (which the case for hydrological purposes) the bias problem disappears, and, if we do not consider the three worst points, the results are again very satisfactory: slope equal to 0.95 and correlation coefficient $R^2 = 0.91$. In the case of the Petit Morin, the observed slope and offset are satisfactory: respectively 1.1 and less than 2 %, but the correlation is not so good: $R^2 = 0.63$ whereas it was equal to 0.92 during calibration year. However, if the three worst points are removed, R^2 increases up to 0.83. Finally, the case of the Orgeval is very similar to those of the Grand Morin: consistent slope and offset, and satisfying correlation coefficient.

In conclusion, the retrieval of soil moisture values from space and in a rather operational way seems possible. Soil moisture estimation accuracy depends on the SAR noise features. Nowadays, according to SAR/ERS calibration and speckle noise, it can be assumed that the backscattering coefficient error bar is equal to $\Delta\sigma_0 = 1$ dB. Then, the soil moisture error bar is $\Delta\sigma_0 \times g'(\sigma_0)$, where g is the $(W_s, \tilde{\sigma}_0)$ relationship. In the present case of linear approximation $g'(\sigma_0)$ is constant and ranges from 2.5 to 2.9; inducing a W_s accuracy better than 3 %.

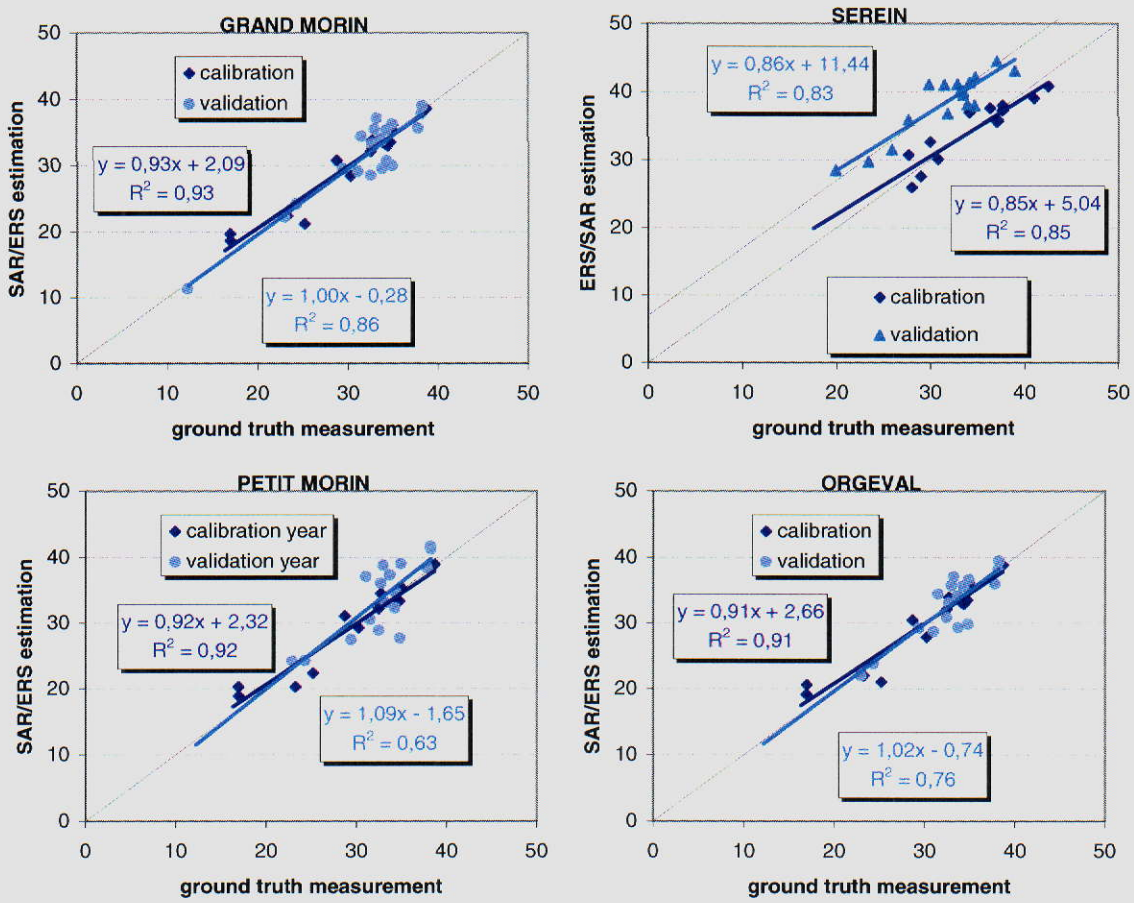


Figure 18: Soil moisture index estimated from SAR/ERS data versus the soil moisture derived from ground measurements, for: a) the Grand Morin, b) the Serein, c) the Petit Morin, and d) the Orgeval. The points corresponding to the calibration year, and those corresponding to the operational period have been plotted and fitted separately.

The soil moistures indices are available at the ERS/SAR acquisition dates excepting during the dense vegetation period. Figure 19 shows, respectively for the Grand Morin, the Serein, and the Petit Morin watersheds (the Orgeval, which is a sub-basin of the Grand Morin, and therefore has a very close behaviour, has not been represented) the evolution, from January 1999 to April 2001, of the processed SAR signal values $\tilde{\sigma}_0$, and the corresponding soil moisture indices $\tilde{W}_s = g(\tilde{\sigma}_0)$.

Finally, we check the usefulness of the target selection. Figure 20 shows a comparison between the processed SAR signal and the average of the backscattering coefficient over the whole watershed, or the average only excluding forest and grassland areas in the case of Serein. The points corresponding to the sparse vegetation period clearly appear as the points farthest from the bisector (straight line). They generally correspond to an underestimation of the soil backscattering coefficient (estimated by the processed SAR value), due to the attenuation effect of the vegetation cover. The points close to the straight line correspond to the bare soil period points. They are slightly closer when forest and grassland areas are not considered in the average. In conclusion, this brief study clearly illustrates the loss of sensitivity when including these areas in the estimation of soil moisture index

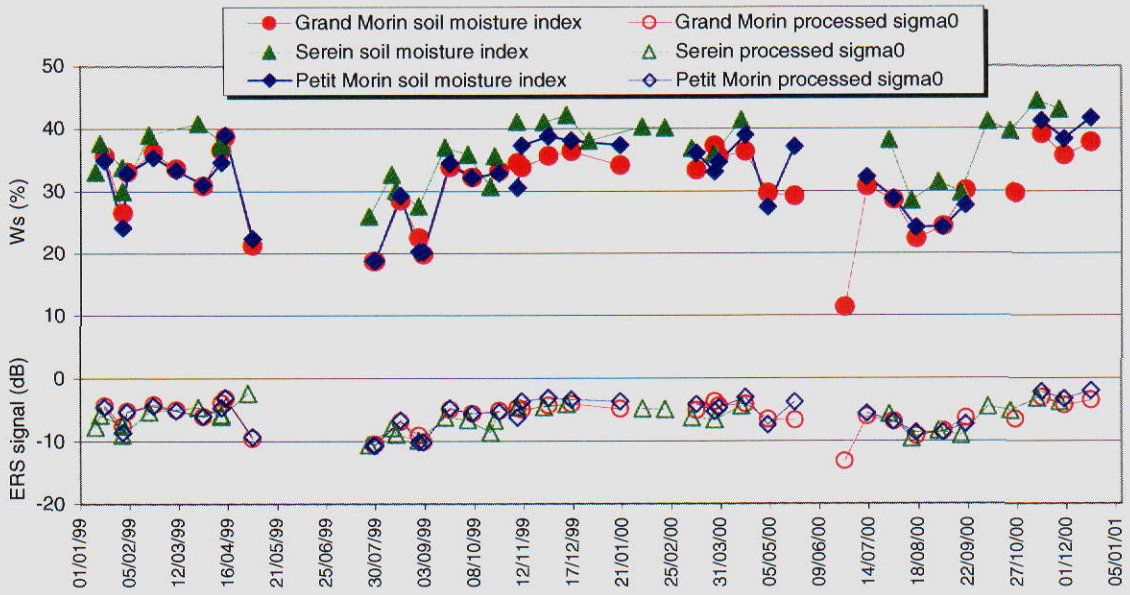


Figure 19: 2.5-year evolution of watershed soil moisture indices respectively over the Grand Morin, the Serein, and the Petit Morin catchments.

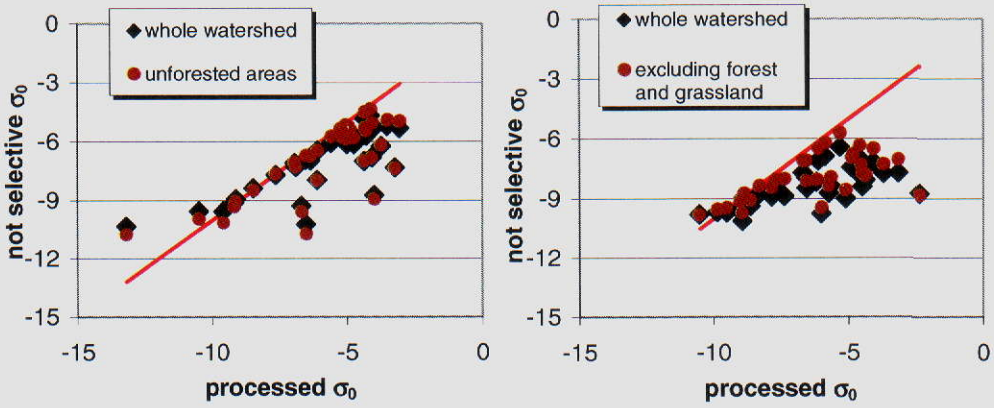


Figure 20: Comparison of the processed SAR signal with the average of the backscattering coefficient over the whole watershed and with the average excluding forest and grassland areas; case of the Serein.

This study shows the possibility to monitor surface soil moisture from space from ERS/SAR data. During the calibration period (first year), the used methodology was *proposed* by (Quesney *et al.*, 2000). Then, for the following years, it is slightly modified to make it more operational. Indeed, if it is desired to use soil moisture indices effectively to improve hydrological model simulations, one must be able to derive them in an operational way. We use a methodology, which needs a rather large amount of ground truth data and delays for the calibration period, and then, we have another methodology less constraining. During the operational years, the methodology is modified in order (i) to avoid having to carry out a ground truth campaign, (ii) to derive soil moisture indices immediately after SAR image acquisition. To achieve these purposes, first, we develop an empirical crop growth model, which is valid only during vegetation growth and which relates the stalk and leaf geometrical dimensions to crop total height. Second, we modify the classification process and the searched classes according to the date of the considered SAR image (e.g. no distinction between crop still in the bare soil state).

In the framework of the AIMWATER project, it has been tested on three different watersheds, and a sub-basin. These basins exhibit different soil types and compositions. They are all agricultural northern European areas, but the percentages of the different crop types are different. Thus, they allow the robustness of the empirical results to be studied (method and relationship between the processed backscattering coefficient and the surface soil moisture at watershed scale).

• The Arade basin (Portugal)

- Database

- A pair of optical images, from SPOT and LANDSAT satellites.
- Time series of ERS-SAR images.
- Time series of volumetric soil moisture content measured by TDR (Time Domain Reflectometry) sensor, to find the relationship between the backscattering coefficient measured by the radar and the measured soil moisture.
- Gravimetric moisture content measurements to calibrate the TDR instrument, and soil bulk density in the different measurement points.
- Vegetation parameters such as plant height, leaf size, stem thick, plant water content, etc., used as input parameters for the radar backscattering model.
- A DEM with 8 meters of resolution, used for pre-processing of the remote sensing images and the elaboration of the classification map and slope/orientation maps.

The preprocessing of optical and radar data has been described with detail previously (see point 3.1).

♦ Soil moisture measurements

To elaborate the soil moisture map it was necessary to find a relation between the signal measured by the radar and the soil moisture. For this reason a temporal series of ground moisture measurements were taken in several representative points of the basin. These measurements were taken with a portable TDR that obtains instantaneously volumetric ground moisture, but before using these data it is necessary to calibrate of the measurements for the type of soil in our study area. The TDR sensor measures the dielectric constant of the soil, and the relation between the value of the dielectric constant of a soil and its water content depends on the chemical composition of that soil, what makes necessary a local calibration of the TDR signal.
For the TDR calibration, soil samples were taken in the measurement points. These samples were weighted and then put immediately into an oven 24 hours at 105 oC. The dry sample is weighted again, and the difference between the wet and dry weights gives the soil moisture content. This value is divided by the dry weight and then multiplied by the bulk density of the sample in order to calculate the volumetric moisture content, which can be compared to the measurement obtained with the TDR. Plotting the soil moisture values obtained with TDR versus the ones obtained with the soil samples we derive the linear regression for the TDR calibration.
Once the TDR is calibrated, it is possible to use the soil moisture measurements obtained with the TDR to derive the relation between the soil moisture and the backscattering signal measured by the radar.

- Extraction of Parameters

As inputs to the AIMWATER hydrological model, we had elaborated, from all the data available, maps with a resolution of 1 km, as required by the models, even if we worked initially with a resolution of 8 meters. The maps elaborated were a temporal series of moisture maps, land-cover classification map, elevation map and slope and orientations map. The low resolution of the final products gives us some capability to adopt approximations, taking into account that we were working with 8 meters resolution to produce outputs with 1 km resolution.

Land cover classification

To produce the classification map, we had used the three bands of SPOT, and the equivalent bands in the case of LANDSAT, as well as the elevations and slopes and orientations derived from the DEM.
The main soil uses within the Arade area are forest and bushes/herbaceous, besides there is some agricultural exploitation on the river valleys. Then, with the classification we tried to discriminate the next classes: forest,

bushes/herbaceous, crop soil, bare soil, and water surfaces. For this reason we used a supervised method to carry out the classification, and we selected the minimum distance method for simplicity and robustness.
Two classifications have been derived: a winter land-use classification with the SPOT image and another for the summer with the LANDSAT image. The resulting classifications confirm what we had advanced, the main soil use in the study area is the forest and bushes/herbaceous with some areas of agricultural exploitation. Comparing both classifications we can see that there is not an important difference between the land use in winter and in summer. There are only few more crop areas and bare soil in summer, as expected.
The fact that satellite images are georeferencered as result of the geometric correction with de DEM allows the validation of the classification. Taken a series of pictures in several coordinates over the basin we can locate the placements of the pictures over the classification and compare the true land cover use reflected in pictures with the information contained in classification. In the next figure we can see a pair of examples of these pictures and their corresponding area in the land cover classification. The first picture shows a bare soil area surrounded of some bushes/herbaceous vegetation with a bite of forest, which is confirmed in the classification corresponding to this placement. Second picture is a small crop area surrounded by forest mixed with bushes/herbaceous vegetation with small areas of bare soil, the same information given by the land cover classification. We can conclude that the pictures taken validate the information reflected in the land cover classification.

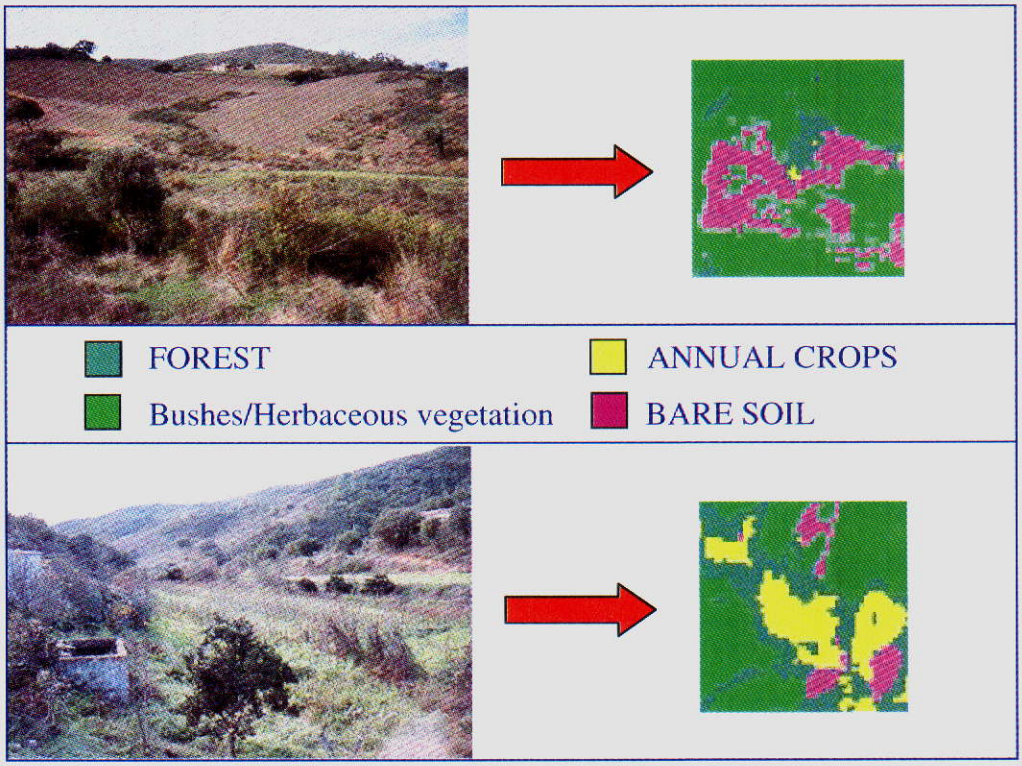


Figure 21.- Validation of the land cover use classification by taking pictures over the basin..

Soil moisture maps

♦ Relation between soil moisture and radar backscattering

To derive the soil moisture maps the most important is to find a proper relationship between the measured soil moisture and the backscattering signal measured by the radar. This relationship is supposed to be lineal for a range of soil moisture values. We derived the relationship with the radar images without vegetation compensation because the ground measurement points where soil moisture was measured were selected by their low vegetation cover, with no vegetation effects in the radar signal for the ground measurement points.

To obtain the relationship we have used the calibrated moisture measurements from the TDR in the different points of the basin. We averaged these measurements obtaining a moisture value for all the basin in a date as nearest as possible to the date of acquisition of each one of the SAR images. Since we know the coordinates of the measurement points, we can locate them within the radar images thanks to the geometric correction. Then we averaged the SAR pixels around each ground measurement point obtaining a backscattering value in dB from the image. The averaged moisture value is plotted versus the dB value. By doing this for the different SAR images we obtained a series of points that cover a range of moisture and their corresponding dB values. A linear regression of these points (see Figure 22) is derived and then used to obtain the moisture map from the ERS-SAR image:

$$\sigma_{dB} = 0.1687 \cdot V_{moist} - 12.553$$

where V_{moist} is the value of volumetric soil moisture and σdB is the backscattering measured by the radar.

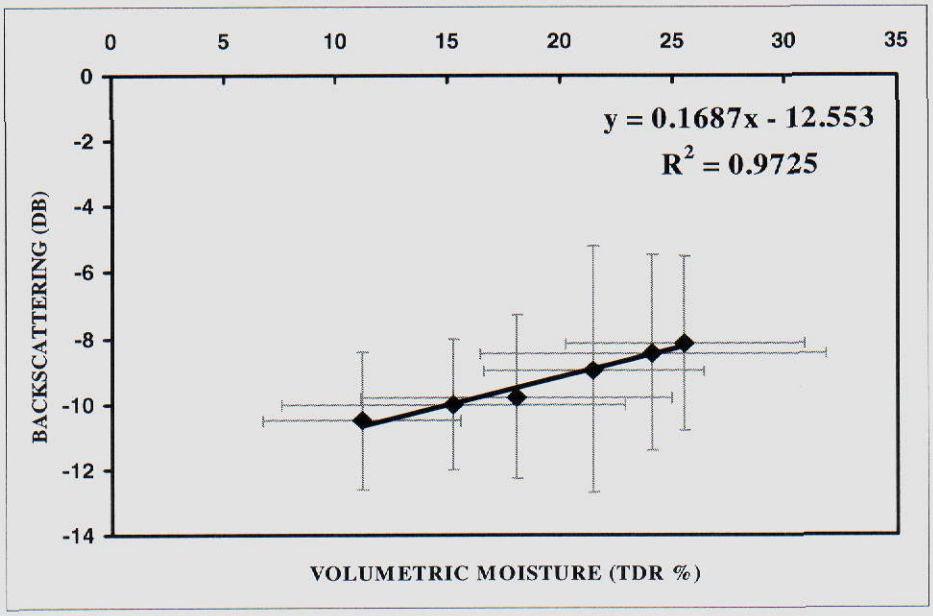


Figure 22.- Relationship between soil moisture measurements and the backscattering signal measured by the radar.

If we compare our derived linear expression with the ones obtained in other study areas, such as for example the other basins involved in the AIMWATER project, we can see that our slope is relatively lower than the others (Figure 23). This difference is due to the difference of soil type. The soil type is different in the North French basins and in our area in South Portugal. Nevertheless we tested the results obtained by running a theoretical backscattering model for a series of typical soils.

The theoretical backscattering model used here uses as input soil moisture, soil roughness, vegetation parameters such as plant height, leaf size, Leaf Area Index (LAI), stem thick, plant water content, etc., and returns as output the corresponding contributions to the backscattering value. In this case we tried to simulate with the model the conditions of our actual measurement points, analysing the mutual influence of soil moisture and soil roughness. By running the model for different soil moisture and roughness values we obtained several theoretical relations between soil moisture and backscattering signal. The slope of these relations were very similar to the slope of the experimental relationship (Figure 24), so we can conclude that our relationship is correct and the difference with the other areas can be perfectly attributed to the differences of soil type.

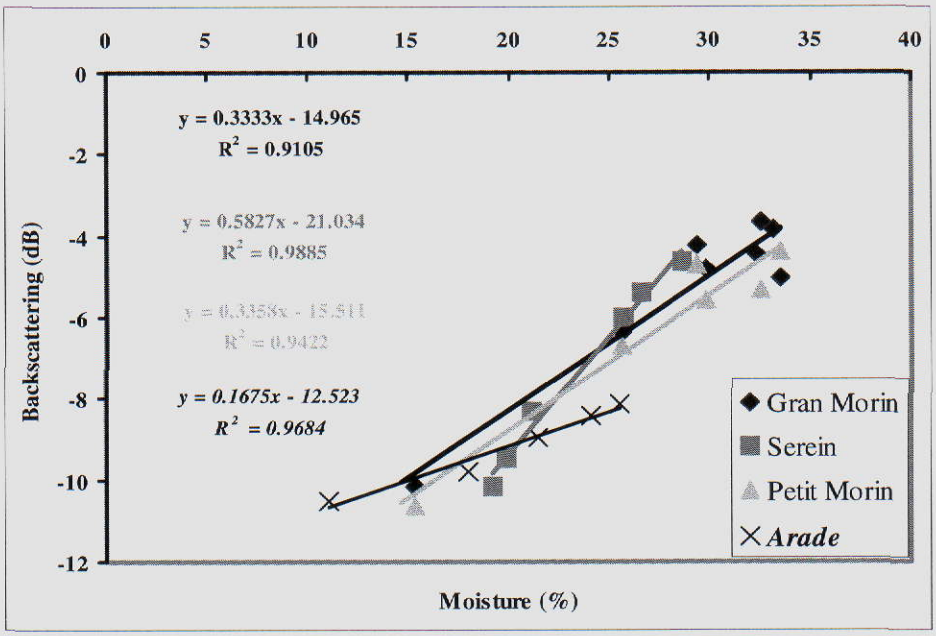


Figure 23.- Relationships between the soil moisture and backscattering measured by radar developed for the French basins involved in the AIMWATER project.

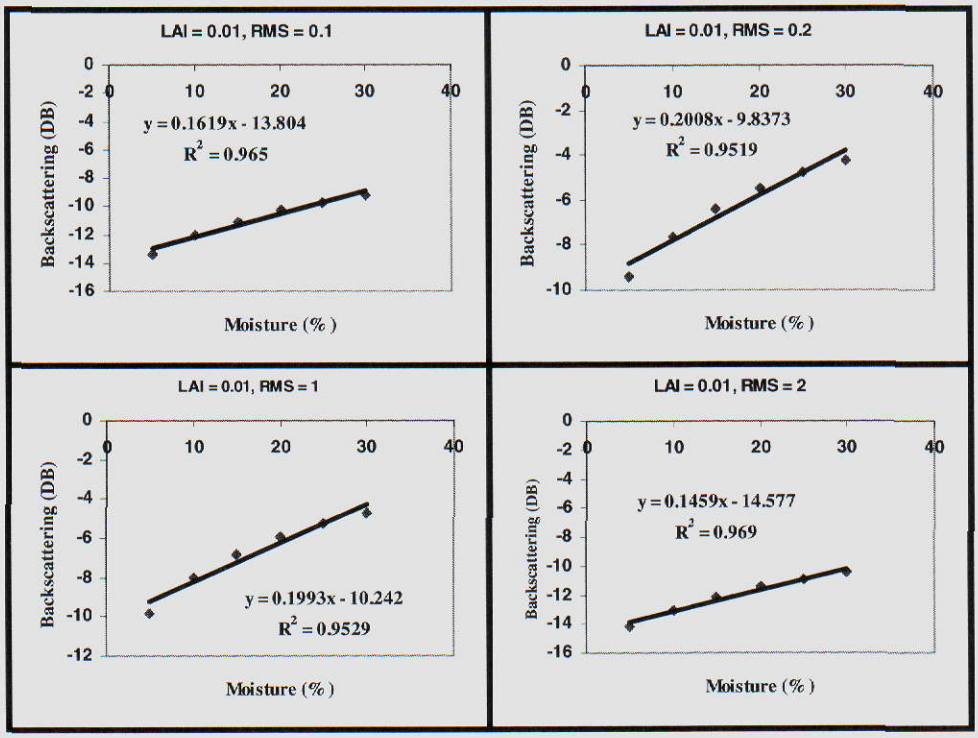


Figure 24.- Outputs of the theoretical backscattering model for the validation of the experimental relation between the backscattering measured by the radar and the soil moisture. The RMS value is a parameter used to define the soil roughness properties in the model.

♦ Compensation of vegetation effects derived from a theoretical backscattering model

Before applying the derived expression to get the soil moisture from the radar image, vegetation effects must be compensated. According to the radiative transfer model we used, the signal measured by the radar σ_{meas} is the sum of the signal reflected by the vegetation layer σveg and the signal reflected by the soil σsoil. This last component is attenuated when it crosses the vegetation layer before and after reflected by the soil surface. Thus, the vegetation effect is double: an attenuation of the soil signal and an intrinsic contribution (Karam et al., 1992):

$$\sigma_{meas} = \sigma_{veg} + \sigma_{soil} \cdot e^{-2\tau}$$

where τ is the optical thickness of the vegetation layer. The correction of this effect is only possible if we know the values of σ_{veg} and τ, and here is where we make use of the optical data and the theoretical backscattering model. The applied correction is based on the following expression:

$$\Delta\sigma(dB) = \sigma_{soil}(dB) - \sigma_{meas}(db) = \frac{20 \cdot \tau}{\ln 10} + 10 \cdot \log_{10}\left[1 - \frac{\sigma_{veg}}{\sigma_{meas}}\right]$$

The model that we had used gives, as a function of the inputs, the backscattering values for each one of the different layers considered: the total value (σ_{meas}), the contribution of the soil after crossing the vegetation layer ($\sigma_{soil} \cdot e^{-2\tau}$) and the vegetation contribution (σ_{veg}) (Ferrazzoli, 1995). We gave as input parameters to the model the vegetation parameters included in the database, and run the model for different LAI values to obtain an expression to relate the values of σ_{veg} and τ with the LAI. This is the key aspect to compensate for the vegetation effects. With the knowledge of the corresponding LAI for each one of the pixels in the radar image we can compensate vegetation effects by applying the previous expression. It is in this point where we make use of a synergistic approach to use radar and optical data simultaneously. From the optical data we derived the NDVI (Normalised Difference Vegetation Index) vegetation index, related to the LAI, so that we can calculate in this way the LAI value for each one of the pixels in the radar image. The relation between NDVI and LAI can be expressed as (Carlson, 1997):

$$NDVI = A - B \cdot e^{-C \cdot LAI}$$

where A = 0.78, B = 1.3 y C = 0.65 are related to the type of vegetation present in the area (Asrar, 1989) (Baret and Guyot, 1991).

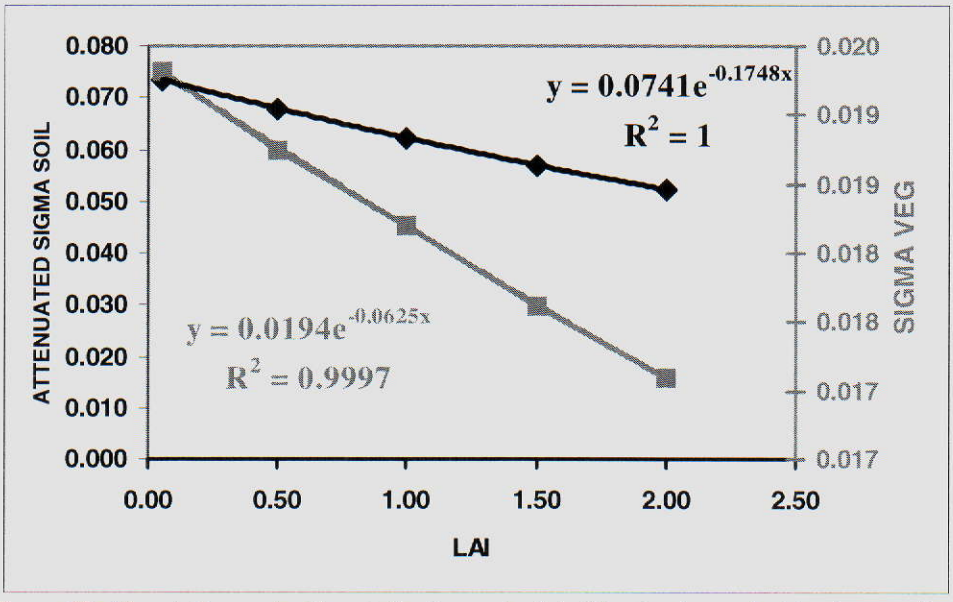


Figure 25.- Vegetation and soil contributions calculated by the theoretical backscattering model in function of the LAI index to compensate the vegetation effect in the signal measured by the radar.

After running the model for different LAI values we have obtained the following expressions to calculate σ_{veg} and τ (see Figure 25):

$$\sigma_{veg} = 0.0194 \cdot e^{-0.0625 \cdot LAI}$$

$$\sigma_{soil} = 0.0741 \cdot e^{-0.1748 \cdot LAI} = \sigma_{xoil} \cdot e^{-2 \cdot \tau} \Rightarrow \tau = 0.0874 \cdot LAI$$

With these expressions we have compensated the effects of vegetation in radar images and then we have applied the relation between radar backscattering and soil moisture obtained previously to calculate finally the temporal series of soil moisture maps.

Elaboration of the inputs for the hydrological model

We have worked with a resolution of 8 meters by pixel in the processing of all remote sensing data and in the original elaboration of the inputs for the AIMWATER hydrological model: temporal series of moisture maps, land cover classification, elevation map and slope and orientations map. But the hydrological model really do not work in this resolution so high. The model works with inputs at a resolution of 1 km by pixel, so we have resampled our final products at this resolution. Figure 26 shows a soil moisture map and the land cover classification at the resolutions of 8 meter by pixel and the corresponding inputs for the model with 1 km by pixel. To work with this low final resolution has allowed us some degrees of freedom in the approximations adopted during the pre-processing of the data, that, although relevant at 8 m resolution, are not significant at 1 km resolution.

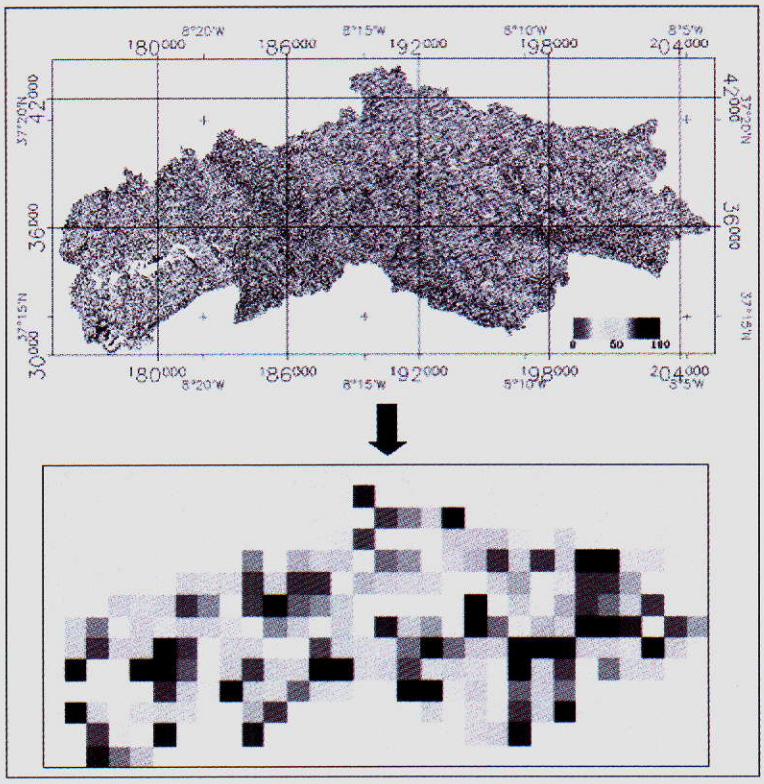
- Conclusions

With the objective of deriving soil moisture indicators to be used as inputs in hydrological models adapted to help the management of multipurpose reservoirs in the Arade area in South Portugal, a full methodology has been developed that makes use of SAR, optical and DEM data in a synergistic way.

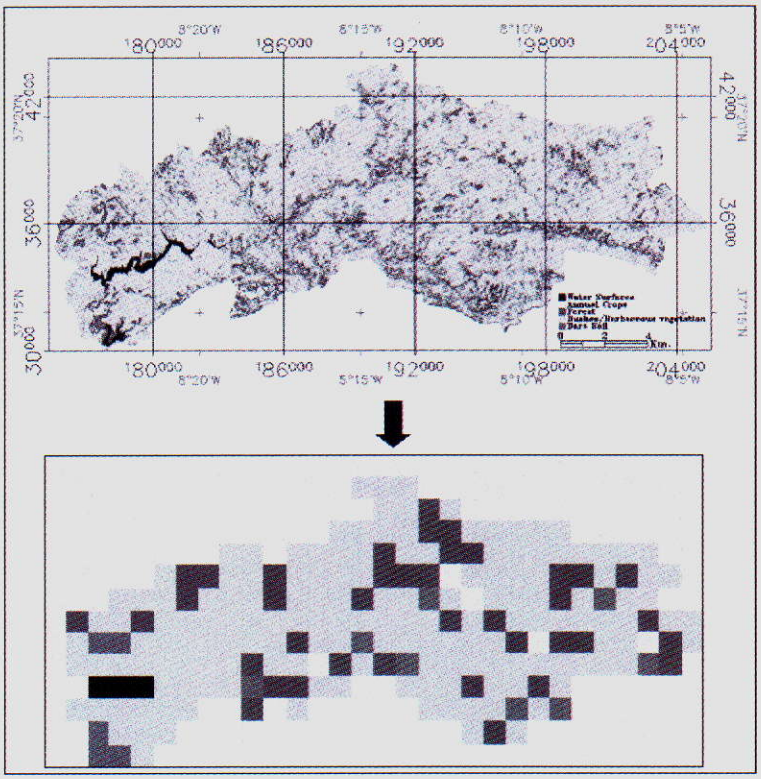
A relationship between the soil moisture and the SAR backscattering signal for our study area has been derived by using a time series of ERS-2 SAR images and ground measurements in several reference sites. This relation is different in slope and intercept from the one obtained for the French basins involved in AIMWATER project. We have validated the derived relationship by comparison with a theoretical backscattering model, and the difference in the relationships is attributed to the different type of soils present in each one of the different areas.

Due to the need for compensation of vegetation effects, synergistic use of optical and radar data is used to derive the soil moisture maps. By means of a theoretical backscattering model we have derived expressions relating the LAI (obtained by means of an optical vegetation index) with the contribution of vegetation in the radar signal, to apply the vegetation compensation for each one of the pixels within the radar images. Due to the low temporal variation that the vegetation presents in the study area we only used two optical images for the compensation of vegetation effects in the ERS time series of images.

The input parameters for the AIMWATER hydrological model: soil moisture maps, land classification, elevation map and slope and orientation map, have been resampled to the required resolution of 1 km by pixel as the final remote sensing products for the Arade basin.



(a)



(b)

Figure 26.- (a) First moisture map at the working resolution (8m) and at the resolution of 1km as input to the hydrological model. (b) Same for the SPOT land cover classification.

Origin: U. Valencia/CETP/Cemagref
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

References

Autret, M., Bernard, R., Vidal-Madjar, D., 1989, "Theoretical study of the sensitivity of the microwave backscattering coefficient to the soil surface parameters", International Journal of Remote Sensing, vol.10, no.1, pp.17-179.

Asrar G., Myneni R.B., Konemosh E.T., 1989. *Estimation of plant-canopy attributes from espectral reflectance measurements*. Theory and applications of optical remote sensing, Asrar G. (Ed), 252-292.

Baret F. and Guyot G., 1991. *Potentials and limits of vegetation indices for LAI and APAR assessment*. Remote Sensing of Environment, 35:161-173.

Bezdek, J.C., Ehrlich, R., and Full, W., 1984, "FCM: the fuzzy c-means algorithm", *Computer and Geoscience*, **10**(2-3):191_203.

Beaudoin, A., Le Toan, T., Gwyn, Q.H.J., 1990, "SAR observations and modeling of the C-band backscatter variability due to multiscale geometry and soil moisture", IEEE Transactions on Geoscience and Remote Sensing, vol.28, no.5, pp.886-895.

Carlson T.N., Ripley D., 1997. *On the relation between NDVI, fractional vegetation cover, and Leaf Area Index*. Remote Sensing of Environment, 62:241-252.

Choudhury, B.J., Schmugge, T.J., Chang, A., Newton, R.W., 1979, "Effects of surface roughness on the microwave emission from soils", Journal of Geophysical Research, vol.84, pp.5699-5706.

Cognard, A.L., 1996, "Suivi de l'état hydrique des sols par télédétection spatiale (radar et thermographie infrarouge) et modélisation hydrologique à l'échelle du bassin versant", Thèse de Sciences, Université Paris XI Orsay, France.

Cognard, A.L., Loumagne, C., Normand, M., Olivier, P., Ottlé, C., Vidal-Madjar, D., Louahala, S., Vidal, A., 1995, "Evaluation of the ERS 1/synthetic aperture radar capacity to estimate surface soil moisture: two-year results over the Naizin watershed", Water Resources Research, vol.31, no.4, pp.975-982.

Dubois, P.C., van Zyl, J., Engman, T., 1995, "Measuring soil moisture with imaging radars", IEEE Transactions on Geoscience and Remote Sensing, vol.33, no.4, pp.877-895.

Engman E.T., 1996, Remote sensing applications to hydrology : future impact. Hydrol. Sc. Bull. 41(4), 637-647.

Entekhabi, D., Nakamura, H., Njoku, E.G., 1994, "Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations", IEEE Transactions on Geoscience and Remote Sensing, vol.32, pp.438-448.

ESA, 2001. *Calibration of the ERS SAR PRI product*. http://earth.esa.int/.

Evans, D.L., Jeffrey, J.P., Stofan, E.R., 1997, "Overview of the Spaceborne Imaging Radar-C/ X-band Synthetic Aperture Radar (SIR-C/ X-SAR) missions", Remote Sensing of Environment, vol.59, no.2, pp.135-140.

Ferrazzoli P., Bracaglia M., Guerriero L., 1995. *A Fully Polarimetric Multiple Scattering Model for Crops*. Remote Sensing of Environment, 54:170-179.

Fung, A.K., Li, Z., Chen, K.S., 1992, "Backscattering from randomly rough dielectric surface", IEEE Transactions on Geoscience and Remote Sensing, vol.30, no.2, pp.356-369.

Geman, S., and Geman, D., 1984, "Stochastic relaxation, Gibbs distribution and Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**(6):721-741.

Griffiths, G.H., and Wooding, M.G., 1996, "Temporal monitoring of soil moisture using ERS-1 SAR data", Hydrological Processes, vol.10, pp.1127-1138.

Jackson, T.J., 1993, "Measuring surface soil moisture using passive microwave remote sensing", Hydrological Processes, vol.7, pp.139-152.

Jackson, T.J., and Le Vine, D.E., 1996, "Mapping surface soil moisture using an aircraft-based passive microwave instrument: algorithm and example", Journal of Hydrology, vol.184, pp.85-99.

Jackson, T.J., and Schmugge, T.J., 1991, "Passive microwave sensing of soil moisture under vegetation canopies", Water Resources Research, vol.18, pp.1137-1142.

Jackson, T.J., Schmugge, T.J., Wang, J.R., 1982, "Vegetation effects on the microwave emission from soils", Remote Sensing of Environment, vol.36, pp.203-212

Karam, M.A, Fung, A.K., Lang, R.H., Chauhan, N.S., 1992, "A microwave scattering model for layered vegetation ", IEEE Transactions on Geoscience and Remote Sensing, vol.30, no.4, pp.767-784.

Kasischke, E.S., Melack, J.M., Dobson, M.C., 1997, "The use of imaging radars for ecological applications - a review", Remote Sensing of Environment, vol.59, no.2, pp.141-156.

Kaufman, Y.J., and Gao, B.C., (1992), Remote sensing of water vapor in the near-IR from EOS/MODIS, IEEE Trans. Geosci. Remote Sensing, 30:871-884.

Kerr, Y. et al., 1997, "RAMSES- Radiométrie Appliquée à la Mesure de la Salinité et de l'Eau du Sol", Réponse à l'appel à propositions « Expériences et missions scientifiques spatiales de taille intermédiaire » du Centre National d'Etudes Spatiales.

Kerr, Y. et al., 1998, "SMOS – Soil Moisture and Ocean Salinity", Answer to the call for Earth explorer opportunity missions of the European Space Agency.

Kite G.W. and Pietroniro A., 1996, Remote sensing applications in hydrological modelling. Hydrol. Sc. Bull. 41(4), 563-591.

Kuusk, A. (1994), A multispectral canopy reflectance model, Remote Sensing Environ., 50:75-82.

Kuusk, A., (1991), Determination of vegetation canopy parameters from optical measurements, Remote Sensing Environ., 37:207-218.

Le Hégarat-Mascle, S., Bloch, I., Vidal-Madjar, D., 1997, "Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing", *IEEE Transactions on Geoscience and Remote Sensing*, **35**(4):1018-1031.

Le Hégarat-Mascle, S., Quesney, A., Vidal-Madjar, D., Taconet, O., Normand, M., Loumagne, C., 2000, "Land cover discrimination from multitemporal ERS images and multispectral LANDSAT images: a study case in an agricultural area in France", *International Journal of Remote Sensing*, **21**(3):435-456.

Le Hégarat-Mascle, S., Alem, F., Quesney, A., Normand, M., Loumagne, C., 2000, "Estimation of watershed soil moisture index from ERS/SAR data", *Proceedings of EUSAR2000*, in Munich, Germany, on May 23-25, 2000, pp.679-682.

Le Vine, D.M., Wilheit, T.T., Murphy, R., Swift, C.T., 1989, "A multifrequency microwave radiometer of the future", IEEE Transactions on Geoscience and Remote Sensing, vol.27, pp.193-199.

Mather, P.M., 1999, "Computer Processing of Remotely-Sensed Images". Ed John Wiley & Sons.

Meyer, P., 1993, "Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment", ISPRS Journal of Photogrammetry and Remote Sensing, **48**(4): 17-28.

Mo, T., Schmugge, T.J., Wang, J.R., 1987, "Calculations of microwave brightness temperature of rough soil surfaces: bare field", IEEE Transactions on Geoscience and Remote Sensing, vol.25, pp.47-54.

Moran, M.S., Vidal, A., Troufleau, D., Inoue, Y., Mitchell, T.A., "Ku- and C- band SAR for discriminating agricultural crop and soil conditions", IEEE Transactions on Geoscience and Remote Sensing, vol.36, no.1, pp.265-272, 1998.

Moreno, J., "Spectral/spatial integration effects on information extraction from multispectral data: multiresolution approaches", European Symposium on Satellite Remote Sensing, Rome, Italy, 26-30 September 1994, EOS/SPIE, vol. 3214, pp. 324-338, 1995.

Njoku, E.G., and Entekhabi, D., 1996, "Passive microwave remote sensing of soil moisture", Journal of Hydrology, vol.184, pp.101-129.

Oh, Y., Sarabandu, K., Ulaby, F.T., 1994, "An inversion algorithm for retrieving soil moisture and surface roughness from polarimetric radar observation", IGARS'94, pp. 1582-1584.

Pampaloni, P., and Paloscia, S., 1986, "Microwave emission and plant water content: a comparison between field measurements and theory", IEEE Transactions on Geoscience and Remote Sensing, vol.24, pp.900-905.

Quesney, A., Le Hégarat-Mascle, S., Taconet, O., Vidal-Madjar, D., Wigneron, J.P., Loumagne, C., Normand, M., 2000, "Estimation of watershed soil moisture index from ERS/SAR data", *Remote Sensing of Environment*, **72**(3):290-303.

Rakotoarivony, L., Taconet, O., Vidal-Madjar, D., 1996, "Radar backscattering over agricultural bare soils", Journal of Electromagnetic Waves and Applications, vol.10, no.2, pp.187-209.

Schmugge, T.J., Jackson, T.J., Kustats, W.P., Wang, J.R., 1992, "Passive microwave remote sensing of soil moisture: results from HAPEX, FIFE and MONSOON'90", Journal of Photogramm. Remote Sensing, vol.47, pp.127-143.

Schmullius, C.C., and Evans, D.L., 1997, "Synthetic aperture radar (SAR) frequency and polarization requirements for applications in ecology, geology, hydrology, and oceanography: a tabular status quo after SIR-C/X-SAR", International Journal of Remote Sensing, vol.18, no.13, pp.2713-2722.

Swift, C.T., 1993, "ESTAR- the Electronically Scanned Thinned Array Radiometer for remote sensing measurement of soil moisture and ocean salinity", NASA Techn. Memo., 4523. NASA, Washington, DC.

Taconet, O., Vidal-Madjar, D., Emblanch, C., Normand, M., 1996, "Taking into account vegetation effects to estimate soil moisture from C-band radar measurements", Remote Sensing of Environment, vol.56, pp.52-56.

Tsang, L., Kong, J.A., Shin, R.T., 1985, Theory of Microwave Remote Sensing, Wiley, New York.

Ulaby, F.T., Moore, R.K., Fung, A.K., 1986, Microwave Remote Sensing Active and Passive, Eds. Artech House, inc.

Ulaby, F.T., Razani, M., Dobson, M.C., 1983, "Effects of vegetation cover on the microwave radiometric sensitivity to soil moisture", IEEE Transactions on Geoscience and Remote Sensing, vol.21, pp.51-61.

Wang, J., Hsu, A., Shi, J.C., O'Neil, P., Engman, T., 1997, "Estimating surface soil moisture from SIR-C measurements over the Little Washita River watershed", Remote Sensing of Environment, vol.59, no.2, pp.308-320.

Wang, J.R., and Choudhury, B.J., 1995, ., "Passive microwave radiation from soil: examples of emission models and observations", in B.J. Choudhury, Y.H. Kerr, E.G. Njoku and P. Pampaloni (Editors), Passive Microwave Remote Sensing of Land-Atmosphere Interactions, VSP Publishing, Utrecht.

Wang, J.R., Schmugge, T.J., Shiue, J.C., Engman, E.T., 1989, "Mapping surface soil moisture with L-band radiometric measurements", Remote Sensing of Environment, vol.27, pp.305-312.

Wigneron, J.P., Chanzy, A., Calvet, J.C., Bruguier, N., 1995, "A simple algorithm to retrieve soil moisture and vegetation biomass using passive microwave measurements over crop fields", Remote Sensing of Environment, vol.51, no.3, pp.331-341.

4. ASSIMILATION METHODOLOGIES FOR HYDROLOGICAL MODELS

As a prerequisite to the implementation of the best suited models for assimilation purposes over the area under study, the first phase of the project was devoted to the selection of the most appropriate models of the acquired set of possible rainfall-runoff modelling approaches. The second phase was focussed on implementing those recommended models over the area under study and investigating the possibilities of linking these models with catchment characteristics. This step was essential before linking model internal state variables with observed ones for assimilation purposes. The last phase of the project was focused on assimilation methodologies and comparing these methodologies in an operational context.

4.1 Model suitability

The selected models were chosen to meet certain criteria e.g.: be operational to satisfy customers' requirements, able to use remote soil moisture data, be reliable in terms of stream flow simulation, be easy to use and to implement and not too demanding in terms of input.

Considering these previous criteria different model structures were recommended in the context of the AIMWATER project. This section describes the criteria used to select rainfall-runoff models, the selected model structures, their hydrological processes and their main differences. Results in terms of simulated stream flow are analyzed and some conclusions are drawn with special attention given to the links between model structure/efficiency and model variables/basin characteristics.

4.1.1 Selection of the most appropriate models

A number of models were selected because they included several characteristics deemed to be important with respect to user needs; the models should be:

(a) already reliable in terms of their results during streamflow simulation and/or forecast to deal with customer requirements,
(b) as simple as possible to come in continuity with the existing modelling tools (if any) of the customers,
(c) easy to use, to understand and to implement for the customers,
(d) not too demanding in terms of input (data) and cost,
(e) able to use the information brought by soil moisture data derived from EO,
(f) able to rely on the expertise and past experience of the modelling group for effective development.

Nine tuneable parameters were believed to be sufficient to describe the rainfall-runoff (R-R) relationship for a wide range of hydrological conditions. Original model structures were thus modified when the initial number of parameters was larger than nine or when the structures were too simple for daily time-steps (several monthly models were tested, with improved routing modules). Care was taken to maintain the originality of each modelling structure. Table 1 lists the 38 models whose performances were analysed, with the number of optimised parameters of the tested version.

This selection brings together a wide sample of the existing mathematical tools used in conceptual or empirical R-R models at the moment. Among the 38 selected are some updated versions of the GR model, which is currently used by the IIBRBS, and a version of the HMS model, which may be included in the PEGASE model for the Seine reservoir management.

All the structures use a soil moisture store (or a comparable device) to take account of antecedent soil moisture conditions, which play a key role on the way effective rainfall is determined.

Models (Authors)	Code	Param.	Models (Authors)	Code	Param.
Abcd (Thomas)	**ABCD**	**6**	*Ihacres (Jakeman et al.)*	*IHAC*	*7*
Arno (Todini)	*ARNO*	*9*	*Blackie and Eeles*	*IHLM*	*9*
Boorman and Bonvoisin (B)	*BOOB*	*6*	*Martine (Thiery)*	*MART*	*7*
Boorman and Bonvoisin (C)	*BOOC*	*5*	*Mhr (Leviandier)*	*MHR0*	*4*
mSFB (Boughton)	*BOUG*	*8*	*Modalp (Arikan)*	*MODA*	*7*
Bucket (Thornthwaite and Mather)	*BUCK*	*6*	*Modglo (Girard)*	*MODG*	*8*
Catpro (Raper and Kuczera)	*CATP*	*8*	*Modhydrolog (Porter and McMahon)*	*MODH*	*9*
Cequeau (Girardet al.)	*CEQU*	*9*	*Nam (Nielsen and Hansen)*	*NAM0*	*9*
Crec (Cormary and Guilbot)	*CREC*	*6*	*O'Donnel and Dawdy*	*ODON*	*9*
Gardenia (Thiery)	*GARD*	*6*	*Pdm (Moore and Clarke)*	*PDM0*	*6*
Georgakakos and Baumer	*GEOR*	*9*	*Sacramento (Burnash et al.)*	*SACR*	*9*
Gr3j (Edijatno et al.)	*GR3J*	*3*	*Sdi (Langford and O'Shaughnessy)*	*SDI0*	*9*
Gr4j (Edijatno et al.)	*GR4J*	*4*	*Sixpar (Gupta and Sorooshian)*	*SIXP*	*8*
Hybrid Gr3j-Gr4j (Edijatno et al.)	*GR4K*	*4*	*Smar (O'Connel et al.)*	*SMAR*	*8*
Gr5j (Ma)	*GR5J*	*6*	*Tank (Sugawara)*	*TANK*	*7*
Grhum (Loumagne et al.)	*GRHU*	*9*	*Tmwam (Bobba and Lam)*	*TMWA*	*8*
Haan	*HAAN*	*8*	*Topmodel (Beven and Kirkby)*	*TOPM*	*7*
Hbv (Bergström)	*HBV0*	*9*	*Wageningen (Warmerdam)*	*WAGE*	*8*
Hms (Morel-Seytoux)	*HMS0*	*9*	*Xinanjiang (Zhaoet al.)*	*XINA*	*8*

Table 1: List of models with authors, code and number of parameters

Country	N° of basins	Origin
Australia	26	Chiew and McMahon (1994)
Brazil	4	Departamento de Eng. Hidraulica, Belo Horizonte (Mato Grosso)
France	307	PLUVIO database (Météo-France, French Ministry for Environment) for rainfall data
		HYDRO database (French Ministry for Environment) for streamflow data
		Météo-France for potential evapotranspiration data
Ivory-Coast	10	ORSTOM - Servat and Dezetter (1992)
United States	37	Mopex, IAHS/WMO/GEWEX, 1999 IAHS Birmingham Symposium
United States	45	ARS database

Table 2: Test catchments

Ranking of models with respect to effectiveness

Models were tested on a sample of 429 catchments (Table 2) (Perrin *et al.*, 2001). This sample includes 1294 calibration periods of three to six calendar years with a whole year of warming-up, and at least two calibration periods for each basin. Using the split-sample test scheme recommended by Klemes (1986), 3204 verification tests were made: for each basin, after calibration for one period, simulation tests were performed on all remaining periods. For example, for a basin with five calibration periods, 20-simulation tests were made. All the models were fed with the same data: daily rainfall and streamflow time-series, and PE data at a daily or 10-day time-step. They were calibrated using a local optimisation technique, called the 'step-by-step' method (Michel, 1989), which optimises the value of an objective function. Five numerical criteria were chosen to assess model performances in calibration and validation modes. Performances in validation mode are more relevant to the actual ability of models to simulate R-R relationship. They show the real worth of the structures and for this reason only the validation results are shown. It is in a mode akin to validation that models are used for practical applications. Figure 1 shows the results with criterion 1 (the Nash and Sutcliffe (1970) criterion calculated on streamflow).
Because one of the potential customers (IIBRBS) is interested in the performance of the model for the Seine river basin, the models were tested on catchments located exclusively in the Seine river basin. Figure 2 shows the results again using criterion 1.

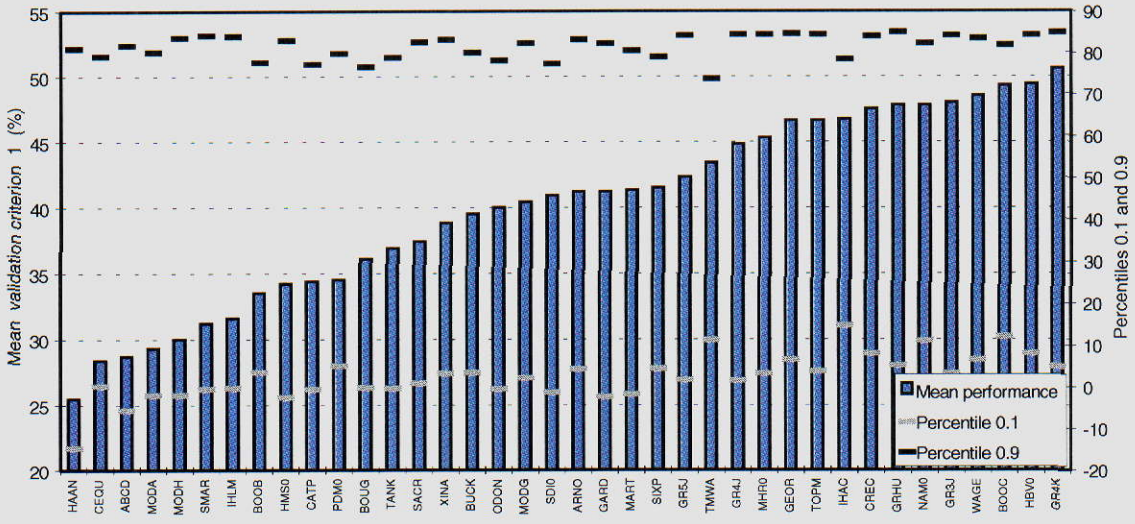


Figure 1: Mean performances in validation (criterion 1).

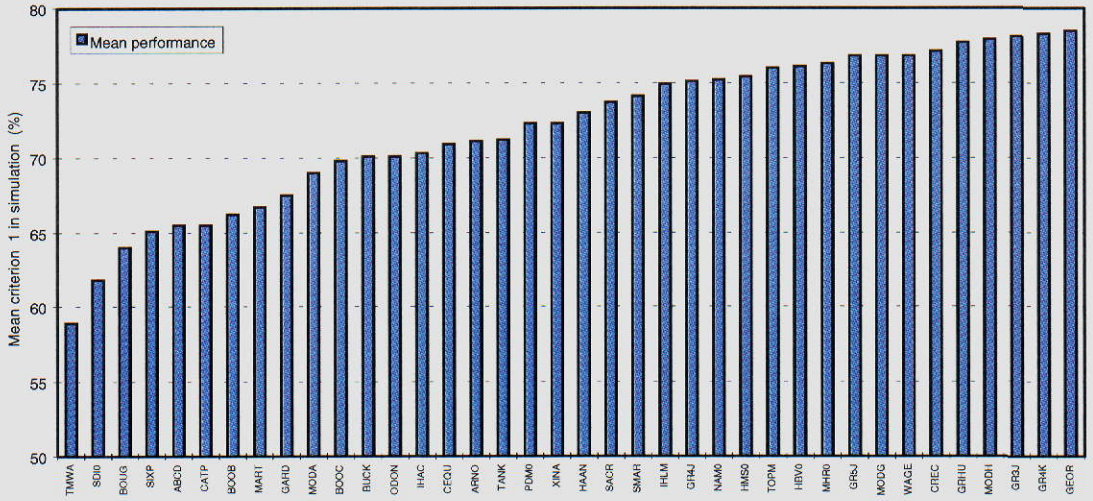


Figure 2 Mean performances in validation (criterion 1) for the Seine river basin.

A few key conclusions can be drawn from this assessment of performances and reliability of rainfall-runoff model structures:

- The majority of the tested model structures performed satisfactorily with some differences associated with their capability to cope with difficult rainfall-runoff relationship in a number of catchments. As expected, performance was better during calibration than simulation. For these catchments, some models are more likely to fail than others, making them less reliable. The more robust models are very valuable for such catchments.
- For most conceptual models, it was possible to find subsets of catchments for which a given model outperformed the others, which means that some models may be more suitable to some conditions than others.
- In calibration mode, complex models having a larger number of parameters achieve the best performances; however this advantage may not hold in simulation mode where models with a low number of parameters yield equally good results.

– Ultimately based on these comprehensive tests and thorough analysis only a few models were considered for comprehensive tests in simulation mode on the basins selected for the project:

- GR models: GR4j model (Edijatno and Michel, 1989) and GRHum (Loumagne *et al.*, 1996):

  A simple version of the GR model is already used by IIBRBS; tests of updated versions provide satisfactory mean performances and reliable results. The simple versions are very parsimonious. The GRHUM model was developed on the basis of GR models, with a modification of soil moisture storage to account for moisture in the topsoil layer, which is essential in the project.

- IHACRES (Littlewood *et al.*, 1997):

  The model yields satisfactory results and proved its robustness on 'difficult' basins. It is parsimonious model, that uses a composite index of antecedent soil moisture conditions and antecedent precipitation, which could be derived from soil moisture data. With only six parameters to optimize it is a relatively easy model to operate.

- TOPMODEL (Beven, 1997):

  Many authors have successfully used this model, which gives reliable results on most test catchments. Simple versions (six or seven parameters) can be used. It is good track record prompted the project to use it as a reference for the assessment of simulation procedures.

- HMS (Morel-Seytoux, 1998):
  This model may be included in the Seine reservoir management model. It has recently been proposed as an alternative to the GR4 model for integration into the PEGASE system used by the French customers (IIBRBS).

4.1.2 Presentation of the five selected models

This section describes the selected model structures, their hydrological process and their main differences in terms of simulated stream flow with a special attention given to the links between model structure/efficiency and model variables/basin characteristics.

- The IHACRES model (Identification of unit Hydrographs and Component flows from Rainfall, Evaporation and Streamflow data)

The IHACRES rainfall-streamflow modelling methodology is the result of collaboration between the Institute of Hydrology, Wallingford, UK and the Australian National University, Canberra (Jakeman *et al.* 1990; Littlewood and Jakeman, 1994).

The IHACRES methodologies have been packaged (PC-IHACRES software and User Guide; Littlewood *et al.*, 1997). In principle, the methodology can be applied at any data time step. There are published accounts of analyses which have used data intervals ranging from as little as 6-minutes, at a spatial resolution of less than 1 ha (0.01km^2), to monthly data at a scale of 10 000km^2. The methodology has been successfully applied to many catchments at a daily data time step. Applications include unit hydrograph analysis for assisting with assessment of the impacts of environmental change (land-use, climate, etc.) and quality assurance of long, strategically important, hydrometric records.

The only input data required are: an unbroken time series of rainfall (and streamflow for calibration); corresponding air temperature (as an indicator of seasonal changes in evaporative demand); catchment size (km^2). The outputs are: streamflow; catchment wetness index; unit hydrographs; hydrograph separation (in many cases) into dominant quick and slow flow components; and indicative uncertainties associated with the unit hydrograph parameters.

In its hydrograph separation mode the model comprises three conceptual storages: a non-linear catchment wetness storage that determines the effective rainfall; and two linear storages in parallel, which route the effective rainfall

to modelled streamflow. In principle, any configuration of such linear storages can be prescribed but research has demonstrated that, given adequate data quality, many catchments are well represented by two linear storages in parallel. Usually, more complex linear routing structures cannot be identified from the information available in the rainfall, streamflow and air temperature data.

In the PC-IHACRES version of the model, the parameter optimization methodology uses an instrumental variable technique to identify the unit hydrograph parameters. The parameters of the non-linear storage are selected by a semi-automatic search of the parameter space. In the version used here, the three-storage IHACRES model structure has been adopted but potential evapotranspiration has been used (instead of air temperature). A time delay parameter (parameter 5) has been added at the output, in order to correct the effects of the linear routing. This modified model structure is known as IHAC (to distinguish it from the full IHACRES methodology). The result is a 7-parameter IHAC model as opposed to the initial 6-parameter version. (Figure 3).

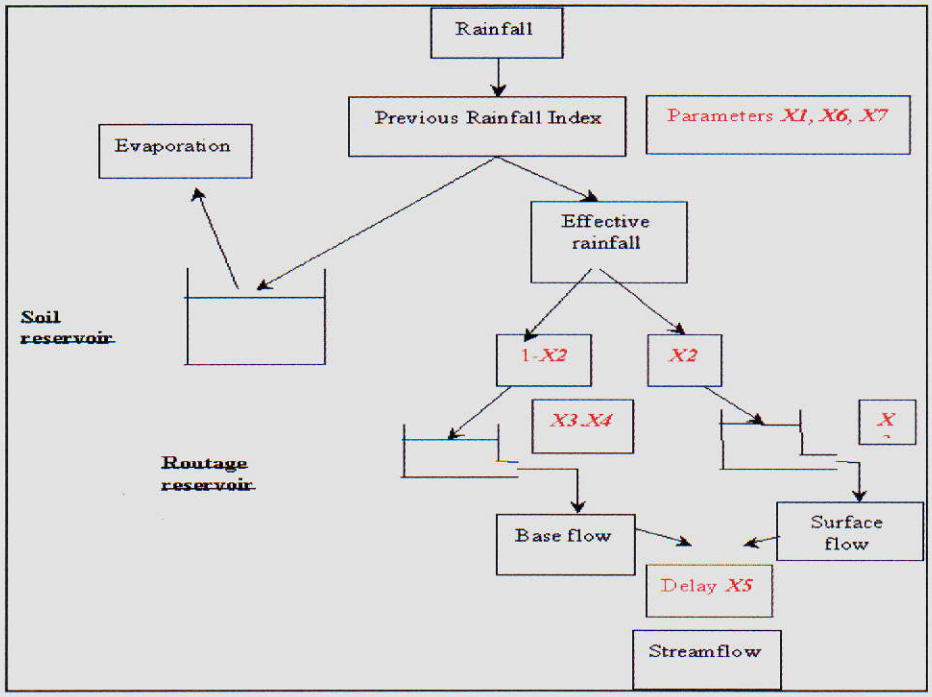


Figure 3 : Structure of IHAC model

- The TOPMODEL model (TOPography-based hydrological MODEL)

TOPMODEL was created in 1979 by K. Beven (University of Lancaster), and M.J. Kirkby (University of Leeds). Its main characteristics are a topographic index, variable contributing areas and the ability to use a daily time step. The input data are: rainfall (and stream flow for calibration), Penmann evaporation, the topographic index distribution over the basin, and the basin surface area. The outputs are: Stream flow.

The model is composed of 3 interconnected stores, two responsible for the non-linear production module: a soil moisture store and an interception store, and one store responsible for the quadratic routing of excess rainfall. Parameters can be determined empirically by soil measurements or with an automatic optimization procedure.
To estimate the spatial distribution of the topographic index, a Digital Terrain Model (DTM) is used. In the version selected here, this distribution is calculated with a logistic function, which depends on two optimized parameters. That allows using a global version of the model, without spatial discretization. We also suppose that runoff is only determined by flows on variable contributing areas, without Hortonian flow (Figure 4).

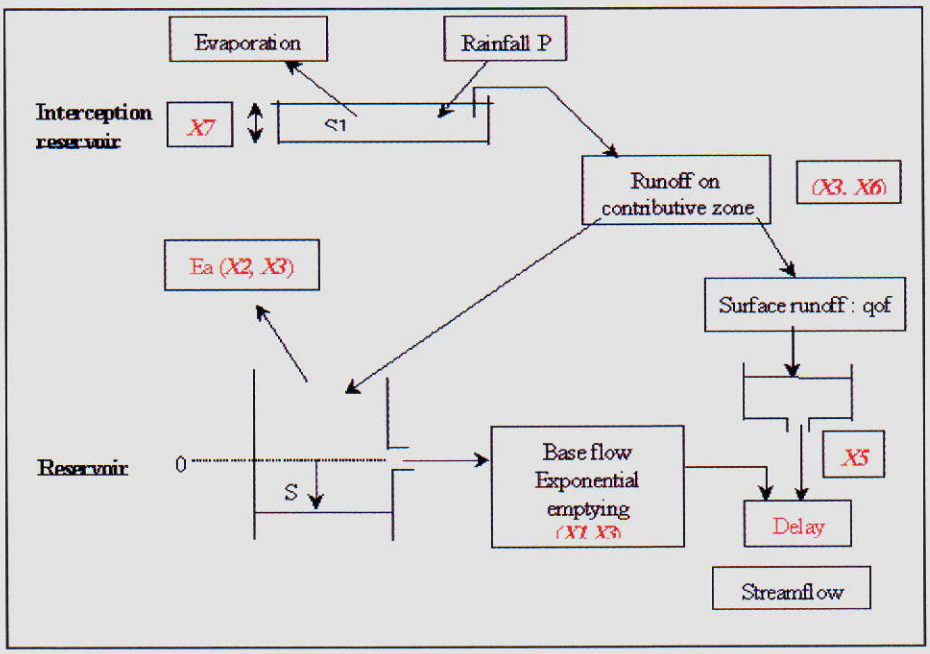


Figure 4: Structure of TOPMODEL model

- The HMS model (H. Morel-Seytoux model)

H.J Morel-Seytoux, has recently proposed the HMS model as an alternative to the GR4 model that can be integrated in the reservoir management model PEGASE, used by the French customers (IIBRBS). The model can be used at a daily time step.

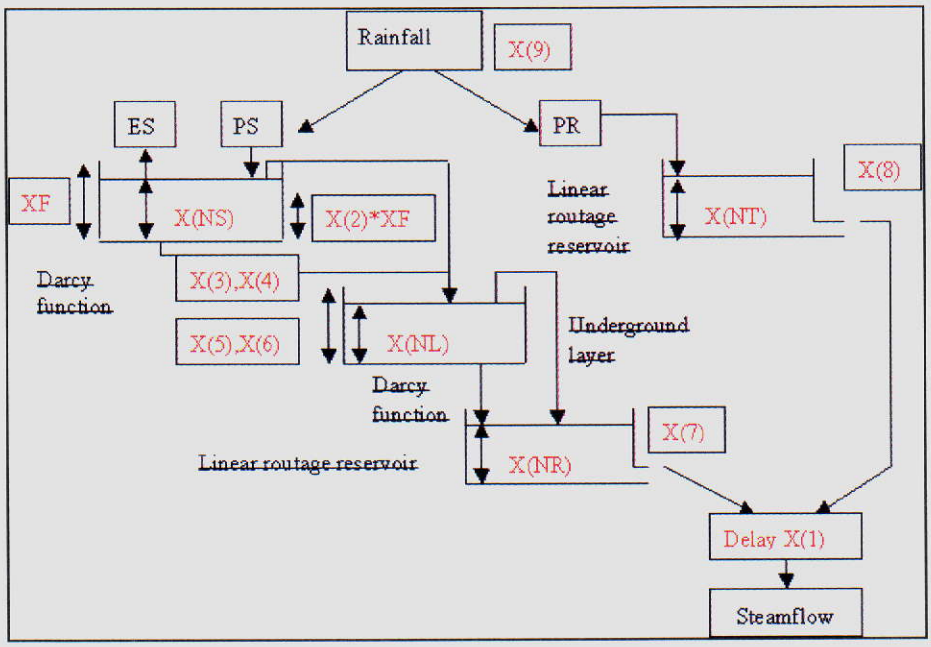


Figure 5: Structure of HMS model

The input data are: rainfall (and stream flow for calibration), evapotranspiration, and basin surface area.
The outputs are: stream flow.
The version tested here is composed of four-reservoir stores. Two of them are responsible for the linear routing process, and the other two are responsible for the production module in the upper soil layer and in the underground soil layer. The drainage process of these two layers depends on a Darcy function and can have overflows.
One of the model characteristics is that the base flow is produced by two reservoirs, one linear and the other non-linear. The version used in the project has 9 parameters (Figure 5).

- The GR4j model (modèle du Genie Rural journalier à 4 paramètres)

A first version of this model was created by Michel in 1983, at Cemagref, Antony, France. Edijatno *et al.* (1989), Nascimento (1995), and Edijatno *et al.* (1999) successively improved the original structure to achieve this current version with only four parameters.
Some interesting applications have been implemented by Makhlouf (1994), for model-parameter regionalisation and environmental changes, by Yang and Michel (2000) for flood forecasting, and by Baudez *et al.* (1999) for a semi-distributed approach. It can be used at a daily time step.
The input data are: Rainfall (and stream flow for calibration), evapotranspiration, and basin surface area.
The outputs are: stream flow
The parameter optimization is performed with a local optimization technique, the step by step method (Michel, 1989). The model is composed of 2 interconnected stores, one responsible for the non-linear production module with a soil moisture store, and the other responsible for the quadratic routing of excess rainfall coupled with two unit hydrographs, one for the surface flow and the other for the base flow component (Figure 6).

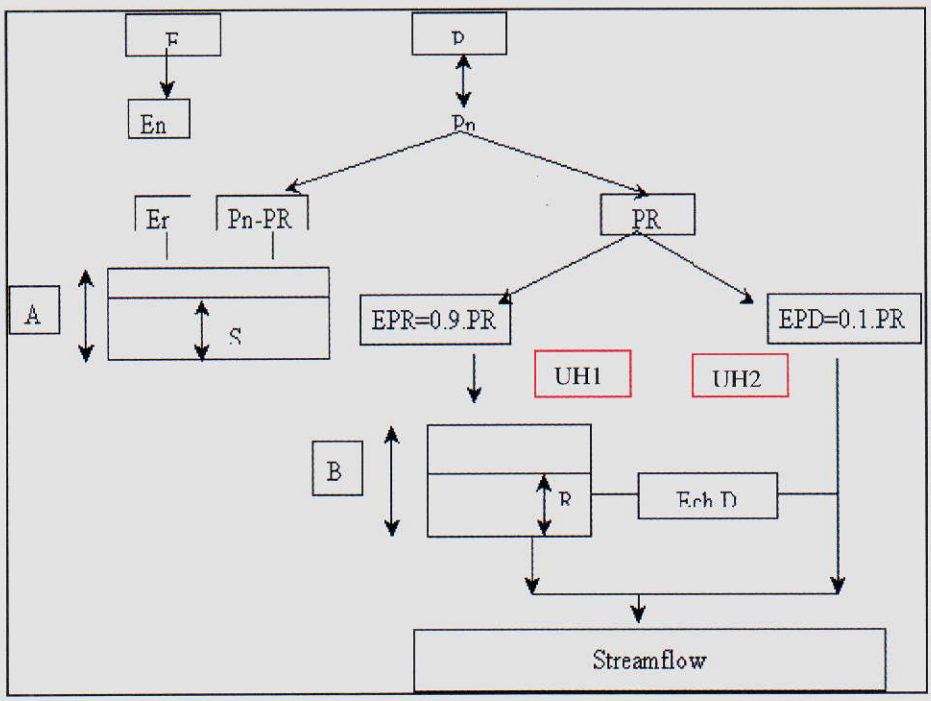


Figure 6: Structure of GR4j model

- The GRHUM model (modèle du Génie rural avec simulation de l'humidité)

This model was created by Loumagne *et al.* (1996), as a result of collaboration between two French organizations, Cemagref and CETP/CNRS.

Origin: CEH/Cemagref
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

The GRHUM structure is derived from the GR4j model, in which the soil reservoir has been replaced by a two-layer reservoir, to introduce exchanges at the interface of soil, land cover and atmosphere, and to simulate the surface soil moisture. It can be used with a daily time step.
The input data are: rainfall (and stream flow for calibration), evapotranspiration, and basin surface area.
The outputs are: stream flow and soil moisture
The parameter optimization is performed with the step by step method. The GRHUM model has the same structure of interconnected storages as the GR4j model. In the version used here the evolution of the vegetation index, which is difficult to estimate in an operational context, has been replaced by a variable that is linked to evapotranspiration. To avoid increasing the number of optimized parameters, the drainage process of the soil reservoir calculated by a Thomas function (Thomas, 1981) has been replaced by a simple maximum function (Figure 7).

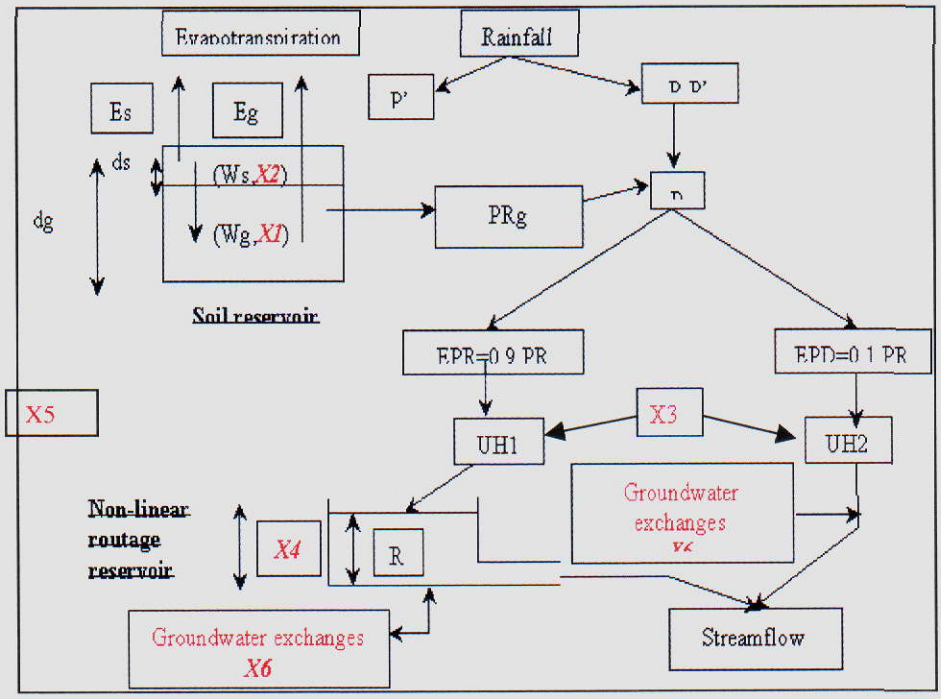


Figure 7: Structure of GRHUM model

- *Structural comparison of the five models*

The five selected models are spatially lumped or can be used in a lumped mode. IHAC, HMS and TOPMODEL are conceptual models, GR4J is considered as semi-empirical, and GRHUM is a mixed model, issued from an empirical structure, and coupled with a more physical soil-moisture scheme.
Their complexity, in terms of optimized parameters, is quite different: GR4J has a very simple structure, with only 4 parameters, GRHUM is a little more complex with 6 parameters, TOPMODEL and IHAC have an intermediary complexity with 7 parameters, and HMS has the most complex structure, with 9 parameters.
Although their structures are quite different from each other they all include a non-linear soil moisture accounting procedure and are subject to the evapotranspiration process.
The soil moisture procedure is performed:

- With a soil store reservoir for the TOPMODEL and GR4 model or a two-layer reservoir for the GRHUM, and the HMS model.
- Or, for the IHAC model, with a function which represents previous rainfall and moisture.

All the structures reproduce two flow components, the overland flow and the base flow, but their partition is not the same for the five models:

In the GR models, the partition is set up to 10% for the overland flow, and 90% for the base flow. In IHAC, where the flow components are labeled "quick and "slow", this partition depends on a parameter, which can vary between 0 and 100%. In TOPMODEL, the partition depends on the soil moisture state based on a topographic curve. In HMS, the partition is given by the SCS curve number method.
Another difference between these models is the existence of a groundwater exchange function in the GR models, which allows them to import or export water into or from the system; in the other models, the basin is assumed to be a closed system with evaporation and rainfall being the sole sources of loss and input.

4.1.3 Modelling process

- Efficiency criterion:

To assess model performances for the Arade and the Seine sub-catchments we have applied the well-known Nash-Sutcliffe criterion to stream flow (Nash and Sutcliffe, 1970). This criterion gives more importance to floods than to low flows, which seems more appropriate in the context of the AIMWATER project given its focus on flood forecasting..
This criterion expressed as a percentage shows the discrepancy between simulated runoff and observed runoff.

$$Nash(Q) = 100 . \left(1 - \frac{\sum_{i=1}^{n} \left(Q_{obs} - Q_{cal} \right)^2}{\sum_{i=1}^{n} \left(Q_{obs} - \overline{Q} \right)^2} \right),$$

Where Qcalc, and Qobs are the simulated stream flow and the observed stream flow respectively. $\overline{Q}$ is the observed runoff average during the calibration period, and n is the number of time step considered.
The parameter stability between each simulation is also a good criterion to evaluate the quality of the different calibrated models.

- Initialisation period and optimization methodology:

For the five models, the initialisation period is one year. After one year, we consider that the reservoir levels are adjusted, and have integrated the previous rainfall and soil moisture conditions.
All the parameters in each model structure (GR4j, IHAC, TOPMODEL, HMS, GRHUM) were calibrated using the step-by-step algorithm developed by Michel (1989).

- Choice of modelling periods:

In order to evaluate the efficiency of a model for a wider range of hydrological conditions over a catchment, it is necessary to perform simulation and cross-control simulation tests for different periods.. A selected time period is used for model calibration, but another time period is used for validation. An inversion is then made on calibration and validation periods.
This technique allows the detection of specific hydrological periods, which exhibit specific basin behaviour.
For each catchment, the models have been run first over the entire period of data availability and then, for more specific periods, during which the basin exhibits a particular behaviour. These specific periods are wet or dry periods. When possible, the same periods have been selected for the sub-catchments, to enable comparison of their behaviour.

4.1.4 Results

- Results over the Seine sub-catchments

The Saulx basin:
Data are available for this catchment from 1958 to 1999. In Figure 8 dry periods are represented in bright blue and correspond to the two-year periods 63/64, 71/72, and 75/76; wet periods are represented in dark blue and correspond to the two-year periods 65/66, 94/95.

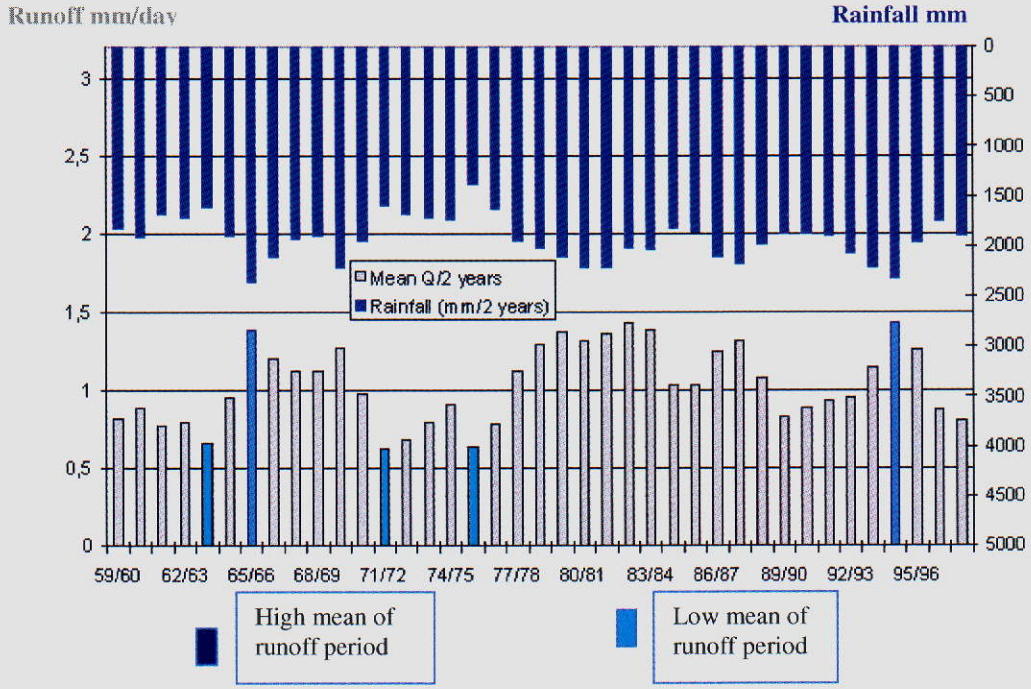


Figure 8: Rainfall and mean daily runoff for two-year periods, on the Saulx basin

Figures 9 and 10, show the mean performance obtained in calibration and validation mode for the five models during the specific dry and wet periods defined above and also for two larger global periods, 59/78 and 79/98 with average characteristics.

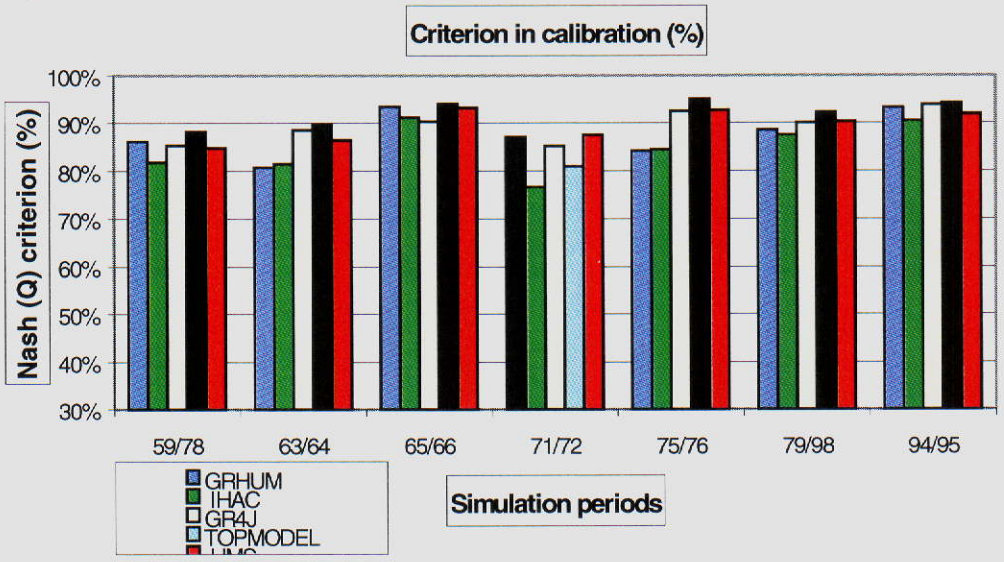


Figure 9: Nash criterion in calibration mode for the five models, on the Saulx basin

Cross-control simulation tests have been performed with similar periods in terms of hydrological characteristics: *global periods* (calibration on the 59/78 period, validation on the 79/98 period and inversely), *wet periods* (calibration on the 65/66 years, validation on the 94/95 years and inversely), *dry periods* (calibration on the 63/64 years, validation on the 71/72 and 75/76 years and inversely). In the case of dry periods the Nash criterion values correspond to the mean over the two validation periods.

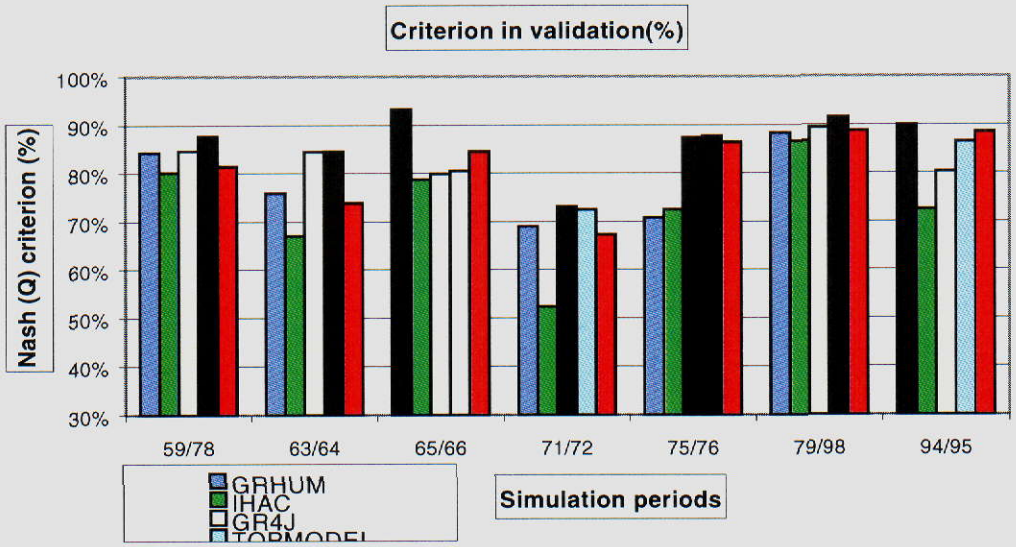


Figure 10 : Nash criterion in validation mode for the five models, on the Saulx basin

The models give a high and close level of accuracy in calibration mode. The level of performance for the Nash–Sutcliffe criterion is quite similar for all the models in calibration mode: about 80% to 90% except for dry periods such as 71/72 or 63/64. The best model for each specific period is represented by black dotted areas.
In validation mode, which is more relevant to the actual model's ability to simulate the rainfall-runoff relationship, performances range from 70% to 90% except for one model IHAC whose performances drop during the 71/72 period The TOPMODEL model has very stable ranking both in calibration and validation mode.

The Serein basin:

For this catchment, data are available from 1958 to 1999. In Figure 11 the dry periods are represented in bright blue and correspond to the two-year periods 63/64, 71/72, 89/90; wet periods are represented in dark blue and correspond to the two-year periods 78/79, 81/82.

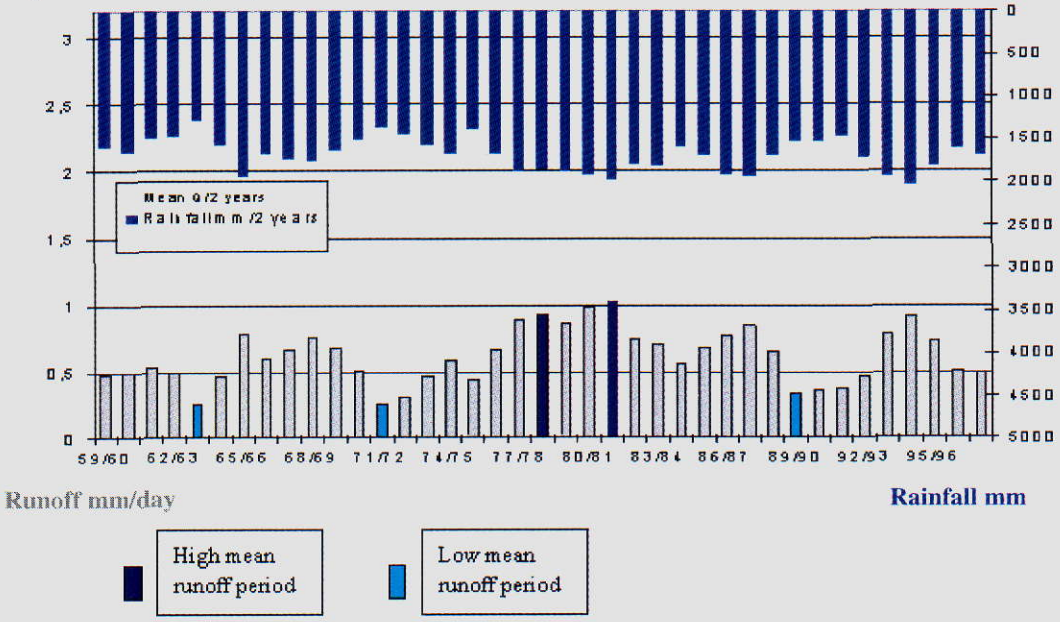


Figure 11: Rainfall and mean daily runoff for two-year periods, on the Serein basin

Figures 12 and 13 show the mean performances obtained in calibration and validation mode for the five models during specific dry and wet periods. Some of them are different from those defined for the Saulx basin (78/79, 81/82, 89/90) but the two larger global periods with average characteristics are the same: 59/78 and 79/98.

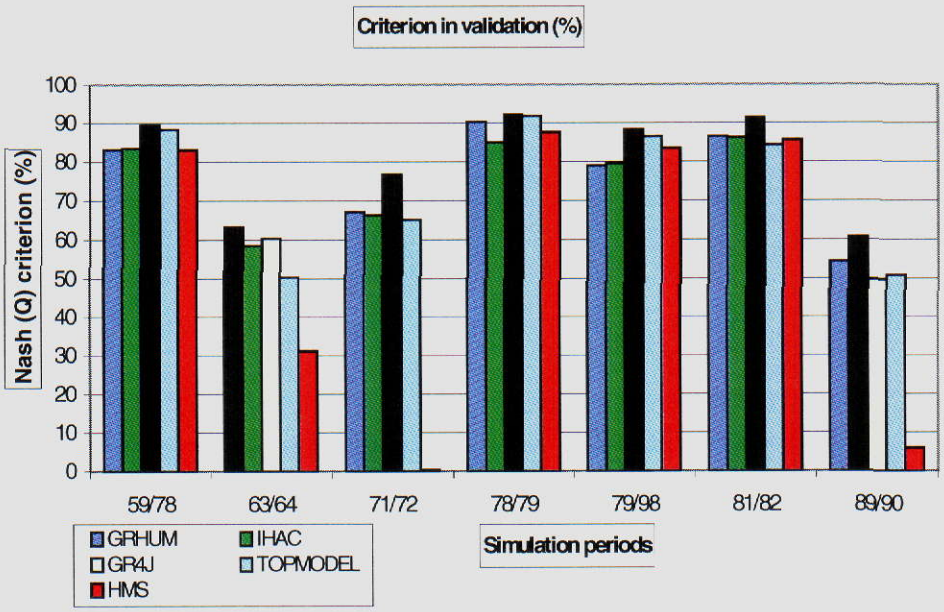


Figure 12: Nash criterion in calibration mode for the five models, on the Serein basin

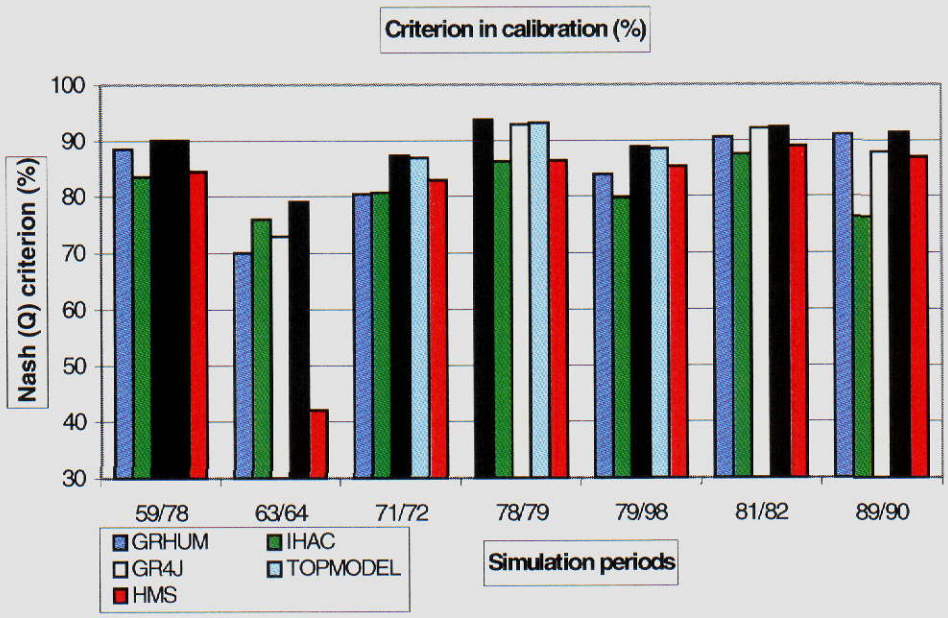


Figure 13: Nash criterion in validation mode for the five models, on the Serein basin

Cross-control simulation tests have been performed with similar hydrological periods:
global periods (calibration on the 59/78 period, validation on the 79/98 period and inversely),
wet periods (calibration on the 78/79 years, validation on the 81/82 years and inversely),

dry periods (calibration on the 63/64 years, validation on the 71/72 and 89/90 years and inversely for each validation period).
In the case of dry periods the Nash criterion values correspond to the mean over the two validation periods. Except for the dry periods (63/64, 71/72, 89/90), the level of performance for the Nash–Sutcliffe criterion is quite similar for all the models in calibration and validation mode: around 80 %. The best model for each specific period is represented with black dotted areas. In validation mode, all the models, and more particularly HMS, are inadequate in modelling the dry periods. Over these periods water balance problems are greater. This characteristic can be explained partly by particular hydrological phenomena over this basin: water losses could be linked to the chalk nature of the catchment substratum, which are not correctly integrated in the model structures.

The Grand Morin basin:

Stream flows in the Grand Morin are directly influenced by releases from the Marne reservoir, so this basin is evaluated using the results from the Petit Morin, a sub-catchment of the Grand Morin. For this catchment data are available from 1958 to 1999 with a gap in the runoff series from 1982 to 1985. During this period no simulation can be performed because the Nash (Q) criterion would not be adequate.
In Figure 14 the dry periods are represented in bright blue and correspond to the two-year periods 91/92 and 96/97; wet periods are represented in dark blue and correspond to the two-year periods 80/81 and 87/88.

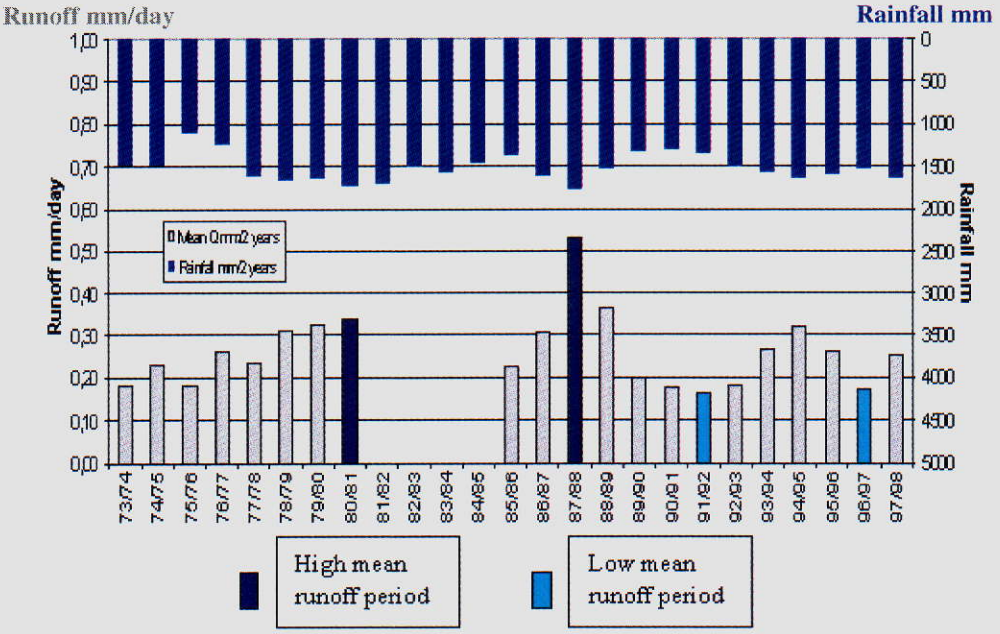


Figure 14: Rainfall and mean daily runoff for two-year periods, on the P. Morin basin

A further problem appears in the runoff data series: before 1978, runoff values never reached 0.9 m³/s, whereas after 1978 this flow was exceeded on many occasions. One explanation may be a change in the rating equation between the two periods.
Figures 15 and 16 show the mean performances obtained in calibration and in validation mode for the five models during selected dry and wet periods.

The cross-control simulation tests have been performed with similar hydrological periods:
global periods (calibration on the 86/91 period, validation on the 92/98 period and inversely), *wet* periods (calibration on the 80/81 years, validation on the 87/88 years and inversely), *dry* periods (calibration on the 91/92 years, validation on the 96/97 and inversely).
The best model for each specific period is represented by black dotted areas. In calibration mode, the level of performance for the Nash-Sutcliffe criterion is quite similar for the five models, around 80%, except for one dry period, 96/97, where the criterion drops to around 70%. However, we can say that the models give a high and

close level of accuracy in calibration mode. In validation mode, most of the models are inadequate in modelling the stream flow, the performance of Nash–Sutcliffe criterion ranges between 44% and 82%. This is more relevant for two models: GRHUM and HMS, during two specific periods 96/97 and 87/88 respectively.

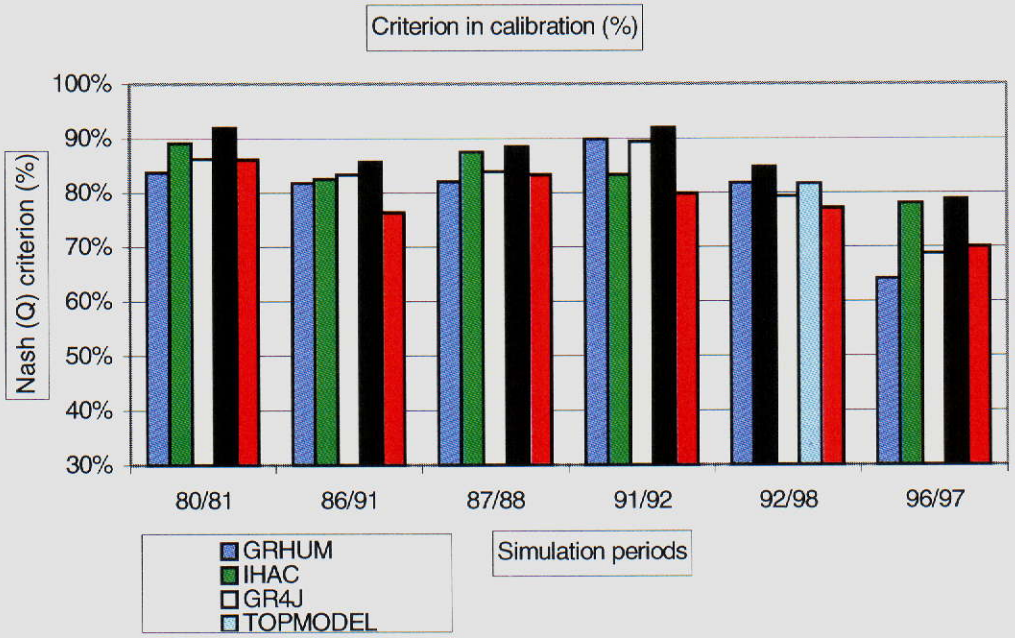


Figure 15: Nash criterion in calibration mode for the five models, on the P. Morin basin

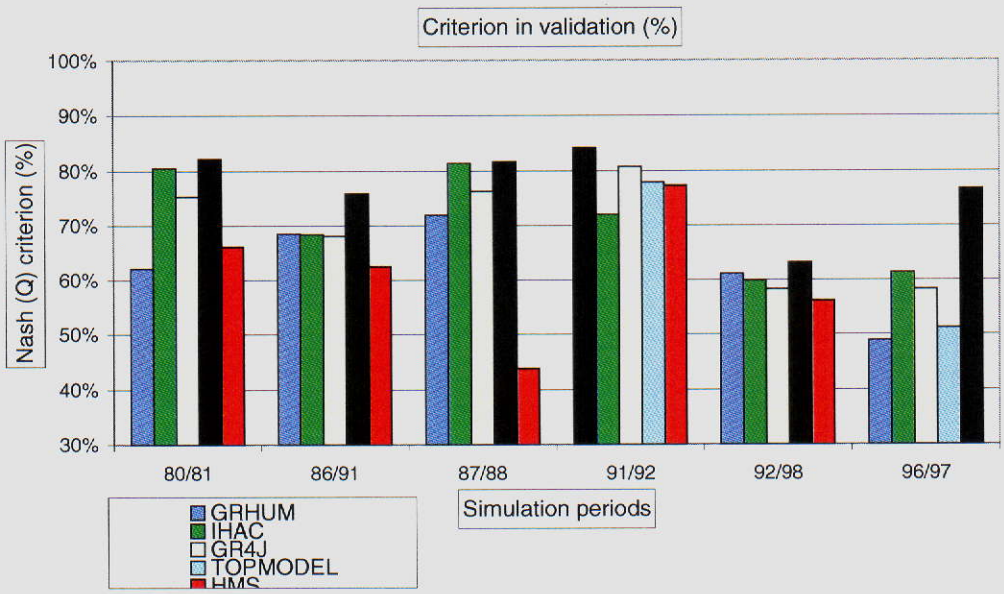


Figure 16: Nash criterion in validation mode for the five models, on the P. Morin basin

In this basin, all the models are inadequate in validation mode, so there is certainly a specific characteristic that implies a particular hydrological characteristic which could be partly explained by the hydromorphic soils observed over this basin coupled with a higher rate of artificial drainage that increases the surface runoff.

- *Results over the Arade basin*

The daily runoff data available at the watershed is from a stream-gauge located at the site of the dam prior to dam construction, called Casa Queimada, with a contributing area of 257.5 km^2. At Casa Queimada the daily runoff is available from November 1933 to September 1951 and the daily rainfall was measured at the same location with only one rain gauge for the same period for which streamflow is available. A selection of dry and wet periods was performed in order to study the behaviour of the Arade sub-catchment during these specific periods. In Figure 17 the dry periods are represented in bright blue and correspond to the two years periods 34/35 and 44/45; wet periods are represented in dark blue and correspond to the two-year periods 36/37 and 41/42, see Figure 17.

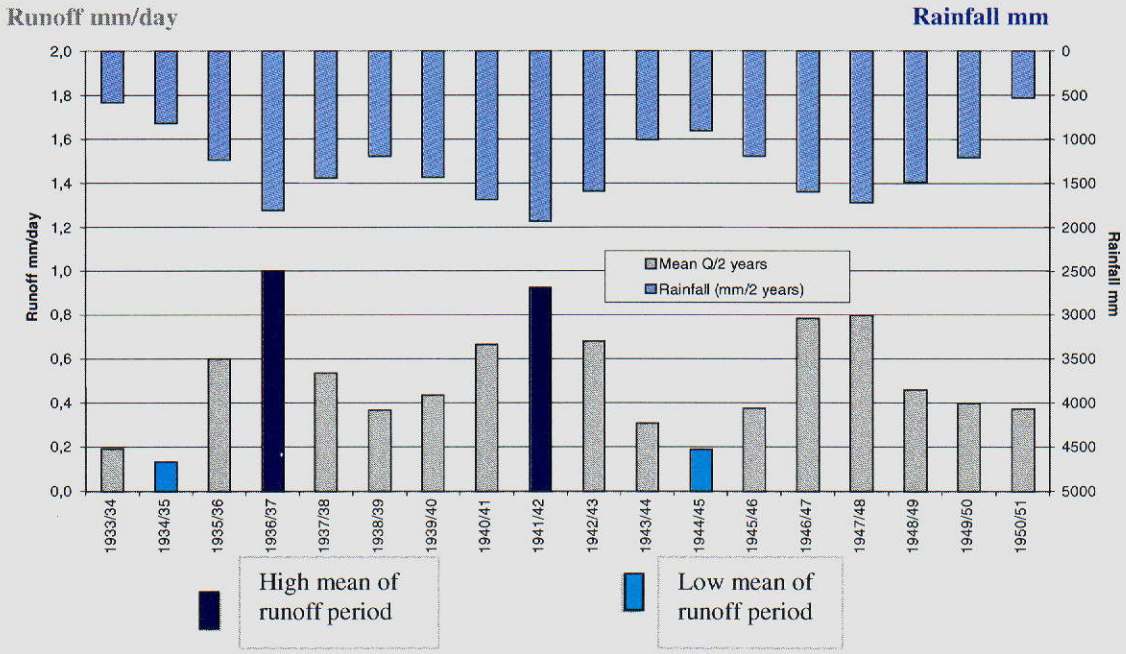


Figure 17: Rainfall and mean daily runoff for two years periods, on the Arade basin

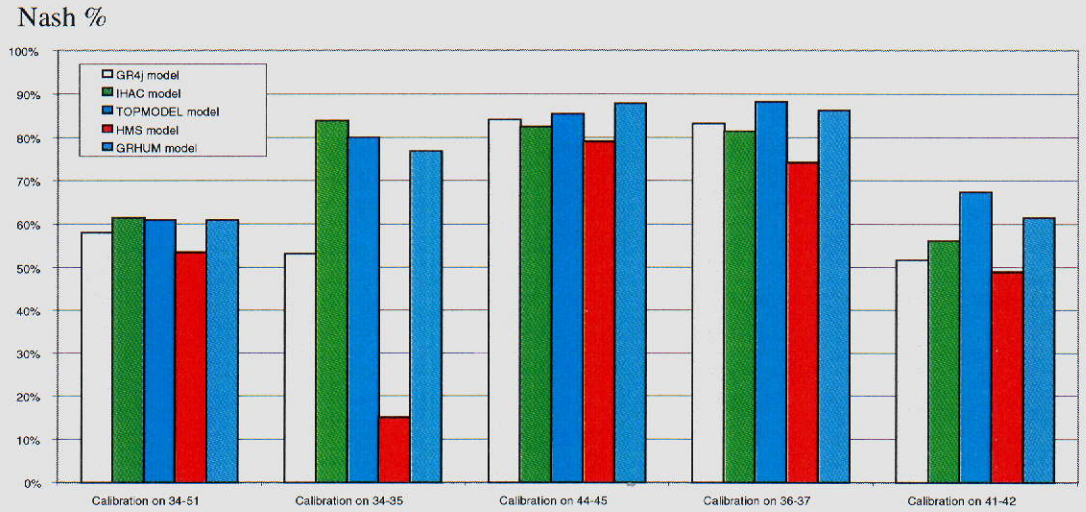


Figure 18: Nash criterion in calibration mode for the five models, on the Arade basin

The low values of runoff averages are due to the ephemeral nature of flow in the Arade river which is frequently dry for almost half the year..
Figures 18 and 19 show the mean performances obtained during calibration and validation of the five models for the specific dry and wet periods defined above, along with calibration results for one larger global period, 34/51 and validation for two smaller global periods 34/44 and 41/51having average characteristics.

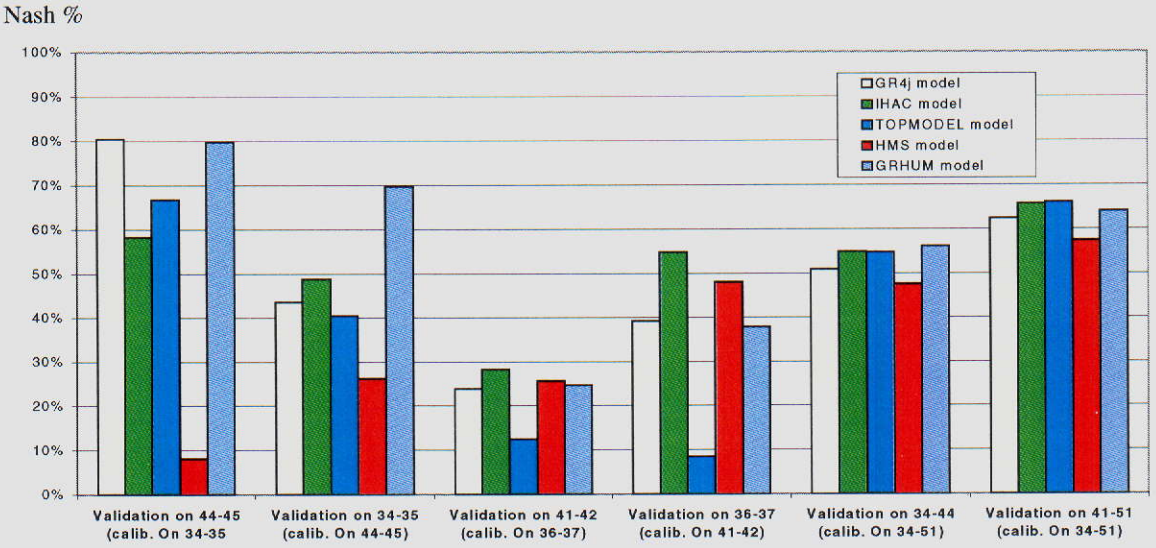


Figure 19: Nash criterion in validation mode for the five models, on the Arade basin

Cross-control simulation tests for periods having similar hydrological characteristics have been performed. These are for:
wet periods (calibration on the 36/37 years, validation on the 41/42 years and inversely) and *dry* periods (calibration on the 34 /35 years, validation on the 44/45 years and inversely). *For the global period*, calibration was performed on the 34/51, 20-years period, and validation on two ten-year periods 34/44 and 41/51.

In calibration mode for global, wet and dry periods, almost all the models give quite a high level of accuracy with a Nash criterion between 60% and 85%, the exception being the HMS model (Nash=15%) for the 34/35 dry period. However this is not the case for the period of 41-42 where all model performances drop to 50%-60%.
In validation mode, performances range from 50% to 80% except for the 41/42 years period, where all the models performances drop to about 20%. This particular time belongs to a disrupted "war period" where data collection problems were encountered. Because of this problem, the average Nash criterion performance in calibration and validation mode was calculated without the 41/42 periods.

- Conclusion on model efficiency

The Seine basin
In total five models (GR4J, IHAC, TOPMODEL, HMS, GRHUM) were tested in the Serein, Saulx, Petit Morin, and Grand Morin subcatchments of the Seine basin.

Figures 20 and 21 show how the five models performed, on average, in calibration and in validation modes respectively.
Figure 20 shows the 20 Nash (Q) calibration averages for each model and Figure 21 shows the 20 Nash (Q) validation averages for each model.

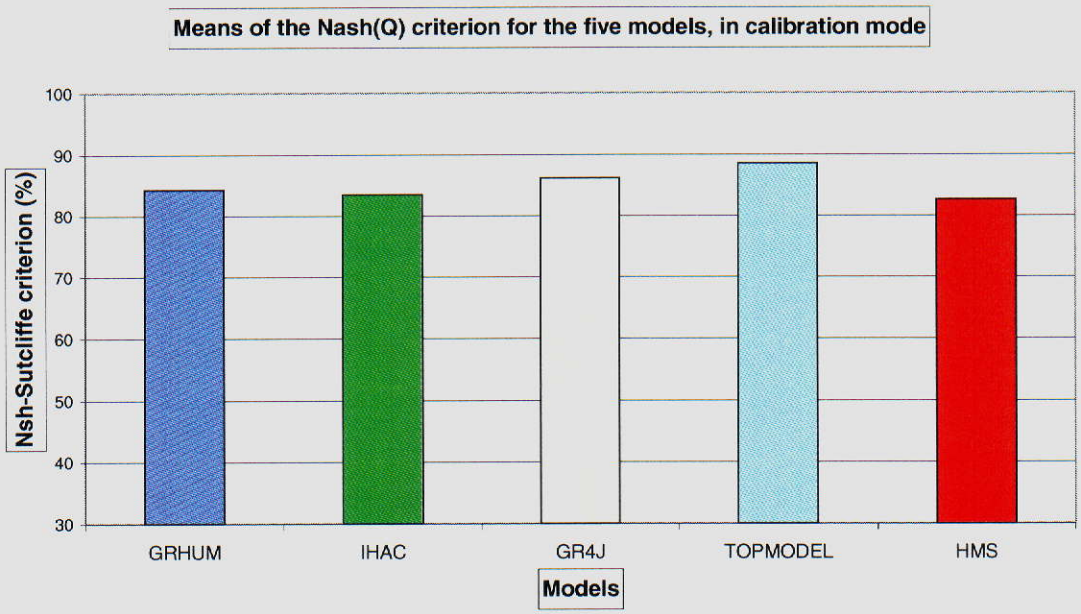


Figure 20: Averages of the efficiency criterion for the five models, in calibration mode on the Seine basin.

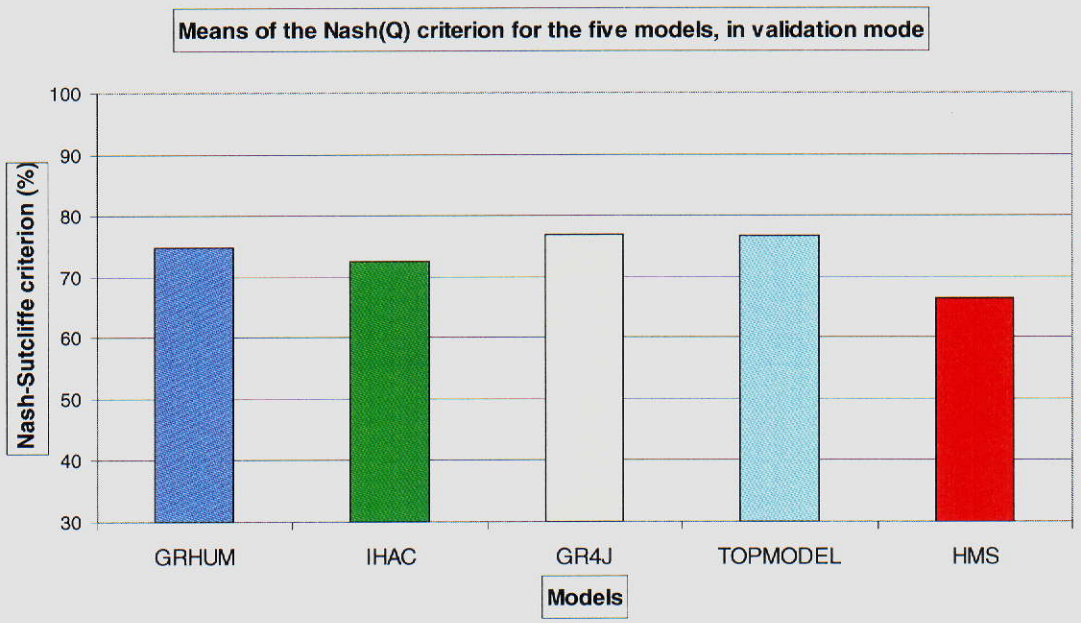


Figure 21: Averages of the efficiency criterion for the five models, in validation mode on the Seine basin.

In calibration and validation modes, the performance for the mean Nash-Sutcliffe criterion is high and close for the five models, although there are structural differences. HMS seems to be less efficient in validation mode, due to its larger number of parameters. Its calibration is too specific over a period, and cannot be applied to other periods.

The Arade basin

In the figures below, the mean performance of the Nash-Sutcliffe criteria for the Arade sub-catchment is presented for each model, in calibration and validation mode. Figure 22 shows the Nash (Q) calibration averages and figure 23 shows the Nash (Q) validation averages for each model.

In calibration and validation modes, the level of performances for the mean Nash-Sutcliffe criterion is close for four of the five models. In calibration the Nash-Sutcliffe criterion is about 70 to 80 % and in validation it is about 60 to 70 %. HMS seems to be less able to simulate non-permanent flow under arid climates. In the case of the Arade sub-catchment the GRHUM model appears to be the most efficient, both in calibration and in validation mode, with a stable ranking.

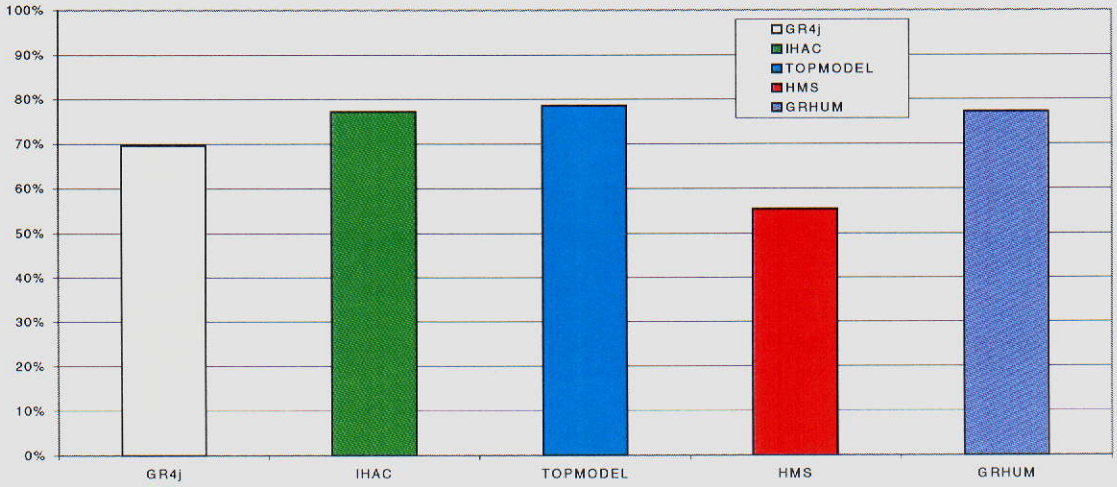


Figure 22: Averages of the efficiency criterion for the five models, in calibration mode on the Arade basin

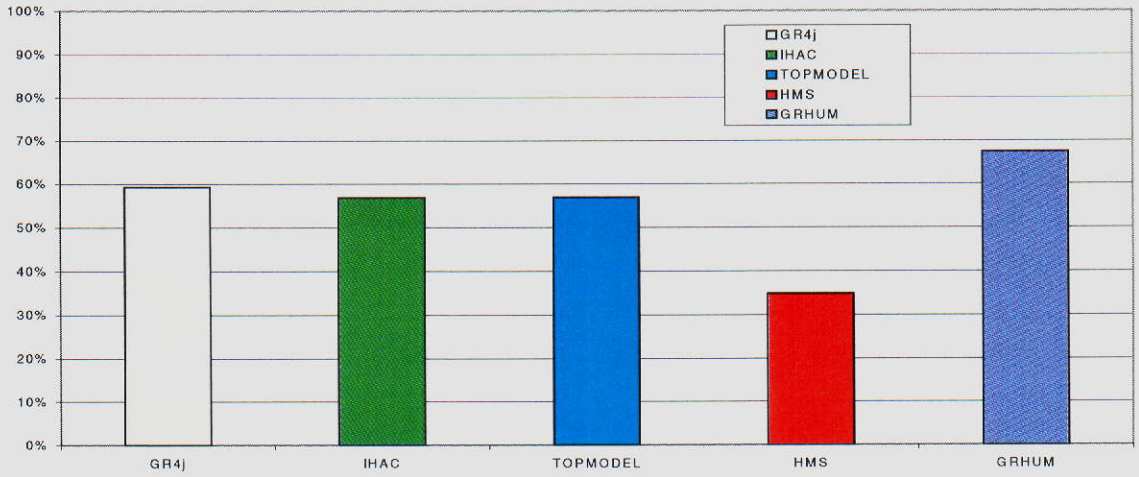


Figure 23: Averages of the efficiency criterion for the five models, in validation mode on the Arade basin

Analysis of these figures and numerous other test results suggested convincingly that the GR and its derivatives such as GRHUM and IHAC models were the best candidates for use in the tests to be conducted including the assimilation of information on soil moisture derived from the EO radar signals. All these models had provided high and close performances in terms of the Nash-Sutcliffe criterion.

Some of the results presented above can be linked to specific geographical and morphological characteristics of the basins, which is the subject of the following section.

4.1.5 Results analysis, regarding basin characteristics

The analysis carried out on the Seine sub-catchment results showed that some differences in the modelling results could be explained by basin characteristics. The main differences that illustrate the link between catchment characteristics and the parameters values of conceptual models, were focussed on the delay parameter of the basin response to a rainfall event and on the exchange parameter or models behaviours dealing with water excess.

- Delay parameter and basin lag-time

The Saulx has the longest lag-time, with a four-day delay, then the Serein with a three-day delay, the Petit Morin and the Arade with only a one-day time-to-peak. These differences appear in the models' delay parameter. In each structure a parameter is optimized in order to account for the lag time between rainfall and stream flow. In the case of TOPMODEL, IHAC and HMS, this parameter is a pure time delay whereas for the others the parameter corresponds to the time base of a unit hydrograph function.
Figure 24 shows the mean time delay parameter obtained by all the models in calibration mode for the three Seine sub-catchments and the Arade sub-catchment.

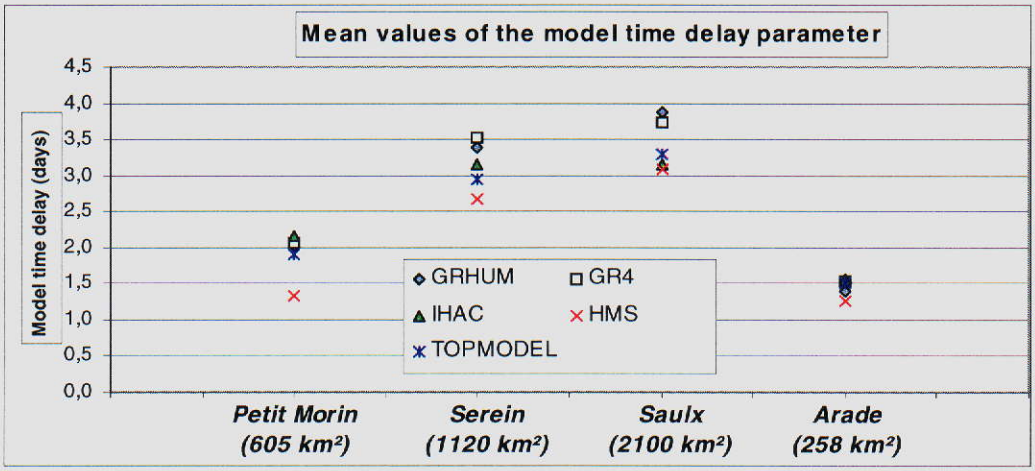


Figure 24: Values of model time delay parameters on the Seine and Arade catchments

Based on these results, it is clear that the Saulx catchment reacts less quickly to a rainfall event than the other catchments. The lag-times, for the selected basins, are approximately 3.5 days for the Saulx, 3 days for the Serein, 2 days for the Petit Morin and 1.5 days for the Arade. These different behaviours can be explained partly by the shape and the size of the catchments. The Saulx (2100 km²) is the largest catchment and the Arade (257.5 km²) is the smallest.
The elongated shape of the Serein basin increases the delay of the basin response and might explain the small difference between the lag-times of the Saulx and the Serein. But the pedology and the land cover can also be responsible for the time-to-peak differences. Indeed, the hydromorphic soils on the Petit Morin basin (which generate surface runoff), can increase the shorter lag-time in this catchment in the same way that the high forest percentage in the Saulx basin can increase the slowest response in the catchment.

- Exchange parameter

Another main difference appears in the parameters of the transfer function, explained partly by the exchange parameter, showing how the models cope with non-permanent flows. For instance, the two models, GRHUM and GR4J have an exchange parameter, which allows them to simulate ground water export or import. For these models, the values of the exchange parameter are really different for both dry and wet periods as presented in Figure 25 (Seine) and Figure 26 (Arade).

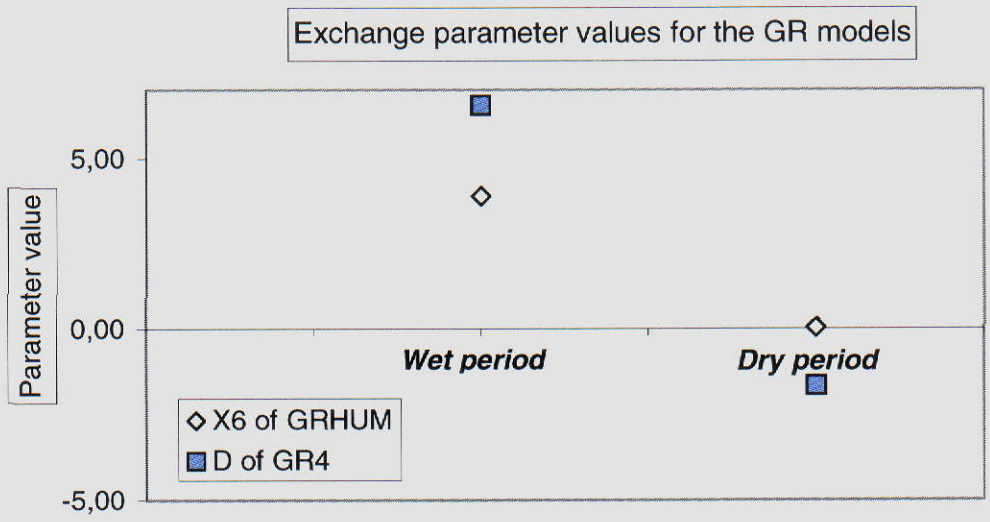


Figure 25: Exchange parameters for the GR models on the Seine.

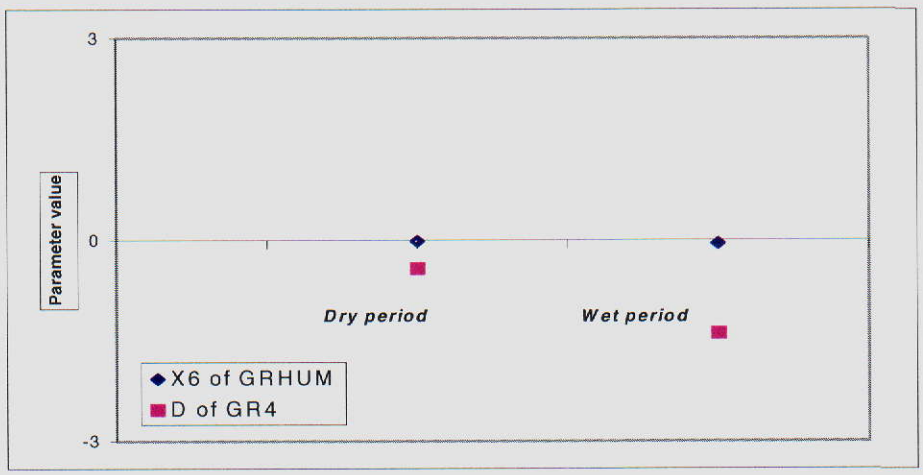


Figure 26: Exchange parameter values for GR models on the Arade

In the Seine sub-catchments, during wet periods the exchange parameter is higher than dry periods. The other models, which cannot simulate water losses except through evapotranspiration, use another way to release water during the wet periods.
Part of the effective rainfall is stored in one of their routing reservoirs, which then empties very slowly thanks to a high value of the outflow parameter. Therefore most parts of the flow are routed through a single store.
It is believed that the catchment water losses could be linked with the karstic phenomena observed in some parts of the Saulx basin.

Over the Arade, for the GRHUM model, the exchange parameter is almost zero, whereas for the GR4 model, the exchange parameter is between –0.5 and –1.5.
In contrast to the Seine sub-cathments it appears that during wet periods the exchange parameter is lower than during dry periods and has negative values, which means there is no loss, but some input of water in the basin.

These results show the difficulty of the GR4 model to deal with the problem of simulating stream flow for an intermittent basin.
Comparing the modelling results and the basin geological and geomorphologic features, allows a better understanding of the selected catchments' behaviour.
Both the Saulx and Serein, have similar pedologies, which can be linked to their similar soil reservoir capacities. The main differences appear in the parameters of the transfer function, explained partly by more higher water losses during wet periods and a longer lag-time for the Saulx basin. The third basin, the Grand Morin, evaluated through the Petit Morin results, a Grand Morin subcatchment, has a shorter time-to-peak, because of its soil characteristic (hydromorphic soils), and its smaller size.

4.1.6 Conclusions

The main conclusions to be drawn from the assessment of model performances and basin characteristics over the studied sub-catchments are:

- A first analysis had shown some differences in geomorphological basin characteristics: karstic phenomena have been identified in the Saulx and Serein catchments as well as a high percentage of drained agricultural land; in the Grand Morin hydromorphic soils are an important feature. Differences in rainfall rates have also been noted, the Saulx basin receiving more rainfall than the other two.

- The modelling of the Seine basin, supported by the five selected models (GR4, GRHUM, HMS, IHAC, TOPMODEL), has demonstrated that in calibration and validation modes all models have provided high and close performances in terms of the Nash-Sutcliffe criterion. In validation, however, some difficulties in modelling the stream flow appeared in the Saulx and the P. Morin catchments, which might be linked, to particular hydrological characteristics such as a karstic nature, a high rate of artificial drainage and the presence of hydromorphic soils. The well-known TOPMODEL has been used as a reference in simulation mode for the assessment of the most suitable models.

- In the Arade basin, the models have good calibration results, but some problems appear in the validation mode especially with the HMS model. For this basin, the GRHUM model seems to be the most appropriate, having a stable ranking in both calibration and validation modes. However, all models experienced some difficulty in correctly simulating the flows of an arid Mediterranean ephemeral river.

- In spite of the lumped conceptual nature of the models, clear links can be made between catchment characteristics and the values of parameters optimized during the calibration procedure. It was shown that models manage to simulate correctly through their delay parameter the lag-time between stream flow and rainfall. It was also shown how the models cope with the excess amount of water by increasing the value of their exchange parameter or the value of their outflow parameter of one of their routing reservoir.

- One of the main issues of the project was to find out which model is the most effective for soil moisture assimilation. As regards to previous results and conclusions, GR4, GRHUM and IHAC will be coupled with an assimilation procedure and their efficiency will be discussed in the chapter below. Because of the less effective results of HMS, this model will not be considered for further assimilation purposes.

4.2 Comparison of assimilation methodologies

The use of remote sensing to predict stream flow is neither a common practice nor widely used. It is very much at the research stage, however it won't be long before it is used widely by the water resources authorities. It is well established that runoff generation is strongly associated with soil moisture status. If the latter can be estimated from satellite back scattering images then one could calculate the runoff volume and the stream flow. The AIMWATER project was formulated with this objective in mind and with idea of helping water reservoirs mangers to better manage their water demands among different sectors of society. Following selection of the appropriate models, the next step is to select the assimilation methodology by which the remote sensing data could be incorporated into the rainfall-runoff models to predict stream flow. The difficulty in predicting floods in a reliable way stems from the lack of accuracy of the models, particularly during unusual hydrologic events. Methods have thus been developed to improve flood prediction in hydrology based on the ability of model to simulate the current observations prior to use it in forecasting mode. In hydrology this operation has been termed updating; in meteorology it is known as assimilation.
The fundamental idea is that if the model predictions diverge from the observations at a given time, there is little chance that future estimations will approach the correct values. The improvement then comes from a correction of the trajectory of the model based on observations during the period preceding the day when a prediction into the immediate or long-term future is desired. The use of observed data to correct the values calculated by the model has been widely described in the literature (O'Connell and Clarke, 1981, Refsgaard, 1997) and refers to four different methodologies used for model updating. However, the use of remote sensing data, especially the soil moisture of the top 10 cm of the soil surface, in the context of forecasting stream flow has not been widely reported and therefore it represents a new challenge for AIMWATER researchers. Ultimately the study focused on three methods, which proved to be more accurate, and could be accommodated within the rainfall-runoff non-linear models used in this project. This work package has dealt with important tasks started after selecting the appropriate rainfall-runoff models followed by selection of the assimilation methods, followed by calibrating, and finally by simulating and forecasting the stream flow. In every activity the user's requirements have been borne in mind and represents the final objective of the work. The main results were published during the final Workshop held in Montpellier (Oudin *et al.*, 2001, Aubert *et al.* 2001, Ragab *et al.* 2001) and the exe files of these methodologies can be found on the web and FTP site dedicated to the project. In the following sections, the assimilation methodologies, model adaptation to incorporate soil moisture, model calibration and model simulation and forecasting will be discussed in detail.

4.2.1 Internal variables updates for R-R models

Shortcomings in the updating methodologies used in rainfall-runoff models are considered to be the main cause of discrepancy between observed and computed stream flow values. These shortcomings can be corrected or improved in a number of ways:

- Errors in rainfall data can be corrected by using effective rather than the gross rainfall; corrections of this type are termed "input updating".
- Inherent deficiencies in the model can be addressed in two ways:
  - After the current state of the model ("state updating")
  - Adjust the model to current conditions by changing its parameter values ("parameter updating").
- Actual stream flow can be taken as the sum of the model output and an error term, which has to be modelled to allow for a prediction about its short-term realizations (error correction). This updating is especially used with black box models, which are not considered here.

A brief description of the three methodologies used in the context of the project is provided thereafter:

- Input updating in forcing mode: the scaling method

The method is mainly based on using the effective rainfall instead of the gross rainfall. The effective rainfall is that part of the gross rainfall that effectively generates the runoff and as such is used as input to rainfall-runoff models. It is envisaged that the model's ability to better predict stream flow will be enhanced by using the

effective rainfall instead of the gross rainfall. The effective rainfall for a given day can simply be obtained by multiplying the gross rainfall of that day by the wetness index of the catchment on the same day. The wetness index is an indicator of the soil moisture status of the catchment, being 1 for a saturated catchment and 0 for a dry catchment. If the wetness index is 1 the catchment is saturated and all rainfall will effectively contribute to the runoff since the gross rainfall will be equal to the effective rainfall. If the catchment is dry e.g. 0.1 then only 10% of the gross rainfall will be effective in generating runoff while 90% will go to satisfy the soil moisture deficit and bring the catchment back to saturation before runoff can be generated. The wetness index is basically a scaled soil moisture value for the entire catchment. Soil moisture values are obtained from remote sensing, TDR or Neutron probe measurements and scaled to a single value ranging from 0 and 1 to represent the catchment wetness status for a given day. Figure 1 shows a schematic diagram of the principles of the approach.

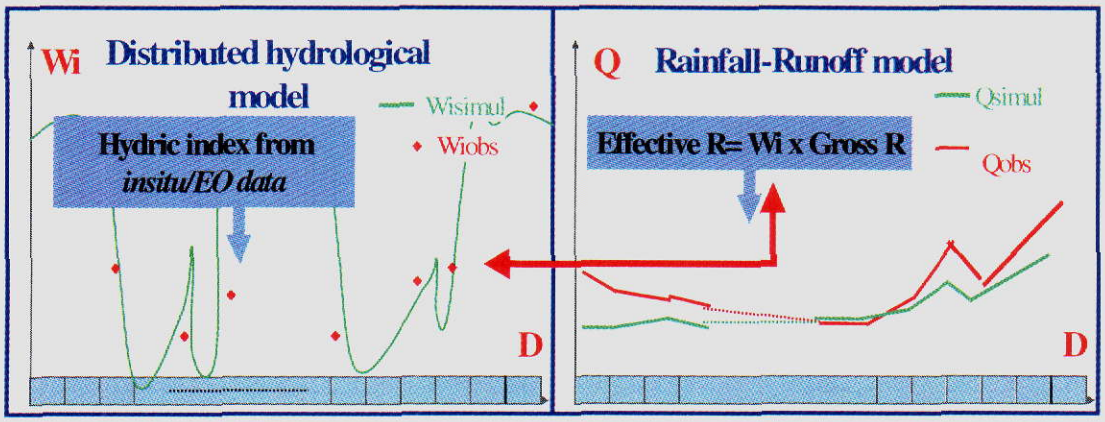


Figure 1: The scaling method

- Parameter updating: a variational method

The inability of a model to produce more accurate stream flow values generally translates into parameter uncertainty. Parameter calibration is the means used by a model structure to adjust to a given set of data so that parameter updating seems to be a natural way to amend the inaccuracy in stream flow value (Figure 2).

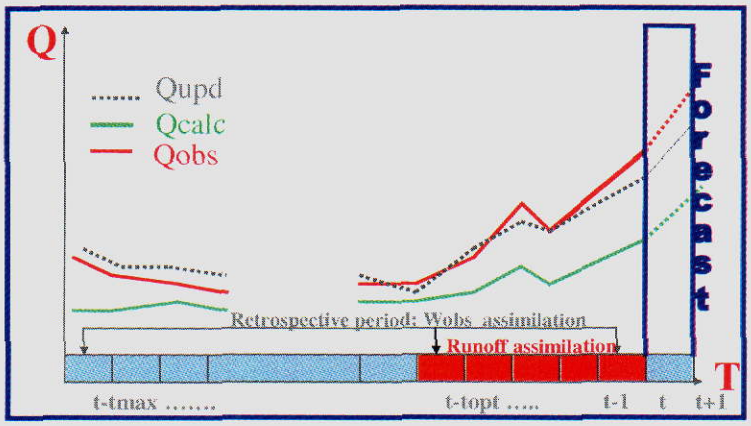


Figure 2: The variation method

The methodology set up by Yang and Michel (2000) has been described thoroughly in (Weise *et al.*, 2001). The most important features of the method are:

- Prior analysis of errors impairing flood discharges is carried out through a regression relationship:

$$Q_j = \alpha + \beta C_j + \varepsilon = C_{A,j} + \varepsilon \qquad (1)$$

where Q is the observed flow rate, C is the flow rate calculated by the model, j refers to any day when C_j is above the threshold defining a flood, and ε is the error term of the regression; α and β are additional parameters that take into account the potential lack of accuracy of the model at high discharge rates.

- The updating process retraces backward a sufficiently long time for parameter updating to produce a substantial effect at the present time. A 'baseline' set of parameters, obtained by calibration over a long period, is assumed to represent the long-term behaviour of the basin.
- Parameter updating is kept close to the baseline set in order to reduce the last error in stream flow value. Actually, in the absence of additional information, the best thing to do is to alter all parameters around the baseline set, in the smallest way with respect to their uncertainties.
- A hedging measure is imposed in order to force parameters not to depart too much from the baseline set.

This methodology was tested in Yang and Michel (2000) against data from four watersheds upstream from Paris. Results revealed the parameter updating method to perform better than the traditional method of correcting errors using an AR(1) model (Auto-Regressive model of order 1) (Lettenmaier and Wood, 1992).

- *State updating: a sequential method*

This methodology is meant to locally correct the value of the internal variables of the model when an observation is available (Figure 3). At time t, when an observation is available, which does not exactly correspond to the simulated *a priori* value, a correction is made and an *a posteriori* value of the variable, closer to the actual observed value, is obtained. Then the model starts from these new initial conditions and evolves freely until new observations are available.

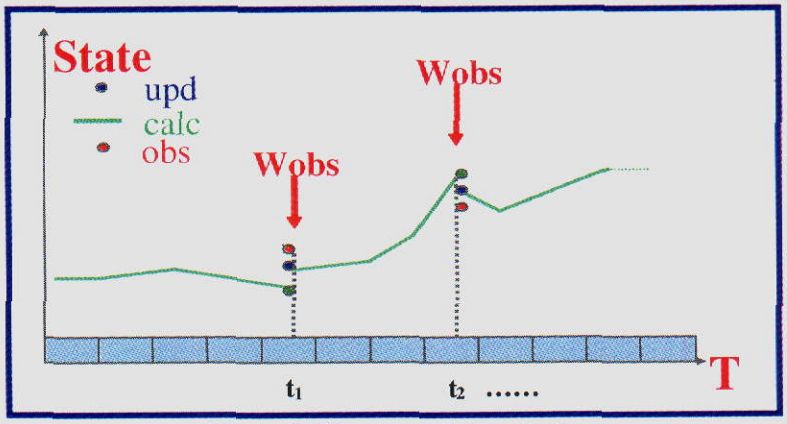


Figure 3: The sequential method

The calculation of the correction factor is an important step in the method. To calculate this correction factor an extended Kalman filter is commonly applied (Kalman, 1960). The method takes account of the errors estimated by the model and the observations, and locally linearizes the model to determine the correction needed in the internal variables. Thus the more the observations are precise, the more the *a posteriori* state of the model will be close to them. A description of the extend Kalman filter is given in (Quesney *et al.*, 2001 and Aubert *et al.*, 2001)

The application of the Kalman filter requires two steps, the adjustment step and the propagation step (Table 1). During the adjustment step, the gain matrix $\boldsymbol{K}_k$ is calculated. It indicates, for each internal variable, the correction factor, that needs to be applied. This calculation takes into account the instantaneous error matrices $\boldsymbol{R}_k$ and $\boldsymbol{P}_k$. This adjustment is then applied to the state vector: the *a posteriori* state vector $\hat{\mathbf{x}}_k$ is equal to the *a priori* state vector $\mathbf{x}_k$ plus the difference between the observation and state vectors multiplied by the Kalman gain. The error covariance matrix is also updated: the *a posteriori* matrix $\hat{\mathbf{P}}_k$ is equal to the *a priori* matrix $\hat{\mathbf{P}}_k$ plus the gain matrix multiplied by the constraint matrix.

During the propagation step, the values at time $k+1$ are calculated: the new state vector is calculated by using the model operator f, and the new error covariance matrix $\mathbf{P}_{k+1}$ is calculated by adding the error at time k $\mathbf{P}_k$ propagated via the tangent linear and the error $\mathbf{Q}_k$ generated during this time step.
The importance of the correction is dominated by the ratio of the errors on the observations and the model. If the uncertainties on the observations are assumed to be small (i.e. $\left|{}^{\mathbf{R}_k}/_{\mathbf{Q}_k}\right| \approx \mathbf{0}$), the Kalman gain will be almost equal to 1 and the internal variables will be completely corrected (forced mode). On the other hand if the observations are unreliable, the Kalman gain will be almost 0 and there will be no corrections.

Adjustment	Propagation
• Kalman gain $\mathbf{K}_k = \mathbf{P}_k \mathbf{H}_k^t \left(\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^t + \mathbf{R}_k\right)^{-1}$ • State updating $\hat{\mathbf{x}}_k = \mathbf{x}_k + \mathbf{K}_k \left(\mathbf{z}_k - h(\mathbf{x}_k)\right)$ • Errors updating $\hat{\mathbf{P}}_k = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right)\mathbf{P}_k$	• State propagation $\mathbf{x}_{k+1} = f(\hat{\mathbf{x}}_k)$ • Errors propagation $\mathbf{P}_{k+1} = \mathbf{M}_k \hat{\mathbf{P}}_k \mathbf{M}_k^t + \mathbf{Q}_k$

Table 1: The extend Kalman Filter

4.2.2 Integrating Soil Moisture: New Information for Updating

To estimate the runoff volume following a rainfall event at the catchment scale, a number of models with different degrees of complexities were developed. The interactions occurring at the soil-vegetation-atmosphere interface, and in particular the variability of the soil moisture status, are dominating factors, which are not directly taken into account in classical hydrological models. The benefit of using soil moisture status as a prognostic variable in the hydrological cycle has been studied by several research teams (Engman, 1990, Loumagne *et al.*, 1991). They have demonstrated how introducing this information in hydrological models can improve stream flow simulation results at a basin scale and control the derivation of the model, in particular in forecasting mode.
The main idea underlining the project is that it could be beneficial for forecasting purposes to introduce a new type of information such as soil moisture into a catchment model in order to update its internal state variable and control its evolution during the forecast. The main feature of the assimilation methodologies set up during the project is that they don't restrict updating to recent stream flow observations, as do classic procedures in hydrology, but also to soil moisture measurements.
Two kind of soil moisture measurement were used in this project for assimilation purposes:
- Point measurements with Time Domain Reflecting (TDR) probes, providing continuous measurements of soil moisture at different depths and
- Soil moisture estimates that can be derived from microwave space borne Synthetic Aperture Radar (SAR) images. Recent studies have shown (Quesney *et al.*, 2000, Le Hegarat-Mascle *et al.*, 2000) that Earth Observation (EO) data could reliably provide spatial and temporal soil moisture information at the catchment scale.

Point measurements and EO soil moisture estimates can be found in the database set up for the project (http://dataserv.cetp.ipsl.fr/AIMWATER/).
The state and parameter updating methodologies procedures can use both *in situ* soil moisture data measurements and EO data soil moisture indicators with filtering techniques to update either internal states or parameters. However, EO data are too scarce for operational use (the satellite passes over the basin, in the same configuration, only once every 35 days). Therefore, TDR measurements were used to set up the methodology and then EO data were used in validation mode.

The scaling methodology can use either in situ information or EO data to calculate the observed wetness index, which would be compared with the modelled wetness index obtained from the distributed catchment scale model DiCaSM during the calibration process. A successful calibration of the DiCaSM model would produce a continuous time series of the wetness index (daily values) to replace the discontinuous / periodical wetness index obtained from remote sensing or weekly point measurements. The modelled daily time series wetness index (0-10) is then used as input in the R-R models (GR4J and IHACRES) to generate the stream flow.

4.2.3 Assimilating Soil moisture: Model's Adaptation

The methods summarized above provide a framework that can accommodate other sources of information. There are many solutions to the problem of dealing with the final error in stream flow value. In the AIMWATER project, the hypothesis is that soil moisture can provide the information to derive the best internal variable updates. The present research is aimed at analyzing the different ways to take soil moisture into account, and to highlight the one that offers the best prospects.
In order to integrate the soil moisture information in the assimilation methods presented above some model adaptation were carried out:

- Scaling method: adaptation of GR and IHACRES models.

According to the Chart in Annex 1. The Wetness Index (WI) is one of the DiCaSM model outputs used as input into two rainfall-runoff models (GR4J and IHACRES). These models have been adapted so that the WI is multiplied by the gross rainfall to obtain the effective rainfall, which is then used as input instead of the gross rainfall. Since these two models have a loss module, which produces something similar to the WI of DiCaSM, a relationship between the WI of DiCaSM and the WI of each other, model has been investigated. For that purpose, the WIs of the original GR4J and IHACRES models that gave the best fit during the simulation exercise, were compared with the WI of DiCaSM. The comparison showed that the WI of the DiCaSM was similar in trend but has higher values that could reach 1 in winter. Subsequently, it was decided to adjust / transform the wetness index that would be at the same scale of the wetness indices of both GR4J and IHACRES in order to ensure better simulation results. Thus the WI was transformed into $WI_{transformed}$ as shown in equation (2). The latter is then multiplied by the gross rainfall to obtain the effective rainfall that replaces the gross rainfall as input to the GR4J and IHACRES models as shown in Figure 4.

$$WI_{transformed} = (WI_{DiCaSM} / B)^{A} \qquad \textit{with A and B greater than 1} \qquad (2)$$

A and B were obtained through the calibration of the modelled stream flow against the observed stream flow for the whole period of stream flow record.

Rainfall-runoff model structures and adaptation

To adapt to the objectives of this study, we wanted to test rainfall-runoff models where production and routing functions can be easily distinguished, so that the production function can be replaced by a function of the soil moisture index calculated by the DiCaSM model without modifying the functions of the transfer module. Two model structures, called here GR3J and IHAC, were selected. They correspond to model structures recommended in earlier work of the AIMWATER modelling group as suitable for use in the context of the project. They are modified versions of the GR3J model proposed by Edijatno *et al.* (1999) and the IHACRES model proposed by Jakeman *et al.* (1990) respectively. A brief outline of the models is given here.

GR3J

The modified model version proposed by Perrin (2000) is used here. The water exchange function has been removed because it is part of the production function, but acts on the transfer routing part of the model. Hence the production and transfer functions of the model can be clearly distinguished. The model scheme is given in Figure 4. The tested version has three parameters (X1: capacity of the production store; X2: capacity of the routing store; X3: time base of the unit hydrograph). Only the two parameters X2 and X3 of the routing module are used in the scaling method.

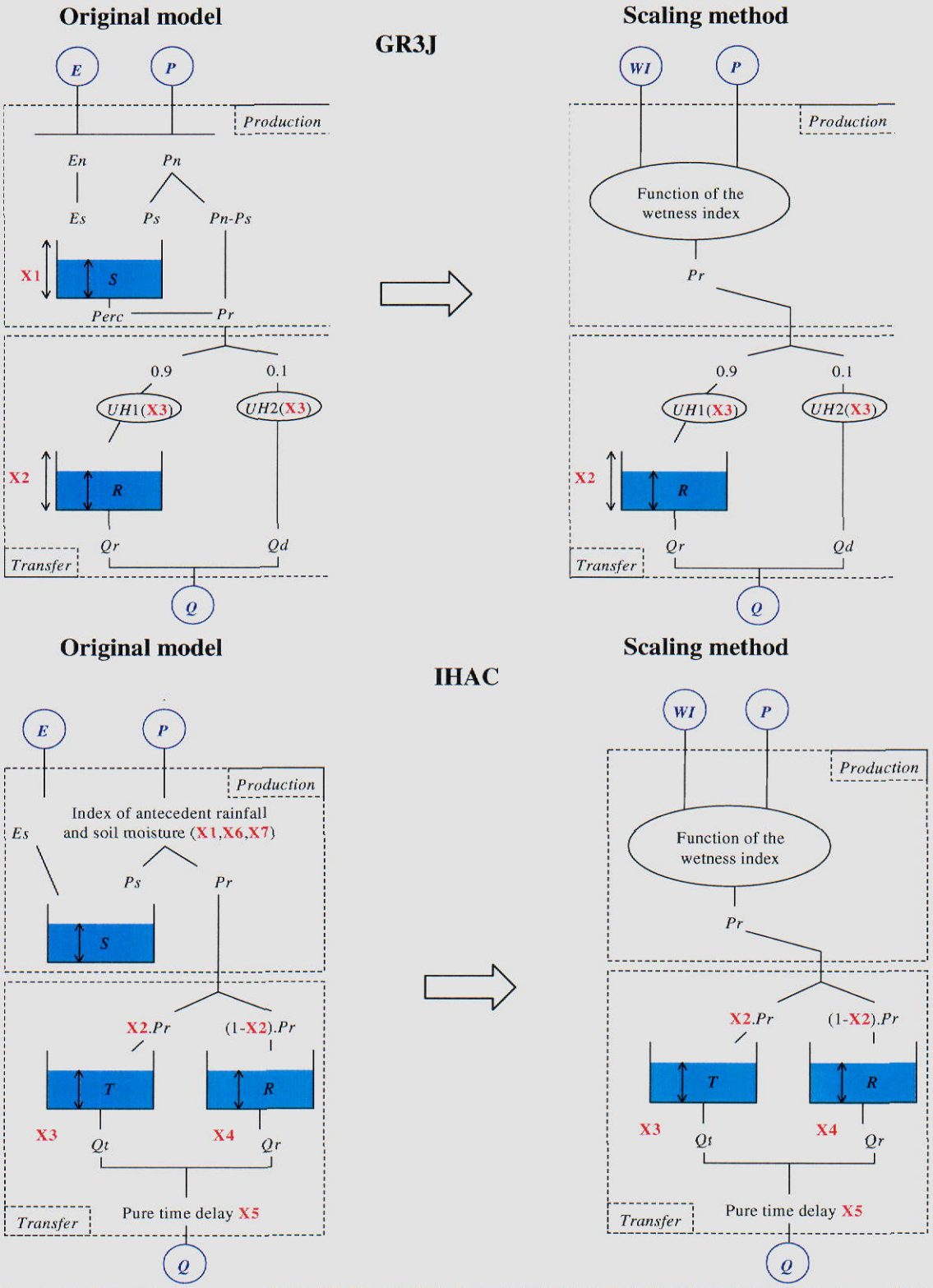


Figure 4: Schematic diagrams of the GR3J and IHAC models (original and scaling method configurations)

IHAC

A modified 7-parameter version was used here. The seven parameters are: X1: volume forcing constant; X2: splitting coefficient of effective rainfall; X3: constant of quick linear routing store; X4: constant of slow linear routing store; X5: pure time delay; X6: modulation factor; X7: catchment drying time constant). It retains the original model structure of IHACRES, in which the time delay is an additional parameter to be optimised. Four parameters (X2, X3, X4, X5) are included in the routing module as shown in Figure 4.
The transfer module of GR3J is globally non linear (unit hydrographs and non linear routing store) but in the IHAC it is linear (two linear routing stores in parallel).

In both models, the soil moisture accounting procedure is replaced in the scaling method by a function of the wetness index (WI), which is the output of the DiCaSM model, while the routing procedure of the models remain unchanged.

- *Variational method: a new constraint*

The most obvious solution to integrate soil moisture information would be to work out a combination criterion, to put together the final errors in streamflow and soil moisture values as determined by the model prior to updating the system. However, this change in the objective function should be minimized. Such a change could blur the comparison between the preliminary updating method and the present one. A more prudent approach is to only add a constraint such as:

$$\frac{1}{n}\sqrt{\sum_{j=1}^{n}\left(Ho_j - Hc_j\right)^2} < \frac{\sigma}{k} \qquad (3)$$

where *Ho* is the observed soil moisture, *Hc* is the modeled soil moisture (explicitly or implicitly), *n* the chosen number of soil moisture values taken into account, σ the standard error of *Hc* values, and *k* a parameter that allows the severity of the constraint to be varied so as to find out the importance that can be attributed to moisture observations. If k is very small (e.g. $k < 0.1$), the constraint fades away, and if k becomes higher (e.g. $k > 1$) the constraint is tighter: *k* is a measure of the interest of complying with observations of soil moisture.
The introduction of this new constraint has been tested on two models among those selected for the project: the first model, GRHUM (Loumagne *et al.*, 1996), has been especially developed to introduce a two-layer soil reservoir that simulates the surface and sub-surface soil moisture. The second one, GR4J (Edijatno *et al.*, 1999) has no explicit counterpart for soil moisture measurements.
Soil moisture values that are taken into account in equation (3) are first estimated, when GRHUM model is used, with an equation similar to equation (1) :

$$Ho_j = \gamma + \delta Hc_j + \eta \qquad (4)$$

where η is the error term with a standard deviation equal to σ.
When the GR4J model is used, observed soil moisture values are correlated with a store content value, for example, with the content (*S*) of the soil-reservoir, if it is the best variable correlated to observed soil moisture. The determination of such a regression equation is restricted to the days *j* when $C_{A,j}$ is above a threshold equal to twice the mean annual streamflow.
At time step *j*, shifts concerning all parameters are carried out successively and a new set is retained, until the next iteration, under two conditions:

a) the new set of parameters reduces the expression :

$$\left|Q_i - C_{B,i}\right| \qquad (5)$$

b) the new set of parameters satisfies the condition described by constraint (3).

If not satisfied, this condition rejects the corresponding parameter set. Although this parameter set could reduce the final error in stream flow, it is considered unsuitable because of the too large difference between the simulated and observed soil moisture. Increasing k from an insensitive low value to a larger one allows for a progressive enforcing of the condition (3) up to a value that would make the most from this new type of data. If, in the process, the method performance was not improved, this would be an indication that the additional information is not relevant for its use in flood forecasting (see flow chart Annex 2).

- ***Sequential method: determination of state and observation vectors***

The implementation of the extended Kalman filter in GR rainfall-runoff models in order to assimilate both soil moisture and stream flow measurements, lies in the determination of the state and observation vectors.

The vector of the observations has two coordinates: $\mathbf{z_k} = (w_k, q_k)$, w_k being the measured soil moisture at time k and q_k being the measured stream flow at time k. The state vector is given by the structure of the model. In GR4J, there are two explicit internal variables: the level s of the soil reservoir and the level r of the routing reservoir. Deep flows over the time delay are also integrated in the state vector. In the Grhum model two levels of the soil reservoir need to be considered *Ws* and *Wg*.

It is then necessary to determine the constraint relation between the observations and the internal variables. For the stream flow, there is identity between the observed flow and the simulated flow, so the observed stream flow is directly determined by the structure of the model: $q = q_r + q_d$. As the stream flow mainly depends on the level of the routing reservoir, we consider in the tangent linear model that q is only linked to b. As GR4J is a conceptual model, representing the physical processes with an extreme simplicity, the observed soil moisture w is related to the level of the soil reservoir but not directly. There is a linear relationship between w and s: $w = h_s \cdot s + h_0$, the coefficients h_S and h_0 being estimated for each basin. For GRhum this relationship is:

$w = h_s . W_s + h_g . W_g + h_0$

The last stage consists in determining the covariance error matrixes. The uncertainties on the observations are of the order of 5% and the errors on soil moisture data and streamflow data are supposed to be uncorrelated. The exact value is not really important, since only the ratio with the uncertainties on the model is significant. The uncertainties on the levels of the soil reservoir s and the routing reservoir b are also supposed to be uncorrelated and different values have been selected in order to test the efficiency of the assimilation procedure (see flowchart Annex 3).

4.2.4 Assimilation process

- *Studied basins and database*

Different sub-catchments within the Seine basin in France and one sub-catchment in the Arade basin in Portugal were selected for this study: the Serein catchment (1120 km²), the Grand Morin catchment (1070 km²), the Orgeval catchment (104 km²) being a sub-catchment of the Grand Morin, and the Petit Morin catchment (605 km²) that lies alongside the northern boundary of the Grand Morin. The Arade basin covers an area of 975 km² and lies close to the southern coast of Portugal. The studied area, which covers 257.5 km², is a sub-catchment located upstream the Funcho dam (see figure 1 in chapter 1).

Catchment	Streamflow	Rainfall	Evapo transpiration	Soil moisture (TDR)	Soil moisture (EO)
Serein	1955-2001	1955-2001 9 raingauges	1962-2001	1999-2001	1999-2001
Petit Morin	1970-2001	1970- 2001 7 raingauges	1962-2001	1997-2001	1997-2001
Grand Morin	1994-2001	1994-2001 8 raingauges	1962-2001	1997-2001	1997-2001
Orgeval	1963-2001	1963-2001 5 raingauges	1962-2001	1997-2001	1997-2001
Arade	1994-2000	1994-2000 1 raingauge	1963-2000	2000-2001	2000

Table 2: Periods of data record for the studied catchments

For the selected catchments, daily time-series of rainfall and stream flow were collected. Ten-day potential evapotrasnpiration data were used along with other meteorological data (wind speed, temperature, sunshine hours, relative humidity etc.). TDR soil moisture data were available over the test period on a 12-hour basis but were used at the daily time-step adopted in this study. Table 2 summarizes the database for the studied area described in (http://dataserv.cetp.ipsl.fr/AIMWATER/).

- *Calibration and test periods*

Variational and sequential methodologies

The GR models were tested for each basin and model parameters calibrated over a long period of record. A local search algorithm was used to optimise model parameter values.

The selected objective function is the Nash and Sutcliffe (1970) criterion calculated on daily stream flow, this criterion focus especially on floods. The Nash criterion can also be calculated on root-square transformed stream flow and logarithmic transformed stream flow, the criteria focussing respectively on mean stream flow and low flows. The calibration and test periods are detailed in table 3.

	Calibration period	Test period	Number of floods
Serein Lead time=2d	1955-1997	1999-2001	12
Petit Morin Lead time=2d	1970-1995	1997-2001	16
Grand Morin Lead time=2d	1994-1996	1998-2001	29
Orgeval Lead time=1d	1963-1995	1997-2001	38
Arade Lead time=1d	1994-1998	2000	7

Table 3: Calibration and test periods over the studied area

Model results in simulation mode for calibration periods are given in table 4. Both models exhibit similar results except for the Grand Morin basin during low flow periods. The Grand Morin stream flows are directly influenced by the releases of the Marne reservoir so the simulation for low flow periods are inadequate.

After calibrating the model parameters over a long period, test periods have been used to assimilate the soil moisture in a forecasting mode.

	GR4j			GRHum		
	Na (Q)	Na (√Q)	Na (lnQ)	Na (Q)	Na (√Q)	Na (lnQ)
Serein	88.2	88.7	81.7	87.6	85.1	71.5
Orgeval	77.8	78.9	70.8	78.3	77.4	68.9
Grand Morin	83.2	46.8	-112.5	86.6	61.7	-35.5
Petit Morin	81.9	84.1	76.2	77.8	78.8	69.5
Arade	58.9	49.5	-43.7	61.4	67.9	37.4

Table 4: Performances of GR4j and GRHum models in calibration mode for the studied area : Nash criteria values (%)

Scaling method

For the four catchments (Table5), soil moisture time-series were simulated by the DiCaSM model for the 5 and 10 cm soil layers, which were of interest in the project. Simulations were carried out using the 1996-1999 and 1996-1998 periods for the Serein and Grand Morin. For the Orgeval catchment, the same wetness index time series as for the Grand Morin was used. Potential evapotranspiration, rainfall and stream flow data for the same periods were used to run the hydrological rainfall-runoff models in calibration / simulation modes. To test the efficiency and robustness of the models, the records were divided into several periods, as shown in Table 5. The models were also tested over the whole period of record. For each period, a warm-up year was used to avoid problems of model initialisation.

The split-sample test scheme was adopted here to test the models. They were successively calibrated for each test period shown in Table 5 and then run in simulation mode for the other test periods.

Catchment	Total period	Sub-period1	Sub-period2	Sub-period3
GrandMorin	1996-1998	1996-1997	1997-1998	
Orgeval	1996-1998	1996-1997	1997-1998	
Serein	1996-1999	1996-1997	1997-1998	1998-1999
Arade	1994-2000	1994-1997	1997-2000	

Table 5: Test periods for the Serein, Grand Morin, Orgeval and Arade catchments

- *Procedure assessment*

Quality assessment in forecasting mode: variational and sequential methods

The assimilation procedure for the variational and sequential methods was run in forecasting mode in order to provide customers with results in an operational context.

In the forecasting mode it has been decided to treat all events during which the flow rate keeps in excess of a threshold equal to four times the mean annual. Even when the threshold is crossed for one time step the forecasting operation is carried out for this step and the result is tallied together with those from all other events. Both models used in association with the forecasting procedure were calibrated against the longest series of data available, thus yielding the baseline parameter set and, after a sensitivity analysis, the standard deviation of each parameter.

For each model, for each day a flow rate is forecast, the squared error of the forecast flow rate is recorded and summed up to form the persistence criterion (Wallis and Todini, 1975). The persistence criterion compares over the whole prediction period Q_{for} at day $j+L$ (where j is the day when the forecast is done) by the assimilation process with a "naïve" prediction supposing that the flow remains constant between day j and $j+L$:

$$\boldsymbol{Pers} = 100 \cdot \left(1 - \frac{\sum_j \left(Q_{for,j+L} - Q_{obs,j+L} \right)^2}{\sum_j \left(Q_{obs,j} - Q_{obs,j+L} \right)^2} \right). \qquad (6)$$

The efficiency criterion compares the forecast Q_{for} done in assimilation mode with Q_{sim} in simulation mode (without assimilation):

$$\boldsymbol{Eff} = 100 \cdot \left(1 - \frac{\sum_j \left(Q_{for,j+L} - Q_{obs,j+L} \right)^2}{\sum_j \left(Q_{sim,j+L} - Q_{obs,j+L} \right)^2} \right). \qquad (7)$$

As to the rainfall rates of the forecast period, perfect *a priori* knowledge was adopted, i.e. actual rainfall was used. Clearly, this is not an operational option, but it was adopted because it was deemed essential to best assess the effect of introducing soil moisture measurements without blurring the picture by additional sources of error

Quality assessment in simulation mode: scaling method

The scaling method was tested in simulation mode. To maintain consistency for all AIMWATER methodologies the classical Nash criterion was chosen as the objective function with which to calibrate the models. Four criteria were used in simulation mode to evaluate results: Nash(Q), Nash(√Q), Nash(ln(Q)) calculated on stream flow, root square transformed stream flow and logarithm transformed stream flow respectively, with the bias defined by:

$$bias\ (\%) = 100. \left(1 - \left| 1 - \frac{\sum_{i=1}^{n} Q_{sim}}{\sum_{i=1}^{n} Q_{obs}} \right| \right) \qquad (8)$$

where Q_{sim} and Q_{obs} are the simulated and observed streamflow and n is the number of time steps. This formula allows averages over several periods.

For each model structure, the assessment was divided into several phases as follows:
1. Assessment of the performances of original model structures
2. Choice of a scaling transformation
3. Assessment of the hydrological models using the wetness index and scaling method

The results are presented for each catchment.

The Serein catchment

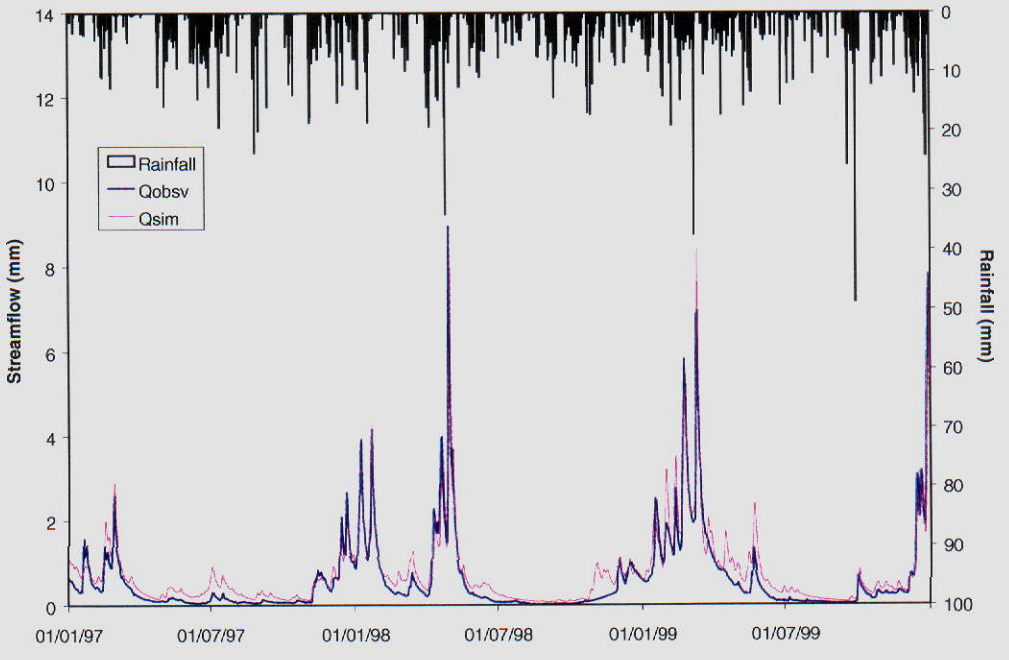
Assessment of the original models

The performances of GR3J and IHAC are summarised in Table 6. In this catchment the GR3J model performed better than the IHAC model; although both are able to reach similarly high performances, the GR3J model seems more robust with higher minimum performances.

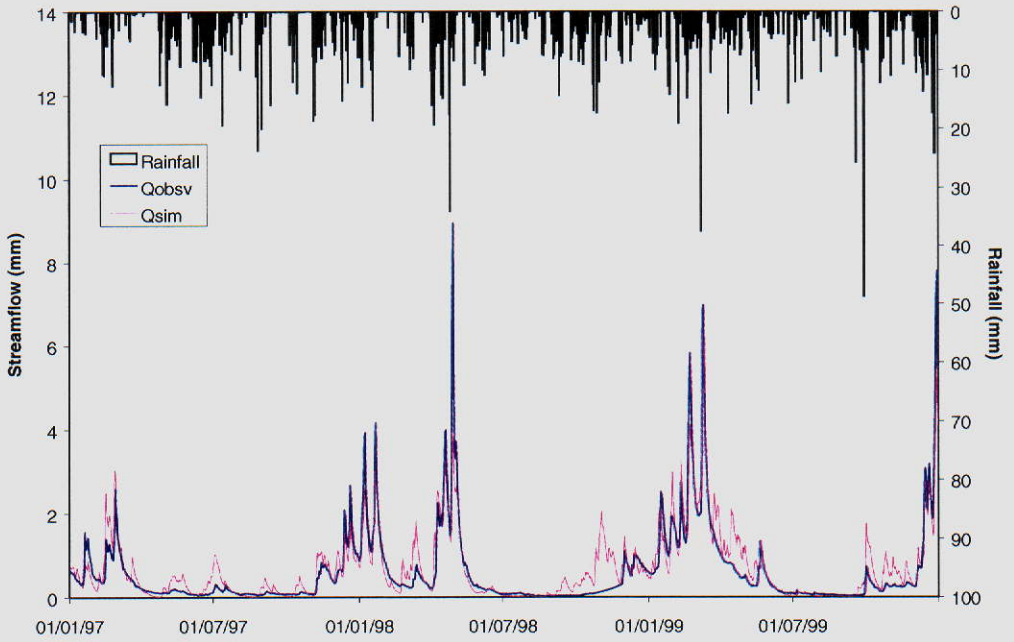
The mean performances will be taken as a reference to assess the worth of the scaling method. Note that the 4-parameter version of the GR model (which includes a water exchange term) performs better than the GR3J version used here, with mean criteria of 85.7, 83.7, 76.6 and 87.1 % respectively. An example of stream flow simulations is shown in Figure 5 for both models.

		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	84.0	77.8	64.7	76.7
	Maximum	91.6	87.3	80.1	90.2
	Minimum	68.3	56.7	36.9	54.3
IHAC	Mean	58.9	62.1	57.6	78.4
	Maximum	86.0	85.2	80.3	98.2
	Minimum	23.8	36.9	27.8	44.8

Table 6: Performances of GR3J and IHAC in simulation mode on the Serein catchment



(a)



(b)

Figure 5: Streamflow simulation on the 1997-1999 period for (a) GR3J model (Nash=89.5 %) and (b) IHAC model (Nash=79.8 %)

Choice of a scaling transformation

Figure 6 shows the wetness index simulations of the DiCaSM model; figure 7 presents the evolution of soil moisture (or index) in the GR3J and IHAC models. In the scaling method, a correspondence was sought between the soil moisture accounting procedure of the original rainfall-runoff models and the externally simulated wetness index, so that the latter can replace the former in the hydrological models.

When comparing Figures 6 and 7, it can be seen that the DiCaSM wetness index frequently reaches values close to 1 in winter time (saturation of the top soil layer), whereas the moisture indices of the rainfall-runoff models, which account for a catchment soil moisture over the whole trench of soil, reaches lower top values.

Therefore *a prior* transformation must be applied to the DiCaSM WI before it is used to substitute the rainfall-runoff model soil moisture accounting procedure. The proposed transformation depends on two parameters (A and B) as given in Equation (2). A and B were calibrated over the whole period of record for both WI5 and WI10. The optimisation was carried out for GR3J and IHAC, since the transformations are dependent on the rainfall-runoff model used (more precisely, the selected routing procedure). Optimised values are given in Table 7. Except in the case of parameter B for IHAC and WI10, all the other values for A and B respectively are very close to each other. Figure 8 illustrates the evolution of the transformed values of WI10 for GR3J and IHACRES. Both curves are very similar.

		A	B
GR3J	WI5	1.14	3.78
	WI10	1.12	3.69
IHAC	WI5	1.16	3.77
	WI10	1.10	4.61

Table 7: Values of parameters A and B of the scaling method for the GR3J and IHAC models, and both wetness indices WI5 and WI10

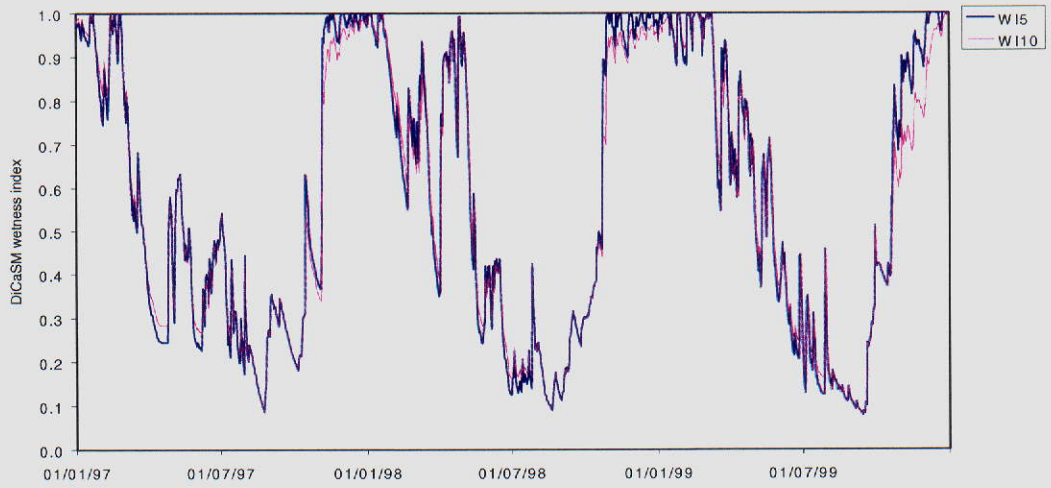


Figure 6: Evolution of the DiCaSM wetness index for the top 5 and 10 cm soil layers (WI5 and WI10)

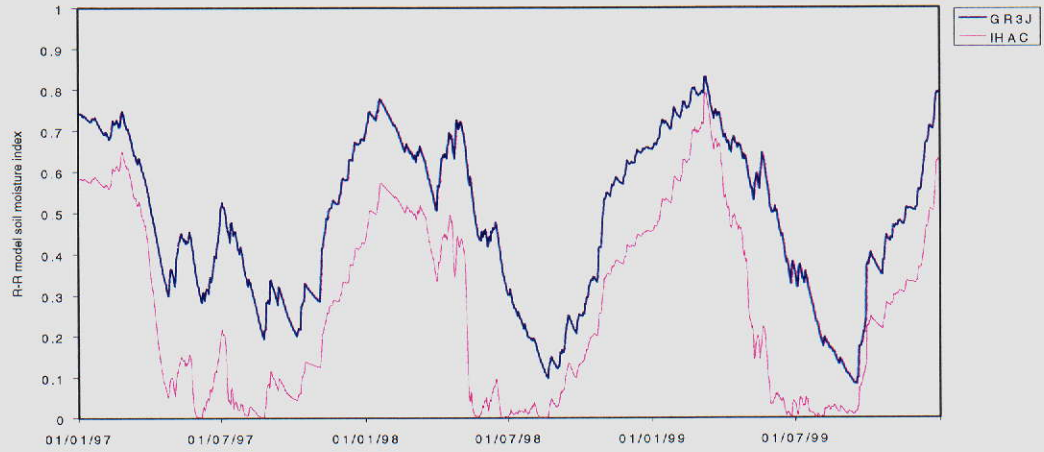


Figure 7: Evolution of the soil moisture indices of the GR3J and IHAC models

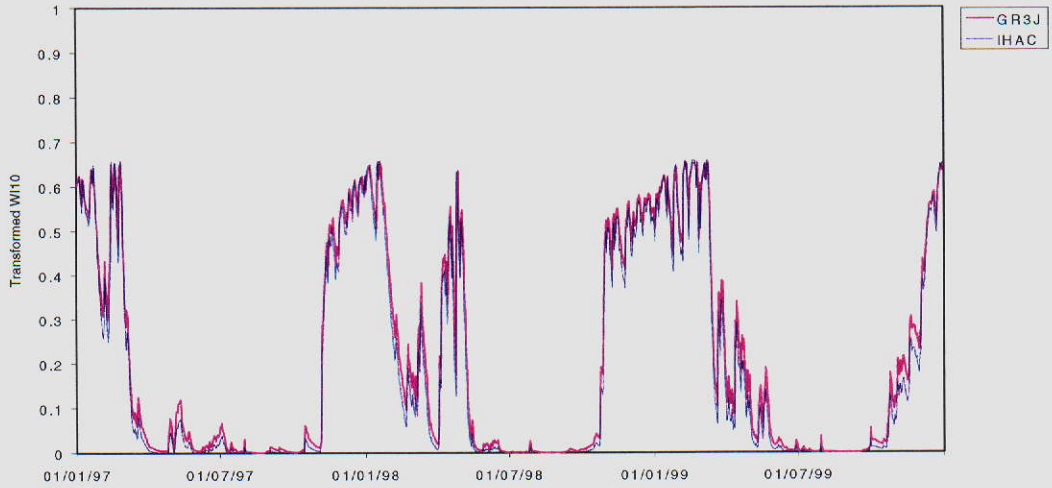


Figure 8: Evolution of the transformed WI10 for GR3J and IHAC

The Grand Morin catchment

Assessment of the original models

Results are given in Table 8. Here performances of both models are almost identical in terms of flood simulation (Nash(Q)). However IHAC results are better than the GR3J model for low flow simulation. The GR3J model could not manage to sustain low flows for the 1997-1999 period, which yields very low values of Nash(ln(Q)) (Figure 9). In contrast, IHAC gave a better fit. Note that the use of the water exchange term (fourth parameter) in the GR model does not change the results, with a mean criteria of 78.4, 54.7, -59.2 and 88.1 % respectively. Optimising the splitting coefficient of effective rainfall (set to 10 % - 90 %) does not change the results either.

		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	78.5	54.8	-58.5	88.1
	Maximum	86.6	66.8	-11.3	96.5
	Minimum	70.3	45.0	-104.4	81.0
IHAC	Mean	79.1	69.3	50.1	93.6
	Maximum	89.9	79.7	65.5	98.7
	Minimum	67.8	60.2	29.8	86.2

Table 8: Performances of GR3J and IHAC in simulation mode on the Grand Morin catchment

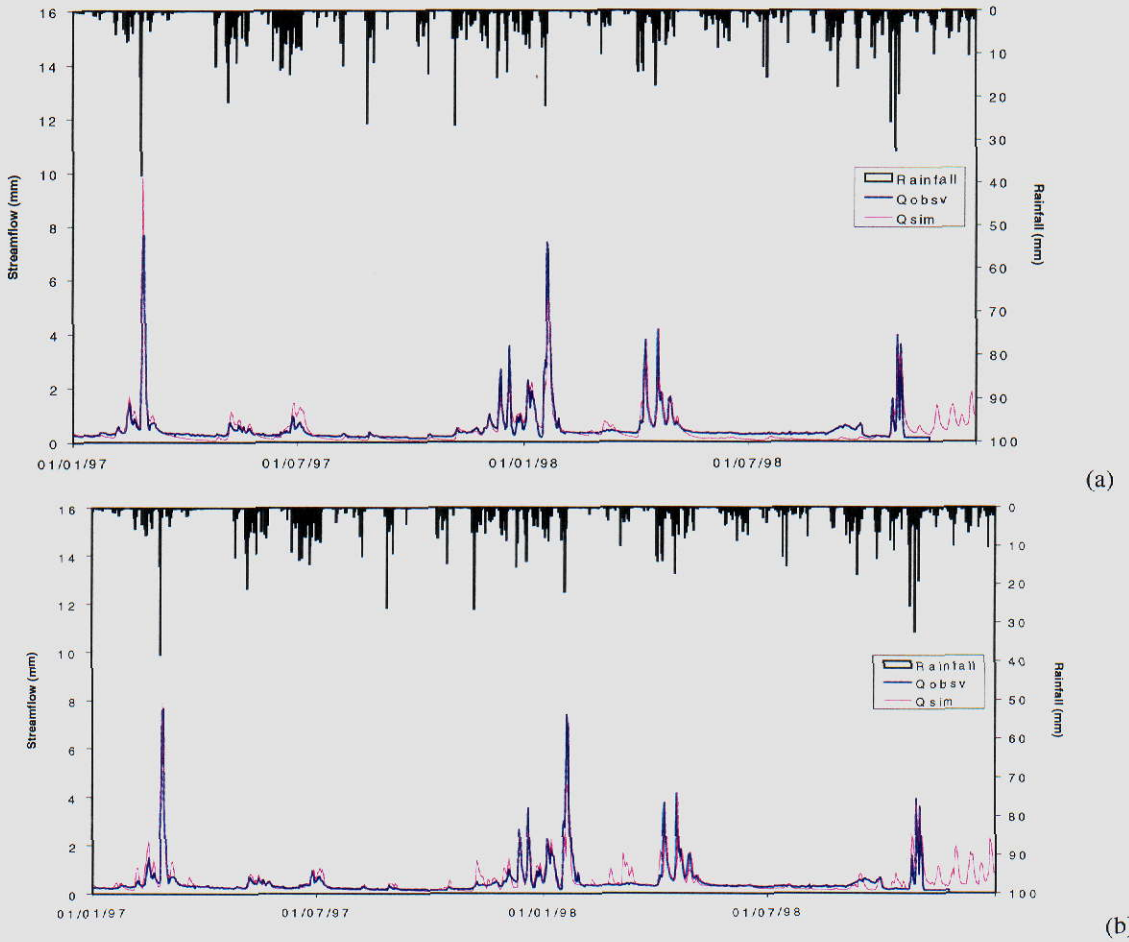


Figure 9: Streamflow simulation on the 1997-1998 period for (a) GR3J model (Nash=82.9 %) and (b) IHAC model (Nash=80.2 %)

Choice of a scaling transformation

The same type of wetness index transformation used for the Serein catchment was applied. Figure 10 shows the evolution of the wetness indices simulated by the DiCaSM model, while Figure 11 shows those of the rainfall-runoff models GR3J and IHAC. Here the DiCaSM time-series reaches 1 less frequently than the simulations of the Serein catchment.

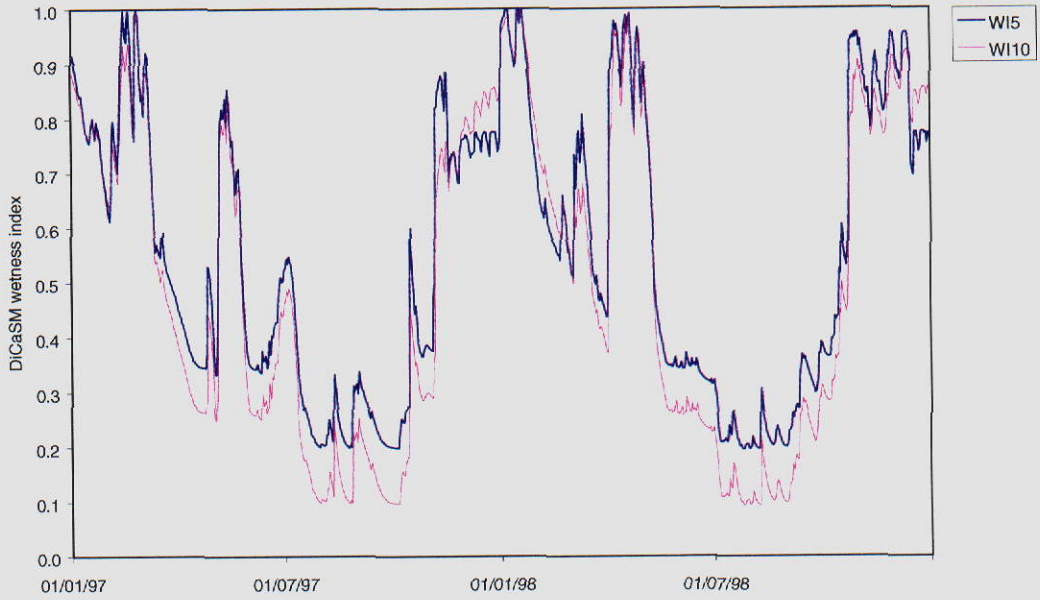


Figure 10: Evolution of the DiCaSM wetness index for the top 5 and 10 cm soil layers (respectively WI5 and WI10)

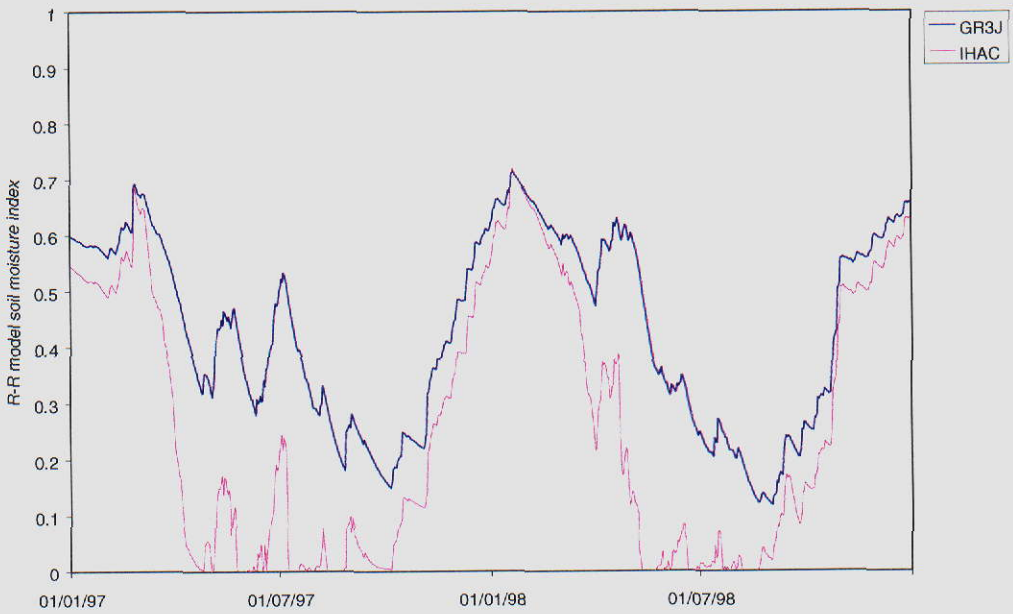


Figure 11: Evolution of the soil moisture indices of the GR3J and IHAC models

The parameters A and B of the scaling method were calibrated for the whole period of record. Parameter values are given in Table 9. They are significantly more heterogeneous than in the case of the Serein catchment.

		A	B
GR3J	WI5	1.42	2.45
	WI10	1.31	2.72
IHAC	WI5	1.04	6.37
	WI10	1.01	5.60

Table 9: Values of parameters A and B of the scaling method for the GR3J and IHAC models, and both wetness indices WI5 and WI10

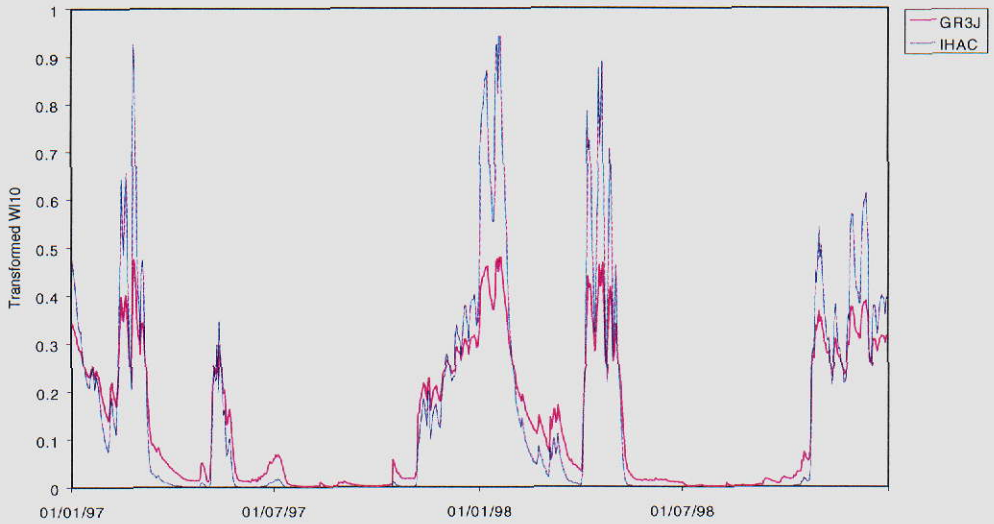


Figure 12: Evolution of the transformed WI10 for GR3J and IHAC

Figure 12 shows that contrary to the Serein catchment, here the scaling formulae yield very different transformed WI10.

The Orgeval catchment

Assessment of the original models

Results are given in Table 10. The GR3J model performs better than the IHAC model. An example of flow simulations is given in Figure 13 for both models. Here there are no problems of low flow simulations in contrast to the Grand Morin catchment. The use of the water exchange term (fourth parameter) in the GR model does not change the results a great deal (except for bias), using mean criteria of 87.1, 86.5, 80.3 and 96.3 % respectively.

		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	87,8	86,0	82,0	88,6
	Maximum	93,0	87,7	83,6	95,7
	Minimum	81,4	83,1	80,7	81,1
IHAC	Mean	71,9	65,7	51,8	87,2
	Maximum	83,9	71,3	66,0	92,9
	Minimum	53,8	53,6	32,1	75,7

Table 10: Performances of GR3J and IHAC in simulation mode on the Orgeval catchment

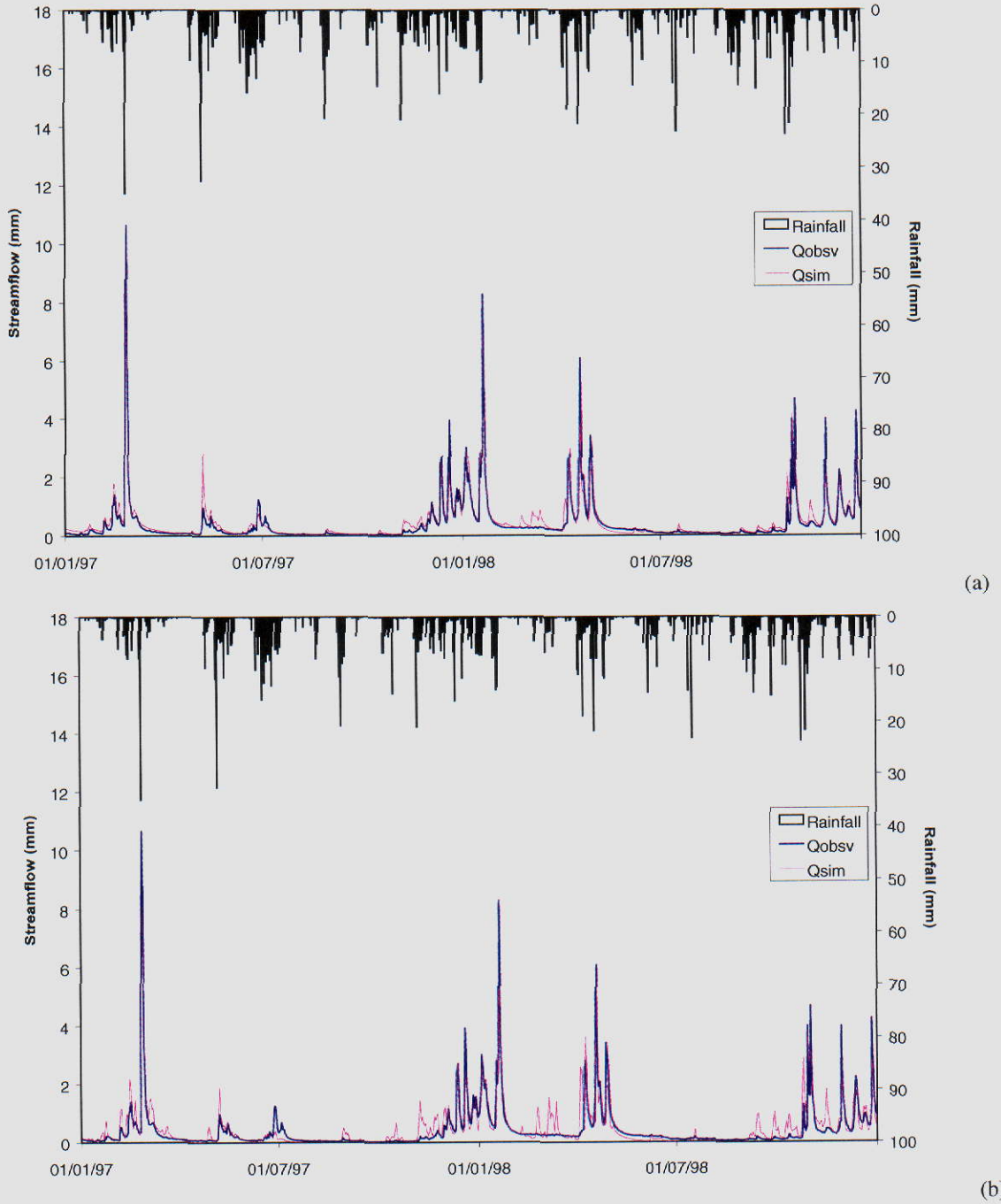


Figure 13: Streamflow simulation on the 1997-1998 period for (a) GR3J model (Nash=89.1 %) and (b) IHAC model (Nash=77.3 %)

Choice of a scaling transformation

The DiCaSM wetness index time series used here are the same as those used for the Grand Morin catchment (see Figure 10). Figure 14 shows the evolution of the soil moisture states of the rainfall-runoff models GR3J and IHAC. They are quite similar to those obtained for the Grand Morin catchment (see Figure 11).

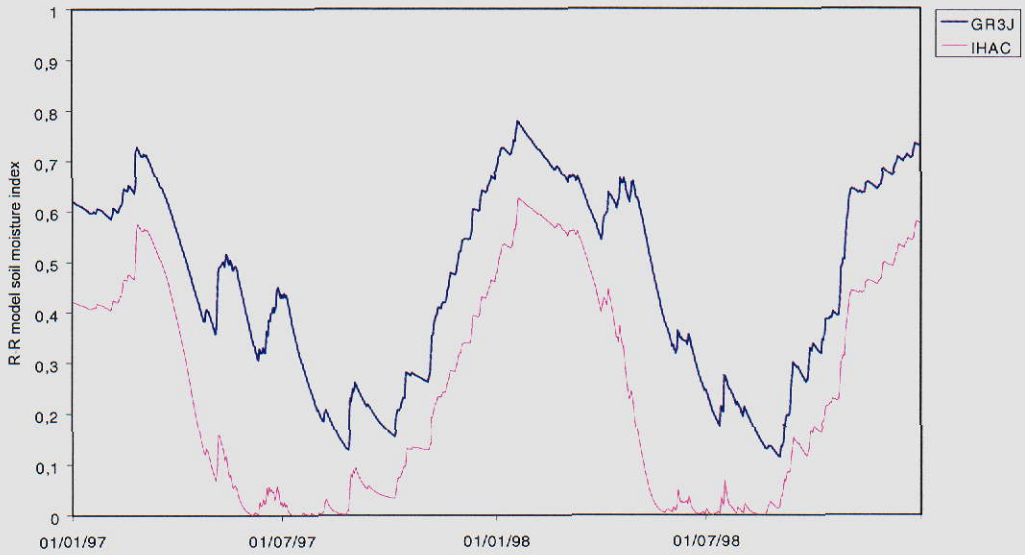


Figure 14: Evolution of the soil moisture indices of the GR3J and IHAC models

The parameters A and B of the scaling method were calibrated on the whole period of record. Parameter values are given in Table 11 and the corresponding transformed wetness index time series for the 10 cm depth are shown in Figure 15.

		A	B
GR3J	WI5	1,20	3,62
	WI10	1,18	3,36
IHAC	WI5	1,10	4,88
	WI10	1,07	4,89

Table 11: Values of parameters A and B of the scaling method for the GR3J and IHAC models, and both wetness indices WI5 and WI10

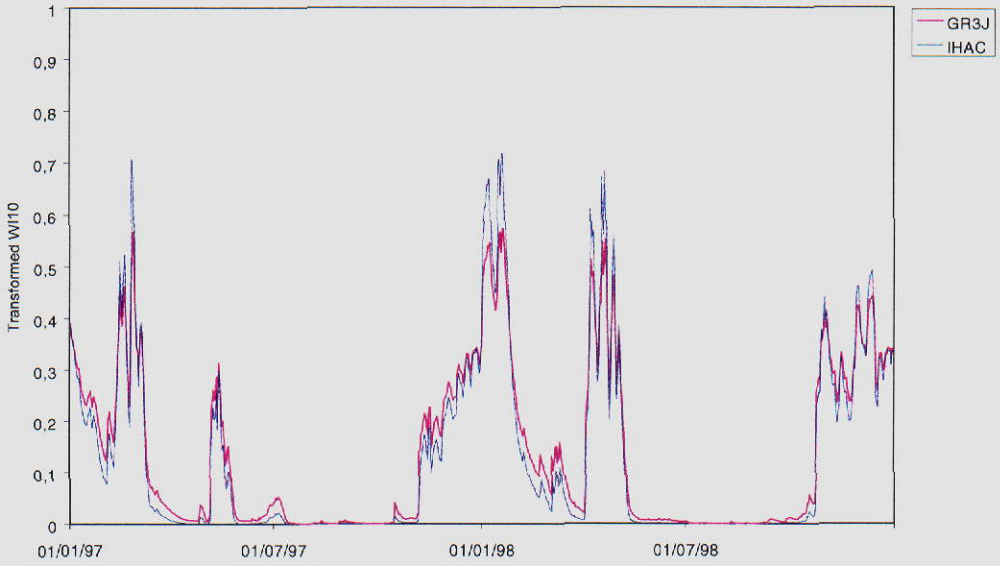


Figure 15: Evolution of the transformed WI10 for GR3J and IHAC

The Arade catchment

In addition to the whole period 1994-2000, two sub-periods were considered, 1994-1997 and 1997-2000. Stream flow data are derived from water height measurements in the Funcho Dam and might therefore lack precision. It should be noted that rainfall intensities here may be considerably higher here than in the Seine basin. The catchment has an ephemeral behaviour, with periods of zero stream flow in the dry season. The average water yield coefficient is less than 5 %. (mean annual runoff: 26 mm; mean annual rainfall: 730 mm; mean annual PE: 2230 mm).

Assessment of the original models

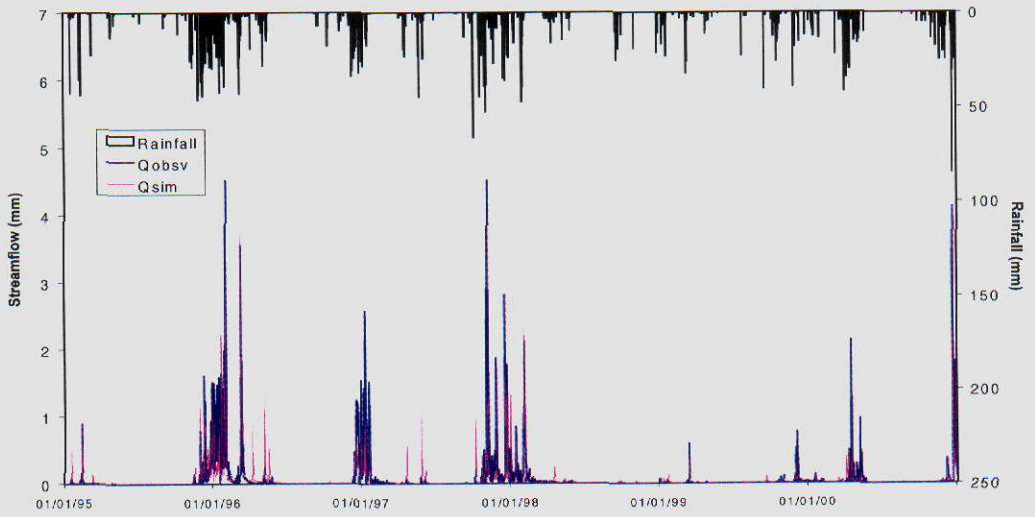
Results are given in Table 12. GR3J and IHAC performed poorly in terms of Nash criterion on this catchment. However, it must be remembered that such a criterion is not well adapted to assess model performances on low-yielding catchments. Although models seem to be unsuccessful for flood simulations (models generally underestimate flood peaks), they manage to give the right order of magnitude of stream flow. There was thus a need to add a complementary criterion, called Nashbis, defined by:

$$\text{Nashbis} = 1 - \frac{\sum_{i=1}^{n} \left(Q_{obs,i} - Q_{calc,i}\right)^2}{\sum_{i=1}^{n} \left(Q_{obs,i} - \hat{M}\right)^2} \qquad \text{where} \quad \hat{M} = \frac{1}{4}\left(\frac{P^2}{P+E}\right) \tag{9}$$

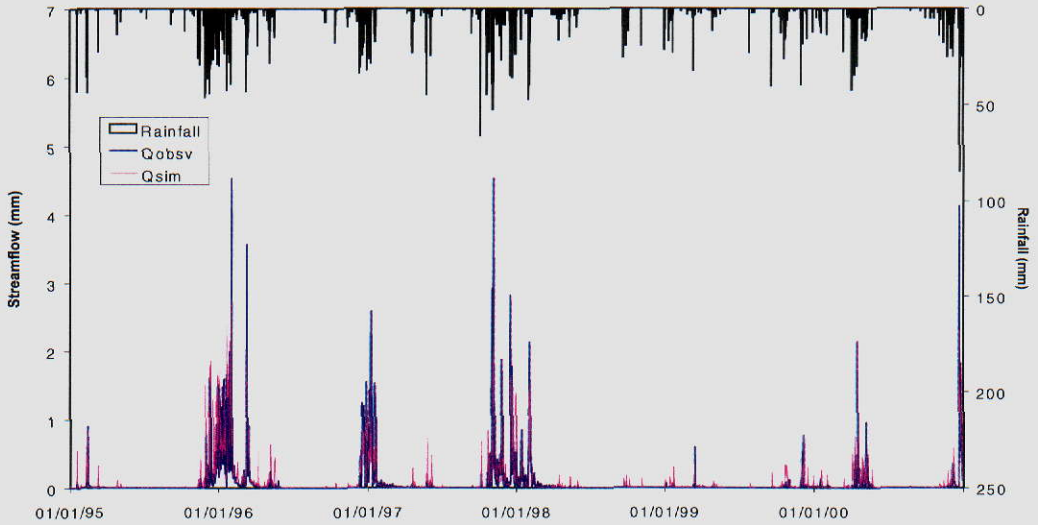
where P is rainfall and E potential evapotranspiration (see Perrin, 2000). With this criterion, results are more satisfactory.

		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Nashbis (%)	Bias (%)
GR3J	Mean	57,1	62,8	56,9	79,5	81,7
	Maximum	66,8	67,3	67,0	87,0	94,9
	Minimum	47,8	57,2	46,1	73,2	71,5
IHAC	Mean	58,6	54,4	39,6	80,2	81,6
	Maximum	68,6	72,5	68,1	87,7	91,7
	Minimum	39,9	40,2	15,2	69,1	70,4

Table 12: Performances of GR3J and IHAC in simulation mode on the Arade catchment



(a)



(b)

Figure 16: Streamflow simulation on the 1995-2000 period for (a) GR3J model (Nash=56.8 %; Nashbis=78.3 %) and (b) IHAC model (Nash=65.8 %; Nashbis=82.9 %)

An example of flow simulations is given in Figure 16 for both models. The use of the water exchange term (fourth parameter) in the GR model does not improve the results, with five mean criteria of 58.2, 33.3, -20.3, 80.1 and 72.5 % respectively. It should be noted that the parameters of the routing modules of both models showed the catchment response to be extremely quick. In GR3J, the capacity of the routing store (close to zero) and the time base of the unit hydrograph (close to the minimum value of 0.5 day) make the response almost instantaneous. In IHAC, the linear store for the quick flow component (86 % of total flow) empties almost instantaneously and the pure time delay is close to zero. This means that both models behave as if they had an almost instantaneous transfer. Their routing module therefore becomes useless and plays no role, which is one reason why model results are very close (model differ only by their production modules). The actual catchment response time is probably less than one day. Therefore, a shorter time-step than a day i.e. hourly could be more suitable to model the flood dynamics in this catchment.

Choice of a scaling transformation

The DiCaSM wetness time series used here are shown in Figure 17.

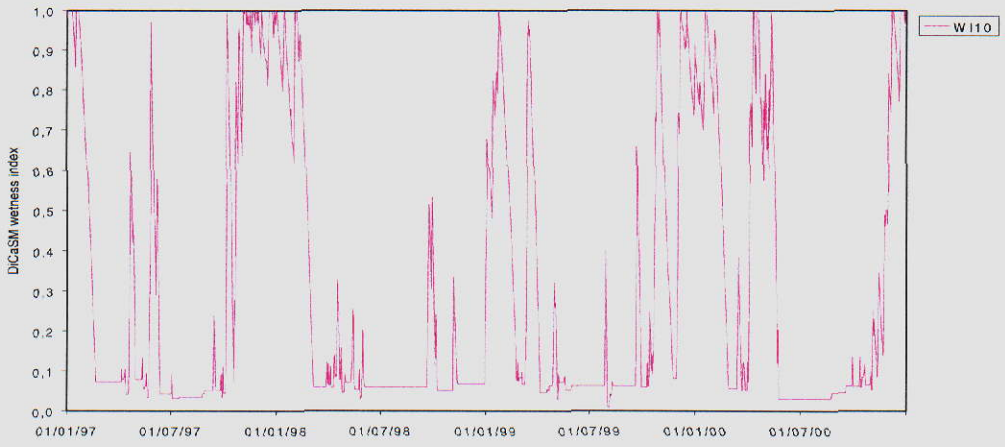


Figure 17: Evolution of the DiCaSM wetness index for the top 10 cm soil layers (WI10)

Figure 18 shows the evolution of the soil moisture status of the rainfall-runoff models GR3J and IHAC. They have markedly different behaviour in both range and shape. The evolution of the soil moisture variables of the IHAC and GR3J models are much smoother than the DiCaSM WI. This is because the optimised soil moisture store capacities (or equivalent) in both models are very large, leading to a smooth evolution throughout the year. Such large capacities indicate that the splitting coefficient of gross rainfall, which depends in both models on these moisture internal variables, does not vary much over the year. So this splitting is not dependent on the moisture status of soils, but rather on the Hortonian process and its possible link to shallow soils and steep slopes of the catchment . The values of the IHAC and GR3J internal variables remain in a narrow range (between 0 and 0.1 for IHAC and between 0.05 and 0.35 for GR3J) whereas DiCaSM values are often close to 1.

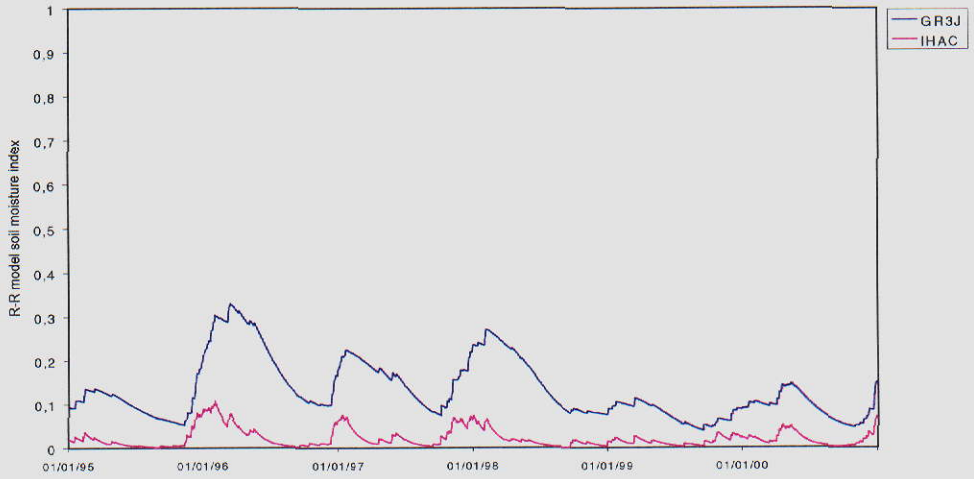


Figure 18: Evolution of the soil moisture indices of the GR3J and IHAC models

Parameters A and B of the scaling method were calibrated over the entire period of record. Parameter values are given in Table 13 and the corresponding transformed wetness index time series for the 10 cm depth are shown in Figure 19. Although both pairs of parameters are different, they yield quite similar transformed WI values.

	A	B
GR3J	2,28	3,36
IHAC	3,14	2,38

Table 13: Values of parameters A and B of the scaling method for the GR3J and IHAC models, and both wetness indices WI5 and WI10

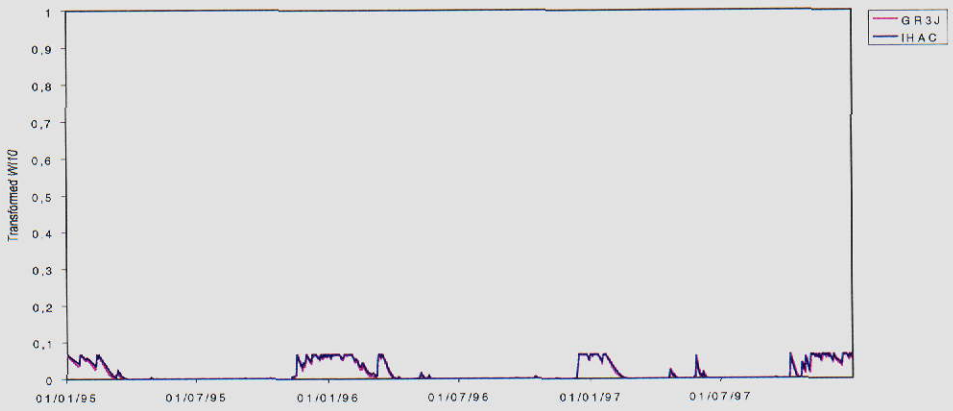


Figure 19: Evolution of the transformed WI10 for GR3J and IHAC

4.2.5 Soil Moisture Assimilation: Results

In the following sections, the results of soil moisture assimilation in the R-R models using three approaches are reported. The results of the scaling methods for the Grand Morin, Serein, Orgeval and Arade are given in the first section while in the second section the results of the parameter updating method used for the Serein, the Grand Morin and Arade basins are shown. Finally, in the third section, the results of the sequential method are reported.

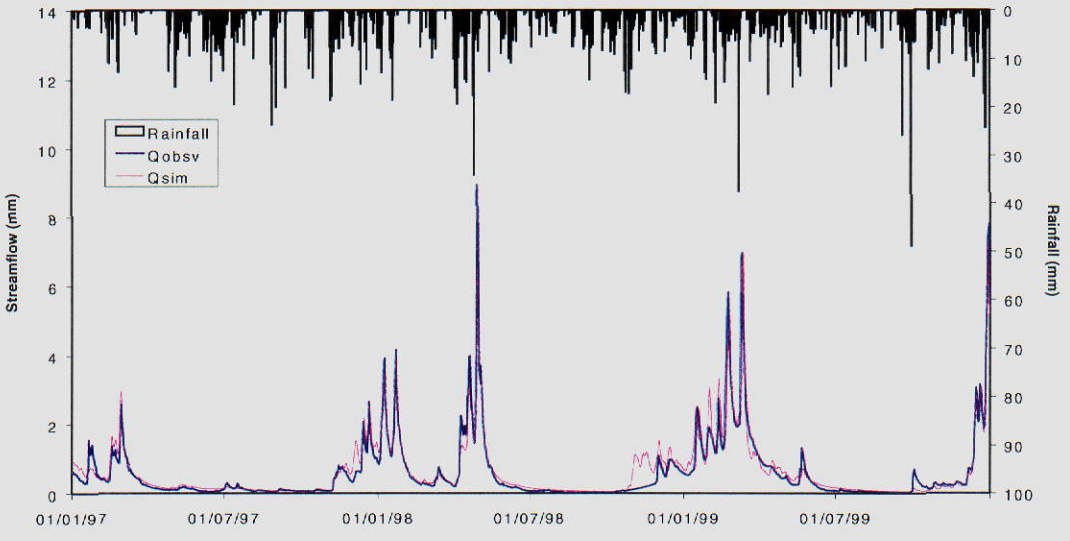
- *The Scaling Method*

The Serein Catchment

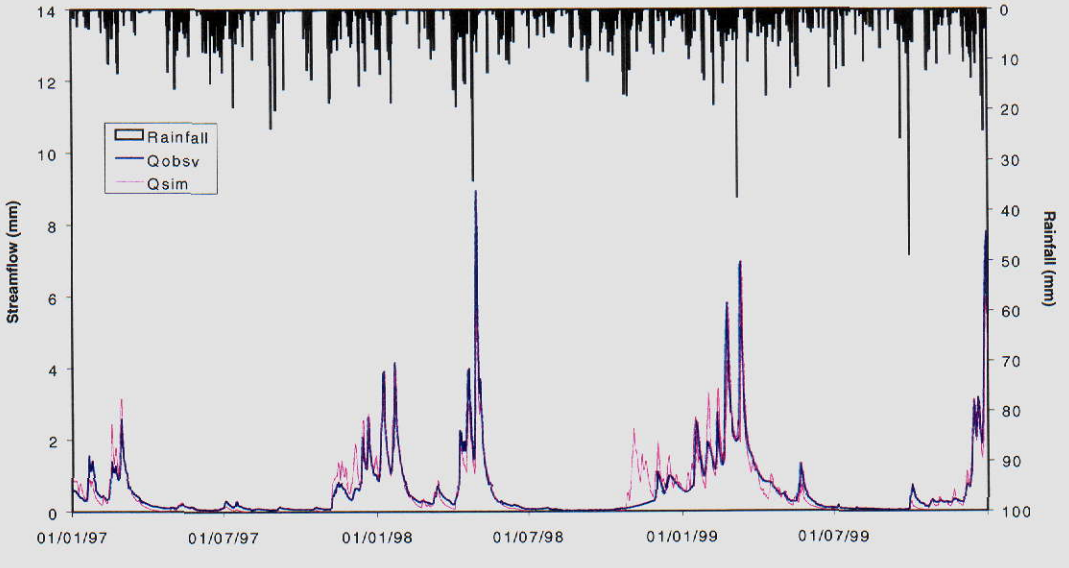
Results presented in Table 14 indicate that better performances can be achieved with the wetness index of the 10 cm layer than with the 5 cm layer (except in the case of IHAC for Nash(ln(Q)). In that case, a significant improvement was obtained in comparison with results of original model (Table 6): the mean Nash criterion is increased by 4.5 % for the GR3J model and by 18.9 % for the IHAC model. The improvements are also important for the other criteria (except in the case of IHAC for Nash(ln(Q)).

WI5		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	82.0	82.8	79.2	87.2
	Maximum	91.6	92.2	88.3	99.9
	Minimum	71.4	72.7	61.4	76.2
IHAC	Mean	65.1	72.1	63.9	88.5
	Maximum	84.6	88.0	86.7	97.0
	Minimum	38.0	51.3	14.8	74.9
WI10		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	88.5	88.4	84.6	90.0
	Maximum	94.1	93.4	90.0	98.0
	Minimum	78.2	84.1	75.1	81.6
IHAC	Mean	77.8	76.2	47.0	92.4
	Maximum	89.7	88.7	81.8	98.6
	Minimum	65.0	57.7	-12.6	79.3

Table 14: Performances of GR3J and IHAC using the scaling method with adjusted WI5 and WI10 on the Serein catchment in simulation mode



(a)



(b)

Figure 20 Stream flow simulation on the 1997-1999 period for (a) GR3J model (Nash=92.0 %) and (b) IHAC model (Nash=86.0 %) with the scaling method and WI10

In spite of this large improvement for the IHAC model in flood simulations, its performances remain a little lower than those of the GR3J model on this catchment, especially for low flow simulations. A better water balance is also obtained for both models (in the case of the GR3J model, the mean bias is not much different from the mean bias obtained with the model version that includes the water exchange term). Figure 20, by comparison with Figure 5, shows the differences obtained in stream flow simulation: low flows seem better simulated than the original models, especially in the case of GR3J.

The Grand Morin Catchment

Table 15 summarises the results obtained by both models. Although the advantage of using WI10 instead of WI5 for the IHAC model is clear, there are less obvious results for the GR3J model. When comparing these results with those of the original models in Table 8, it can be noted that improvements are obtained for the IHAC model when using WI10, whereas the application of the scaling method to the GR3J model did not improve the results. Low flows are still a problem (Figure 21).

WI5		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	67.4	27.8	-170.9	77.0
	Maximum	79.4	43.4	-105.9	77.1
	Minimum	55.7	11.3	-239.1	76.8
IHAC	Mean	75.3	64.2	37.9	87.9
	Maximum	80.5	66.5	45.5	91.1
	Minimum	69.8	59.9	24.7	83.1
WI10		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	72.4	14.4	-318.7	71.5
	Maximum	83.4	26.4	-252.2	71.5
	Minimum	65.8	1.9	-389.3	71.4
IHAC	Mean	86.0	78.0	56.1	89.8
	Maximum	89.4	79.1	61.4	94.4
	Minimum	82.8	76.8	48.8	83.8

Table 15: Performances of GR3J and IHAC using the scaling method with adjusted WI5 and WI10 on the Grand Morin catchment in simulation mode.

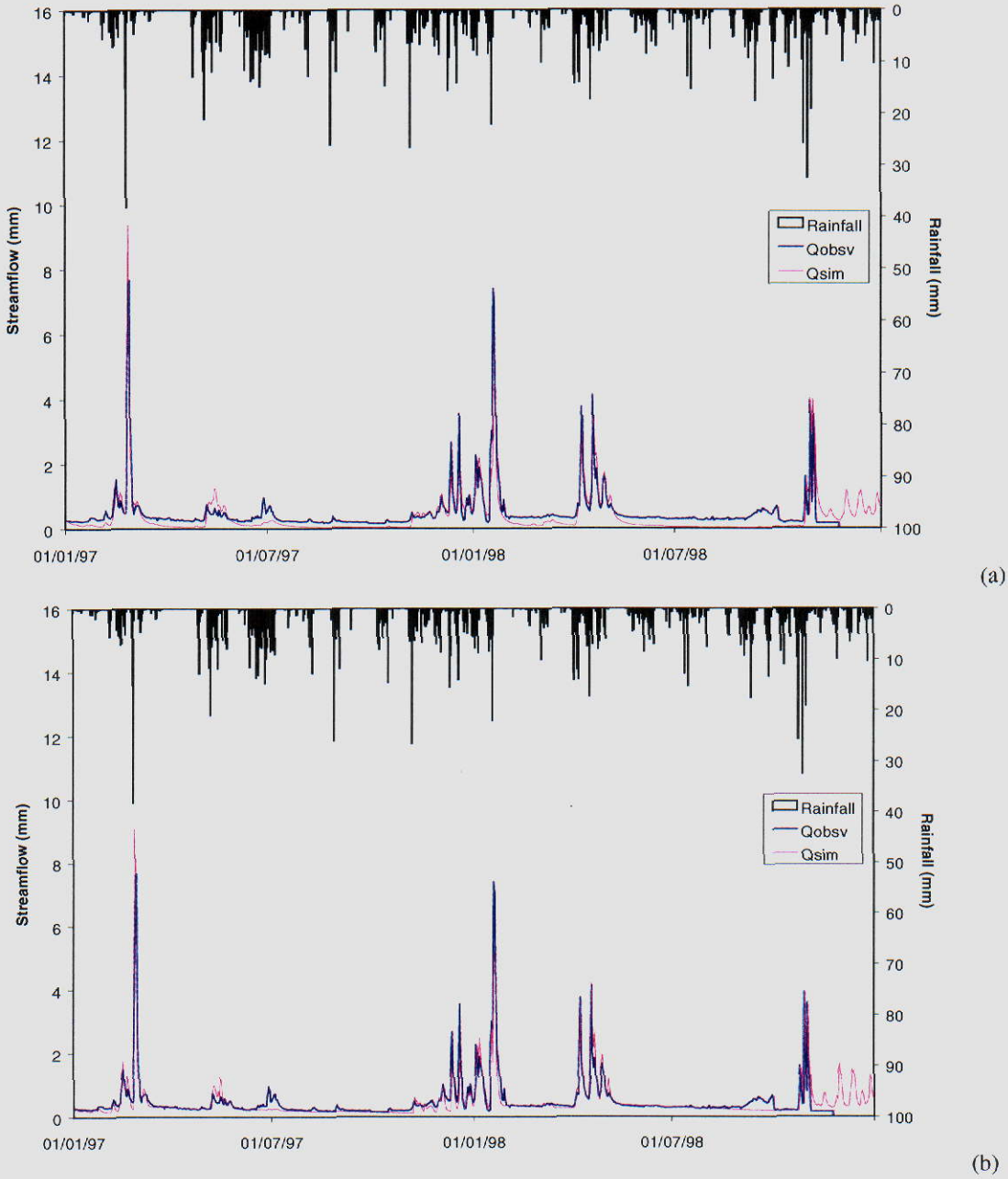


Figure 21: Stream flow simulation on the 1997-1998 period for (a) GR3J model (Nash=75.7 %) and (b) IHAC model (Nash=87.1 %) with the scaling method and WI10

The Orgeval Catchment

Table 16 summarises the results obtained by both models. Here again, it seems more efficient to use WI10 instead of WI5. The GR3J model remains more satisfactory than IHAC model. However, in comparison to Table 10, it can be noted that the use of simulated WI with the GR3J model does not bring any improvement especially for the low flow simulations. In the case of IHAC, a substantial improvement can be achieved on flood peak simulations; however, low flows are less satisfactorily simulated. Examples of stream flow simulations by the scaling method are shown in Figure 22.

WI5		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	77,7	74,2	58,5	88,3
	Maximum	82,6	76,8	70,7	88,9
	Minimum	72,8	72,2	39,9	87,8
IHAC	Mean	73,5	67,1	32,3	94,1
	Maximum	77,7	75,3	52,6	97,1
	Minimum	67,7	61,8	6,6	90,9
WI10		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Bias (%)
GR3J	Mean	86,7	81,0	52,6	88,3
	Maximum	90,4	82,0	69,1	91,0
	Minimum	83,3	79,9	32,4	86,1
IHAC	Mean	85,2	74,7	9,3	90,1
	Maximum	88,9	79,8	35,8	92,8
	Minimum	81,0	71,4	-29,4	87,1

Table 16: Performances of GR3J and IHAC using the scaling method with adjusted WI5 and WI10 on the Orgeval catchment in simulation mode

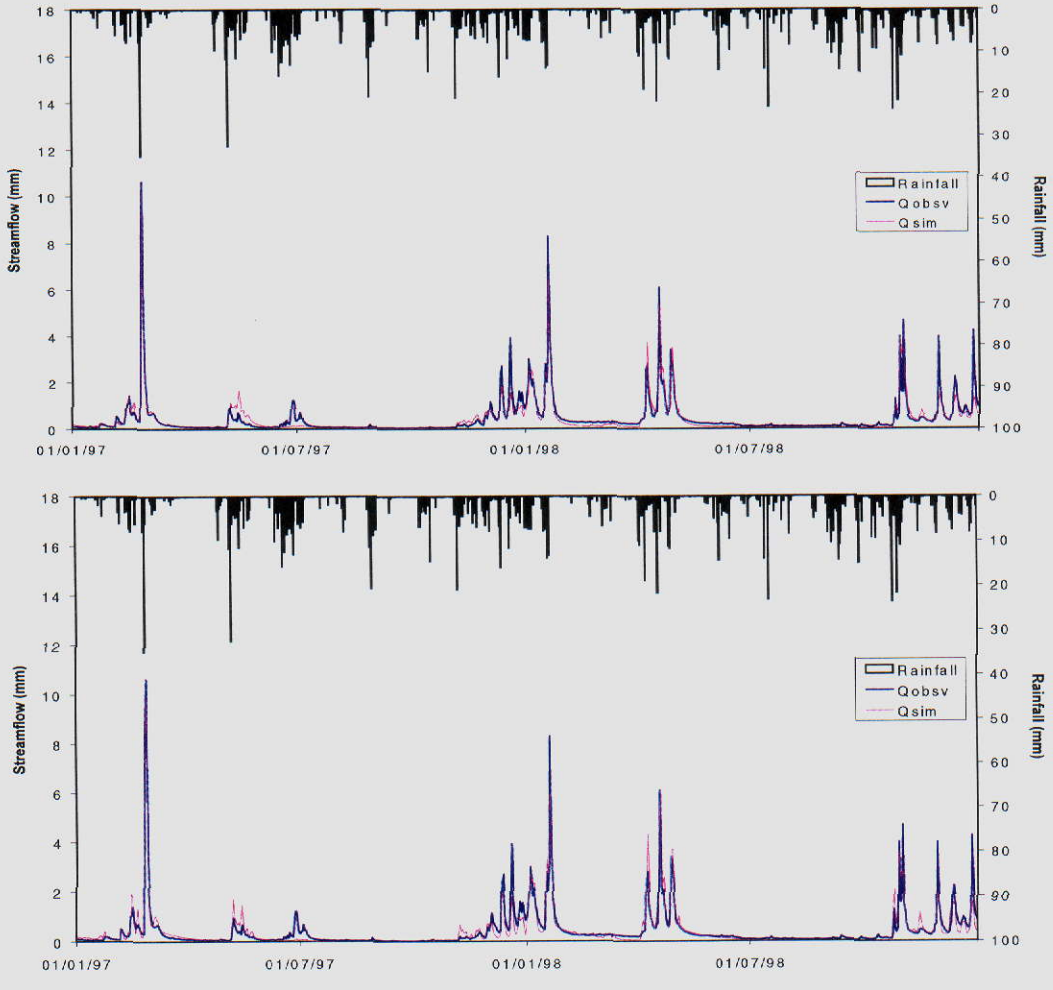


Figure 22: Stream flow simulation on the 1997-1998 period for (a) GR3J model (Nash=87.0 %) and (b) IHAC model (Nash=86.3 %) with the scaling method and WI10

The Arade Catchment

Table 17 summarises the results obtained by both models. Here again, performances of both models are moderate in terms of the Nash criterion. On the other hand the Nashbis criterion remains good and is even slightly improved. However, the improvement brought by the scaling method in this catchment is moderate, as some criteria did not improve. Examples of stream flow simulations by the scaling method are shown in Figure 23.

		Nash(Q) (%)	Nash(VQ) (%)	Nash(ln(Q)) (%)	Nashbis (%)	Bias (%)
GR3J	Mean	66,5	52,8	31,0	87,7	74,1
	Maximum	75,8	66,8	56,5	93,4	94,3
	Minimum	60,0	31,8	-4,2	83,6	44,8
IHAC	Mean	57,2	42,2	16,2	84,6	62,7
	Maximum	60,8	62,6	50,4	89,3	89,6
	Minimum	54,0	14,8	-23,5	81,1	25,0

Table 17: Performances of GR3J and IHAC using the scaling method with adjusted WI10 on the Arade catchment in simulation mode

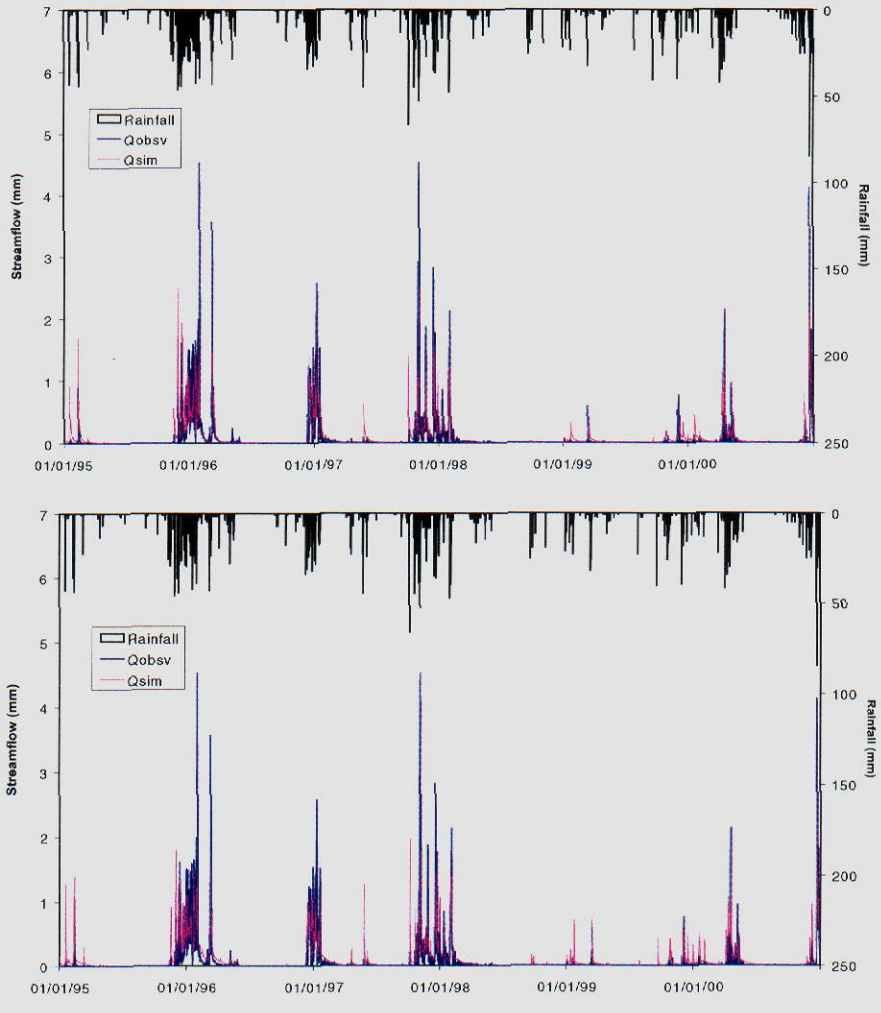


Figure 23: Stream flow simulation on the 1995-2000 period for (a) GR3J model (Nash=63.7 %) and (b) IHAC model (Nash=56.8 %) with the scaling method and WI10

In summary, the results indicate that the scaling method is effective for the Serein catchment, where simulation results were generally improved. For this catchment, GR3J performs better than IHAC with significant improvement being achieved for flood simulation.
In the Grand Morin and Orgeval, the benefit of the scaling method for flood simulations is obvious in the case of IHAC. The original configuration of the GR3J model, however, has difficulty simulating low flows satisfactorily for the Grand Morin and the problem becomes even worse when the scaling method is applied. Overall the IHAC results are better than those of GR3J on this catchment. But these results must be balanced with the observation problems of low flow measurements in the Grand Morin. In the Orgeval, the simulation of low flows is better using the original structures. It seems that use of the DiCaSM 10 cm top soil layer wetness index should be preferred to that of the 5 cm layer in the context of the scaling method in the Seine basin. This is because the 10 cm layer is more representative of overall catchment moisture and, therefore, better suited to the requirements of rainfall-runoff production functions.
The scaling method failed to bring significant improvement to streamflow simulations in the Arade catchment. This may be because soil moisture status has a limited influence on runoff generation in this region, unlike conditions in the Seine basin. Model results indicate the splitting coefficient of gross rainfall is poorly dependent on soil moisture status. Besides, this catchment has a rapid response time, probably less than one day, which makes a shorter time step more appropriate for the study of flood dynamics in this area. The use of a modified Nash criterion, which is better adapted for low-yielding catchments, indicates that reasonable model simulations can be achieved by both original models.

The scaling method was tested on the Serein, Grand Morin and Orgeval catchments in France and the Arade catchment in Portugal. The GR3J and IHAC rainfall-runoff models were applied. These two models have non-linear and linear transfer modules respectively. The time-series of wetness index generated by the DiCaSM model was used to calculate the effective rainfall needed as input for the two models.
1- The WI of 10 cm produced better results than for 5 cm
2- Serein: The use of WI 10 cm has improved the Nash – Sutcliffe values when applying the GR3J and IHAC models.
3- G.Morin & Orgeval: The use of WI 10 cm has improved the Nash – Sutcliffe values when using the IHAC model.
4- Arade: The use of WI 10 cm has improved the Nash – Sutcliffe values when applying the GR3J model. Although the results are encouraging, further investigation is required.

- *The Variational Method*

Results based on assimilating point measurements of soil moisture

A first issue was to identify within the model the store(s) whose content could be linked to observed soil moisture measurements in order to assimilate this new information into the model.
The TDR probe gives a soil moisture measurement at several depths ranging from 10 to 165 cm. Equation (5) was tested for several depths of soil moisture measurements in order to find the optimal depth. We have also tried to make a link between these measurements and modelled moisture values by using several model internal state variables (the superficial store content H_s, the global store content H_g, the routing store content R) separately or by combining them linearly (e.g. $H_s\&H_g$ or $R\&H_s\&H_g$). The same tests were carried out with the GR4j model with internal state variables S and R. Results of these tests are provided here for the GRhum model.
Results show the efficiency of the updating procedure to be dependent on the choice of the store content considered in equation (5) and on the value of parameter k. This is a common feature for all basins. For the Grand Morin basin, figure 25 shows that the global store content H_g provides the best results in the assimilation mode. We can also note that the controlled assimilation methodology (k=500) seems to be as useful as the gradual assimilation methodology (value of k ranging from 1 to 20).

Figure 25 illustrates the optimal soil moisture measurement depth given by the assimilation procedure is dependent on the particular basin: the assimilation methodology is improved by superficial measurements for the Arade, Orgeval, Grand Morin and Petit Morin basins whereas deep measurements improve the methodology for the Serein basin. For a given internal parameter H_c, it should also be noted that the assimilation methodology is more efficient when a good correlation coefficient in equation (3) is obtained for a given depth

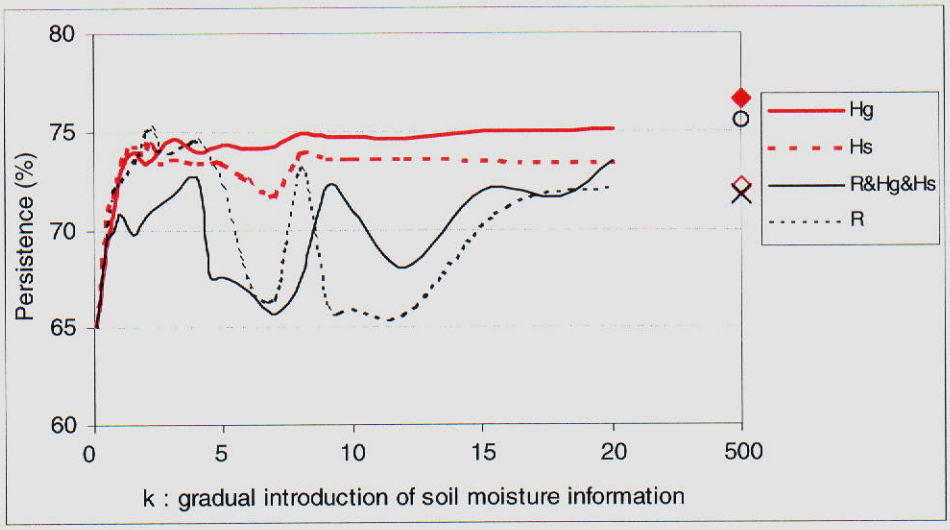
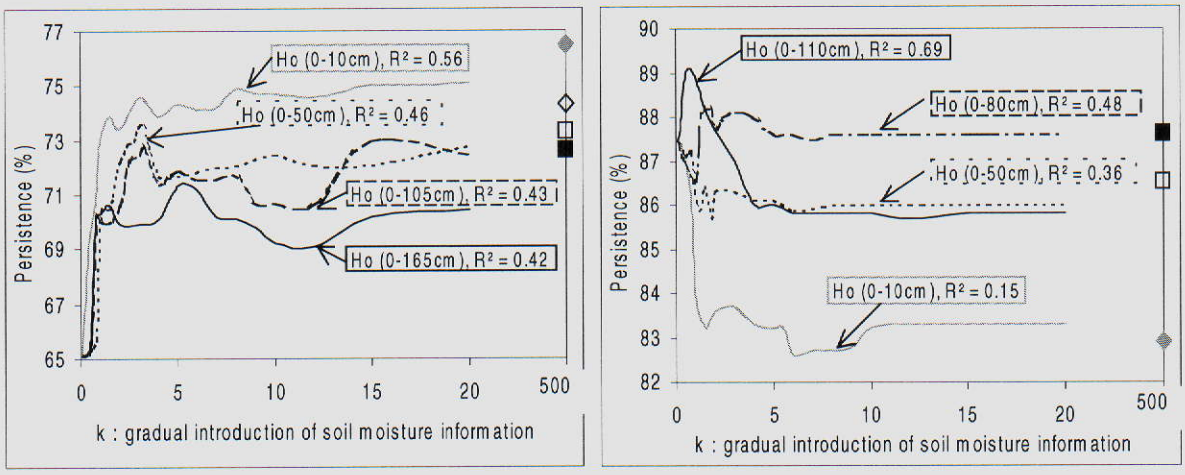


Figure 24: Variability of the persistence criterion with the store content *Hc*



(a) (b)

Figure 25: Relationship between determination coefficient of equation 3 and persistence criterion over the Grand Morin (a) and the Serein (b) at different depths with *Hc* = *Hg*

The main outcome of the different tests is that the assimilation of soil moisture data measured by TDR probes seems to be useful to increase the efficiency of the updating procedure for each basin (figure 26).

In the case of the Petit Morin and Grand Morin basins, the persistence criterion is increased by more than 10 % and the 10 cm depth measurement linked to both H_s and H_g store contents seems to be the optimal assimilation configuration. For the Serein basin, it is better to take the deep soil moisture measurements and the content R of the model routing store. In this basin, the assimilation of soil moisture data seems less efficient: the gain is only 3%. The original method (without soil moisture assimilation but only stream flow assimilation) already performs well for the Serein and these results are difficult to improve. It also appears that the results for the Arade basin are encouraging: the efficiency is improved by 5 % although the database is considerably reduced. In this basin the superficial depth measurement proved to be optimal and can be linked to the superficial store content H_s.

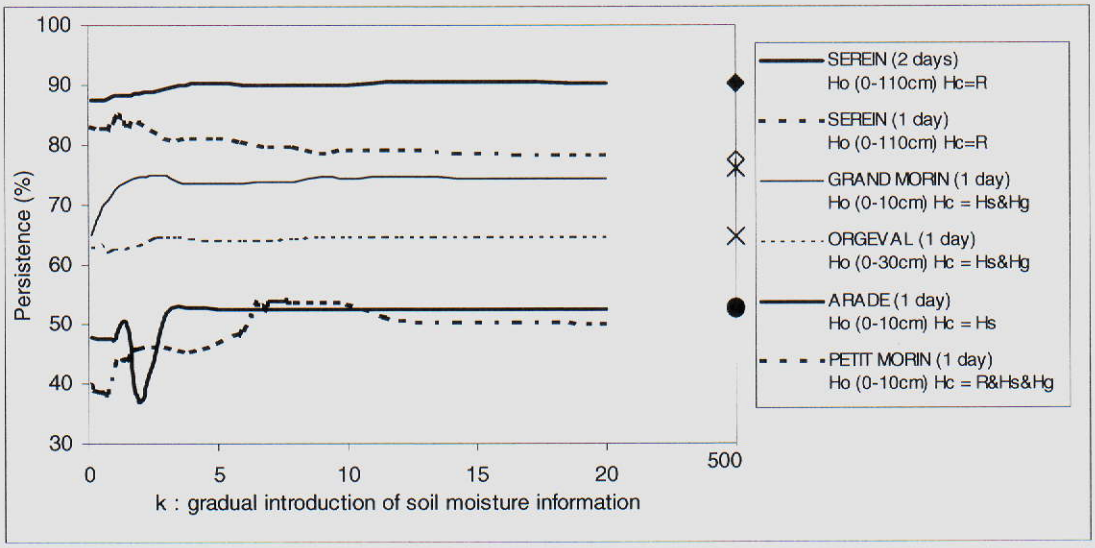


Figure 26: Persistence criterion with the best items taken into equation 3.

For both models GR4 and GRhum a comparison of the results during the test periods for flood events without assimilation and with assimilation of soil moisture and stream flow is presented in table 18.

		Without Assimilation		Assimilation of streamflow and TDR soil moisture		Comparison with assimilation of streamflow alone
Bassins	Models	Nash(Q) %	Persistence %	Efficiency %	Persistence %	Gain in persistence %
Serein L=2 days	GR4j	96.4	86.1	18.0	89.7	4.2
	GRHum	96.2	85.3	36.6	90.6	3.1
Orgeval L=1 day	GR4j	80.2	69.4	-18.3	65.8	5.3
	GRHum	80.2	69.4	-15.2	64.8	1.4
Grand Morin L=2 days	GR4j	88.0	84.8	2.4	85.1	3.8
	GRHum	87.0	82.4	7.9	83.1	10.9
Petit Morin L=2 days	GR4j	85.7	61.1	13.0	66.0	17.6
	GRHum	81.5	50.0	16.3	63.8	13.7
Arade L= 1 day	GR4j	36.4	41.6	14.8	66.5	1.5
	GRHum	9.2	42.5	18.8	53.3	5.6

Table 18: Comparison of results without assimilation and with assimilation of TDR soil moisture and streamflow

From these results it appears that the updating process is efficient except for the Orgeval catchment, where the delay response is less than the one-day time step of the model. Attempting to comply with observed soil moisture

values could provide significant improvement to the initial parameter updating methodology, assimilating stream flow alone (17% over the petit Morin). However, when the model is already efficient in the simulation mode the additional information on soil moisture seems less useful. The comparison of performances in assimilation mode for the GR4j and the GRhum model does not show any improvement when using a structure accounting for superficial and root soil moisture. The interesting feature is that the GRhum structure seems more able to integrate superficial soil moisture than GR4j model: the optimal soil moisture measurement depth to be considered by the assimilation procedure is improved by superficial measurements for the GRhum model and by deep layers for the GR4j model.

Results based on assimilating EO estimates of soil moisture

A first issue was to study the *a priori* required frequency of the EO soil moisture information in order to update the derivation of rainfall-runoff models in the forecasting mode. We have tested several frequencies of TDR measurements over the Petit Morin and the Grand Morin basins. For this purpose, we have removed successively 50 %, 80% and 100% of the available soil moisture data. As expected, figure 27 shows that the efficiency of the assimilation methodology decreases when the soil moisture database is reduced.

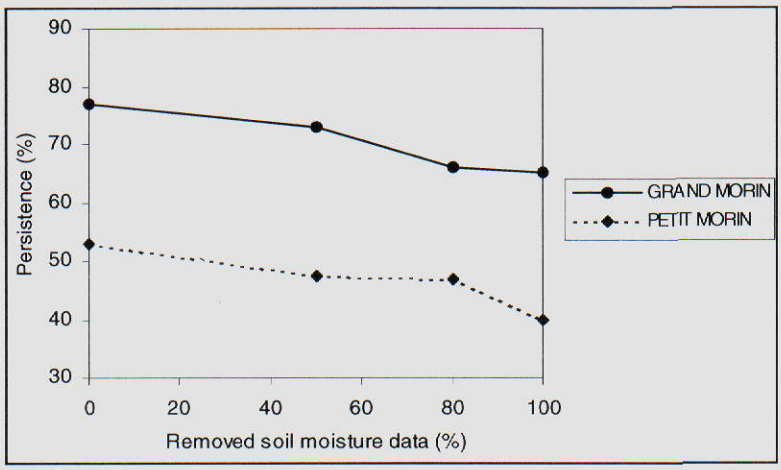


Figure 27: Persistence criterion with a decrease of TDR soil moisture data available

However, the two basins seem to behave differently. The efficiency for the Petit Morin basin drops quickly to 50% when 50% soil moisture data are lacking and then remains stable until 80% of moisture data is removed. For the Grand Morin basin the efficiency decreases slowly up to 50 % of data removal but then subsequently much more rapidly. The differences between the two basins come from the reduced database of soil moisture values: the efficiency is dependent on the concomitance between flood events and soil moisture measurements.

These results show that the assimilation of soil moisture data derived from EO techniques has little chance of increasing the efficiency of the parameter updating procedure when ther is more than a 40% loss of soil moisture measurements.

Currently, the number of the EO data available from the ERS/SAR satellite is about 20 per year, which represents almost 95% days without soil moisture estimation. This situation can be compared with the 80% loss tested above. Therefore, the assimilation of these EO data with the parameter updating method may not provide significant improvements in flood forecasting mode.

Results of EO data assimilation are shown in Table 19. For the Petit Morin basin and the Orgeval basin, the days with EO data do not correspond with the days of flood events. Therefore, the assimilation methodology cannot be carried out on these basins. For the Serein and the Grand Morin basins, the gain in persistence is improved by less than 2%, which is not significant enough.

For the Arade basin, the gain in persistence is improved by more than 5% for the GRhum model. The rate of flood days corresponding with the EO measurements days is higher than for the other basins.

		Assimilation of streamflow and EO data (soil moisture)		Comparison with assimilation of streamflow alone
Bassins	**Models**	**Persistence %**	**Efficiency %**	**Gain in persistence %**
Serein L=2 days	**GR4j**	87.9	13.4	1.7
	GRHum	88.5	21.9	0.9
Orgeval L=1 days	**GR4**	61.4	- 3.5	0.3
	GRHum	63.2	- 0.1	0.3
Grand Morin L=2 days	**GR4j**	81.8	- 9.2	0.3
	GRHum	71.2	- 3.1	0
Arade L= 1 day	**GR4j**	41.7	- 4.2	1.6
	GRHum	53.3	18.8	5.6

Table 19: Assimilation of EO data and comparison with assimilation of streamflow alone

Results showed that the variable parameter updating methodology could be used in a flood-forecasting context to assimilate catchment a soil moisture index into hydrological models. This method exploits information on both streamflow and soil moisture. Rainfall–runoff model performances for flood forecasting are significantly improved when assimilating *in situ* measurements of soil moisture at a daily time-step. However no real significant improvements could be achieved using soil moisture derived from ERS/SAR satellite where the image acquisition frequency is 35 days.
Due to the paucity of available EO data and the low frequency of soil moisture estimation, a definite conclusion cannot be reached until more data are retrieved from the catchment sites and validated over other catchments. However, the future SMOS mission with data acquisition of 3-day frequency should offer a new potential for the use of this information in hydrology.

- *The Sequential Method*

Results assimilating point measurements of soil moisture

As seen in the section above, after the determination of the observation and state vectors given by the structure of the model the first issue was then to determine the constraint relation between the observations and the internal variables on each basin.

Here the results are provided for the GR4j model and over the Serein sub-catchment. Figure 28 represents the soil moisture measured by the TDR probe at 30 cm against the level of the soil reservoir simulated by the GR4j model (without assimilation).

The figure exhibits a linear dependency allowing calculating the coefficients h_S and h_0 of the constraint equation. For the soil moisture measured at 30 cm: $h_S = 0.47$ and $h_0 = 0.15$ were obtained. Different depths have been tested and for the Serein the best relationship was obtained at 50 cm (R^2=0.92) with $h_S = 0.3$ and $h_0 = 0.27$. The other French sub-catchments also exhibited better relationships with soil measurements taken in the root zone. The Arade is the only one, which soil reservoir level S was linked to superficial soil moisture measurements.

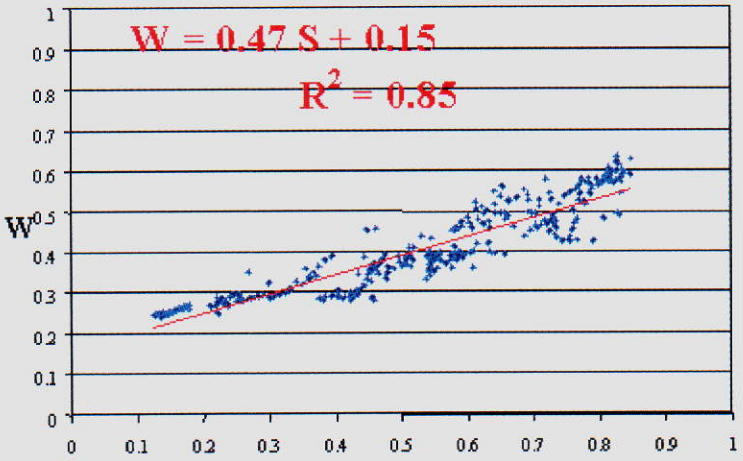


Figure 29: Relationship between the level of the soil reservoir and measured soil moisture over the Serein sub-catchment

The last stage focused on determining the covariance error matrixes, R_k for the uncertainties on the observations and Q_k for the uncertainties on the model. In order to check on the interest of the assimilation procedure, the uncertainties on the observation have been kept constant at the order of 5% (Quesney *et al*, 2000). The errors on soil moisture data and streamflow data are supposed to be uncorrelated. The exact value is not really important, as only the ratio R_k/Q_k, representing the uncertainties on the observations versus the uncertainties on the model, is significant. Different values of the uncertainties on the model have been selected in order to test the assimilation procedure. These values provide the level of confidence in the model. They vary and range from $Q_k=10^{-6}$ (no assimilation, the observation are not considered reliable and the Kalman gain is almost equal to 0) to $Q_k=10^{6}$ (forced mode, the model is not considered reliable and the Kalman gain is almost equal to 1).

The results of the tests with different values for Q_k and only with streamflow data assimilation, over the Serein, are provided in table 20. The forecasting time has been fixed to 2 days, which corresponds to the time response of the Serein catchment. The tests have been performed over the entire period of data. We can see in table 20 that the integration of flow data significantly improves the performances of the model, with the best result in forced mode: this means that the flow simulated by the model has to be entirely corrected to fit the data. This operation should be performed prior to the integration of soil moisture observations.

Q_k	10^{-6}	10^{-3}	1	10^{3}	10^{6}
Pers (%)	61.6	61.4	67.9	68.9	68.9
Eff (%)	0.0	0.1	15.7	19.6	19.6

Table 20: Persistence an Efficiency criteria for streamflow assimilation, over the entire test period

Q_k	10^{-6}	10^{-3}	10^{-2}	5.10^{-2}	10^{-2}	10^{6}
Pers (%)	68.9	70.1	72.8	72.1	71.8	66.5
Eff (%)	19.6	22.6	29.1	27.9	26.9	10.3

Table 21: Persistence and Efficiency criteria for soil moisture assimilation and streamflow forcing, over the entire test period

Table 21 shows the results of the progressive integration of soil moisture in the assimilation process on the Serein, keeping streamflow in forced mode. There is also an increase of both persistence and efficiency criteria, even if the increase is less important. This is not a surprise as the relation between soil moisture and streamflow is not direct; on the other hand, flow assimilation directly correct errors on streamflow.

Another difference is that the best result has been obtained with $Q_k=10^{-2}$, the forced mode $Q_k=10^{6}$ deteriorates the forecast (compared to simple flow forcing). This validates the use of a Kalman filter: simple forcing overcorrects the model. This can be explained by the fact that the link between soil moisture and the level of the soil reservoir (which is the adjusted corresponding internal variable) is not direct, and also by the fact that the observation comes from a punctual measurement, it does not perfectly reflect the mean soil moisture on the whole catchment. Same conclusions can be driven from the tests performed for the others catchments. They show also an increase of both persistence and efficiency criteria when assimilating soil moisture over the entire period.

During flood periods, tests with different values of model uncertainties Q_k have been performed (from 10^{-6} to 10^{6}). The best results have been obtained again with assimilation of soil moisture ($Q_k=10^{-2}$).
Table 22 shows the results for the GR4 model, GRHum providing similar results.
It appears from these results that the updating process is always efficient even over the Orgeval. The efficiency criterion reached 32 % over the Petit Morin.
The persistence criterion is higher with the sequential method than with the variational one except for the Serein. Again the model is already quite efficient in simulation mode (Nash > 90%) the additional information seems less useful but, compared with the assimilation of streamflow alone, using observed soil moisture did not provide the same improvements of the updating methodology. When compared with the variational method, the updating method being quite efficient with streamflow alone.
Also, when assimilating soil moisture, the gain in persistence is less important during flood periods than for the entire test period (table 21).

GR4j model	Without Assimilation		Assimilation of streamflow and TDR soil moisture		Comparison with assimilation of streamflow alone
Bassins	Nash(Q) %	Persistence %	Persistence %	Efficiency %	Gain in persistence %
Serein L=2 days	95.9	84.1	85.1	6.7	0.4
Orgeval L=1 day	80.6	70.0	72.7	9.0	2.4
Grand Morin	68.4	85.5	87.3	12.6	0.1
Petit Morin	84.9	60.4	73.5	31.9	0.3
Arade L= 1 day	62.3	78.0	80.9	13.1	2.9

Table 22. Comparison of results without assimilation and with assimilation of TDR soil moisture and streamflow periods

Results based on assimilating EO estimates of soil moisture

In the former section, daily soil moisture data measured by TDR probes were used. It thus seemed interesting to simulate a time frequency of several days by using only part of the TDR measurements in order to test the adequacy of the procedure with EO data. These tests have been performed on the Serein catchment and over the entire test period. Figure 29 shows the evolution of the efficiency criterion with the time frequency of the soil moisture observations. The efficiency decreases as the gaps in the data increase but remains significantly higher than without assimilation: the efficiency for the Serein basin drops to 50% when the acquisition frequency of soil moisture data is above 10 days. These results show that the assimilation of soil moisture data could increase the efficiency of the updating procedure when soil moisture measurements are missing less than 90 % . Even for a time acquisition frequency as high as 2 weeks the efficiency is higher than without assimilation. As the updating procedure is designed to be fulfilled only when observations are available, sequential assimilation seems well adapted to scarce or incomplete data sets. However, the number of the EO data available from the ERS/SAR satellite is about 20 per year, which represents almost 95% days without soil moisture estimation. Therefore the

assimilation of these EO data with the updating method may not provide significant improvements in flood forecasting mode.

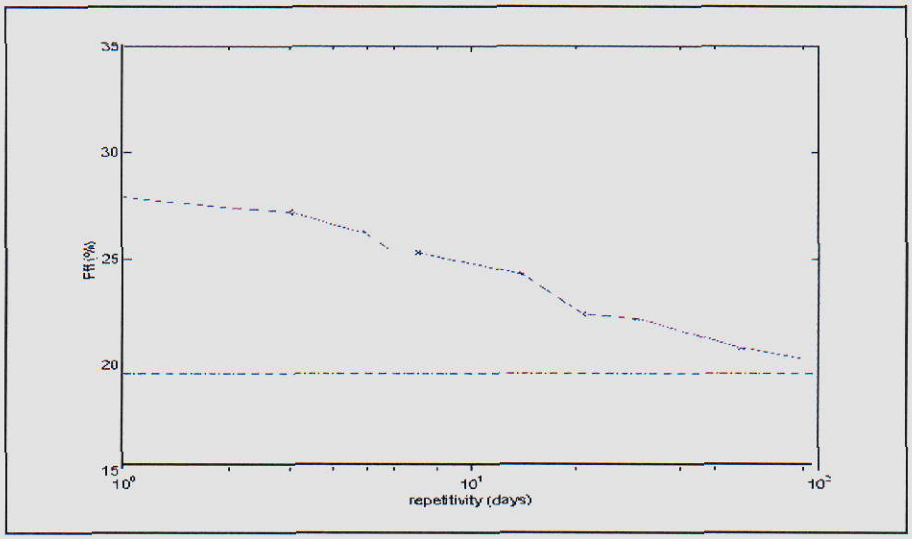


Figure 29: Efficiency criterion versus time frequency of the soil moisture measurements (the dotted line corresponds to the efficiency criterion without assimilation)

Another issue is to know how long the updating lasts: if the effect is longer than the delay between two observations, the assimilation procedure will still be efficient. To investigate this, Figure 30 plots the efficiency criterion over the Serein for forecast times from 1 to 25 days, this simulates the duration when assimilation is noticeable (here the forcing on streamflow data has been removed in order to isolate the effects of soil moisture assimilation, this explains that the criteria are lower than 10%). We can notice that the efficiency criterion remains almost constant until the forecast time equals 12 days: the correction brought to the level of the soil reservoir persists during more than one week. Currently the delay between two EO data acquisitions is about 3 weeks, which seems too long to be efficient.

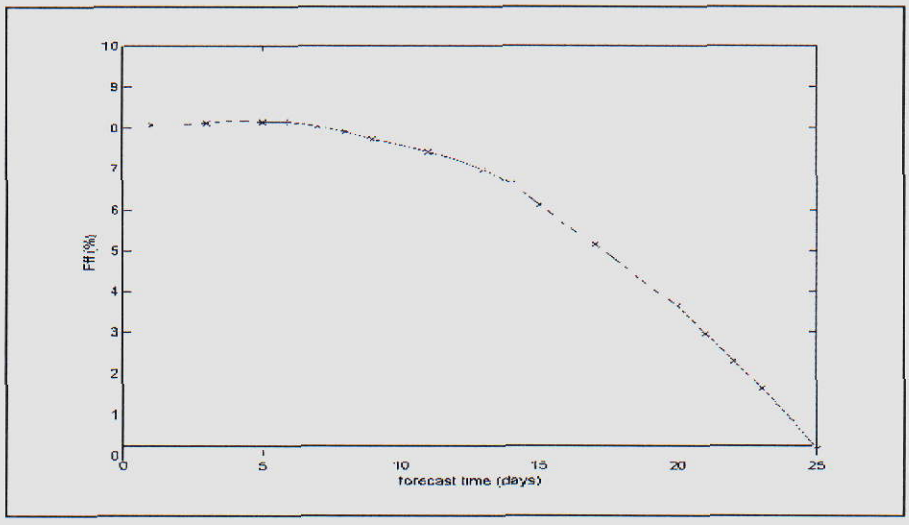


Figure 30: Evolution of the efficiency criterion with the forecast time

Results of EO data assimilation for flood periods are shown in Table 23 over the different sub-catchments. When assimilating soil moisture derived from EO data for the Serein, the Orgeval and the Grand Morin basins, the gain in persistence was improved by less than 0.5%, which was not significant enough.
For the Arade basin and the Petit Morin basin, the gain in persistence was improved by more than 5 %, for these basins flood simulations were not very accurate and the scarce soil moisture information seems interesting enough to update the internal states during flood periods.

GR4j Model	Assimilation of streamflow and EO data (soil moisture)		Comparison with assimilation of streamflow alone
Bassins	Persistence %	Efficiency %	Gain in persistence %
Serein L=2 days	85.1	6.7	0.4
Orgeval L=1 day	70.7	2.5	0.4
Grand Morin L=2 days	87.3	12.6	0.1
Petit Morin L=2 days	82.6	53.9	9.1
Arade L= 1 day	83.2	23.8	5.2

Table 23: Assimilation of EO data and comparison with assimilation of streamflow alone

The sequential updating procedure was first tested with daily *in situ* data on the entire test period. They granted a significant improvement of the forecasting performances of the hydrological model, which confirms that the sequential updating method can be used for streamflow forecasting. The procedure has been tested on flood periods. The efficiency over the different basins could reach 32% when assimilating streamflow and soil moisture but the gain against the assimilation of streamflow alone was less significant than over the entire test period. Compared with the variational method, the total gain against the results without assimilation is higher but compared with the assimilation of streamflow alone it is less significant. However, the tests showed that the updating of the model was efficient for more than one week, this may indicate that the sequential procedure is better adapted to scarce data, like those derived from Earth Observation data in future satellite missions, than the variational one. Currently using ERS/SAR data, results showed little improvements in the updating methodologies due to the paucity of available EO data but should be more efficient for higher data acquisition frequency.

4.2.6 Comparative Analysis and Operational Suitability

The results obtained using the soil moisture data in the three methods have been further analyzed and consolidated. The scaling method summary results are given in Table 24. The results show that the runoff prediction of some catchments has improved by assimilation of soil moisture in the GR3J and IHAC models. Table 25 represents the results as either a " +" sign, which means there has been improvement in the runoff prediction, or a "-" sign to indicate the opposite. The runoff prediction using the IHAC model seems to have been improved when using the assimilated soil moisture in three Seine catchments with slight improvement for the Arade. On the other hand, the GR3J prediction was improved when using the assimilation soil moisture on Serein and Arade with less impact on Orgeval and a negative impact on the Grand Morin. The number of "+" signs are dominating which indicates that the assimilation of soil moisture had improved the R-R model runoff predictions.
For possible operational use of GR3J and IHAC rainfall-runoff models, one should establish a set of consistent values of parameters A&B for each catchment and establish a WI for wet, average and dry years for the same catchment. Subsequently, for a particular catchment and year (dry, wet, average) one could use the specific values of A&B parameters of that catchment to estimate the expected runoff volume for a given event(s).

The scaling effect on runoff prediction of four catchments.

Model	Serein		Grand Morin		Orgeval		Arade	
	Nash-Sutcliffe		Nash-Sutcliffe		Nash-Sutcliffe		Nash-Sutcliffe	
	Original	Scaled	Original	Scaled	Original	Scaled	Original	Scaled
GR3G	84	88.5	78.5	72.4	87.8	86.7	57.1	66.5
IHAC	58.9	77.8	77.1	86	71.9	85.2	58.6	57.2

Table 24. Results of the models simulation without (original) and with inclusion of the soil moisture data using the Scaling methods.

The scaling effect on runoff prediction of four Catchments.

Model	Serein	Grand Morin	Orgeval	Arade
IHAC	+	+	+	±
GR3G	+	-	±	+

Table 25. The effect of soil moisture assimilation on runoff prediction using the scaling method

"+ " means positive effect, " – " mens egative effect and "± " means slight change.

In using point measurements of soil moisture (TDR), the variational method results indicated that the updating process is efficient except over the Orgeval catchment, as its time lag is less than one-day. Incorporating observed soil moisture could provide significant improvement of the initial parameter updating methodology assimilating stream flow alone by as much as 17% over the Petit Morin.

In using the EO, the variational method does not perform better than the point measurement case due to the reduced size of soil moisture data (35 days acquisition frequency). An additional problem is that in the Petit Morin and Orgeval basins, the days with EO and flood data did not correspond. Therefore, the assimilation methodology could not be carried out on these basins. For the Serein and the Grand Morin basins, the gain in persistence is improved by less than 2%, which is not significant enough.
For the Arade basin, the gain in persistence is improved by more than 5% for the GRhum model. The number of flood days corresponding with the EO measurements days was higher than for the other basins. The possibility of operational use of the variational method depends heavily on a good size of soil moisture data sets. This is easily done with respect to point measurements in the field but rather difficult to achieve if EO data is to be used. Subsequently, the possibility of operational use of the EO data in the variational method would require more

frequent satellite images. This could be possible under the future SMOS mission with data acquisition of 3-day frequency. The latter should offer a new potential for the use of this information in operational way.

Compared with the variational method, in the sequential method, the total gain against the results without assimilation is higher but compared with the assimilation of streamflow alone it is less significant. However, the tests showed that the updating of the model was efficient for more than one week, this may indicate that the sequential procedure is better adapted to scarce data, like those derived from Earth Observation data in future satellite missions, than the variational one. Currently using ERS/SAR data, results showed little improvements in the updating methodologies due to the paucity of available EO data but should be more efficient for future higher data acquisition frequency. Table 26 summarizes the results obtained by the variational and sequential method.

GRj model during flood periods		Without Assimilation		Assimilation of streamflow and TDR soil moisture	
Bassins	**Updating methods**	**Nash(Q) %**	**Persistence %**	**Persistence %**	**Efficiency %**
Serein L=2 days	variational	96.4	86.1	89.7	18.0
	sequential	95.9	84.1	85.1	6.7
Orgeval L=1 day	variational	80.2	69.4	65.8	-18.3
	sequential	80.6	70.0	72.7	9.0
Grand Morin L=2 days	variational	88.0	84.8	85.1	2.4
	sequential	68.4	85.5	87.3	12.6
Petit Morin L=2 days	variational	85.7	61.1	66.0	13.0
	sequential	84.9	60.4	73.5	31.9
Arade L= 1 day	variational	36.4	41.6	66.5	14.8
	sequential	62.3	78.0	80.9	13.1

Table 26: Comparison of updating methods for the GRj model with assimilation of streamflows and TDR estimates of soil moisture

4.2.7 Conclusions and Perspectives

The results of the application of the scaling method to the Serein, Grand Morin, Orgeval and Arade catchments were good. Generally, the results show that the runoff prediction has been improved by assimilation of soil moisture in the GR3J and IHAC models. However, the degree of success varied between models and catchments.

Future use of this approach (only if GR3J and IHAC are used) would require further work to establish a consistent set of parameters A&B specific for each catchment (eq.2). These parameters will help to transform the catchment Wetness Index, WI, obtained from the distributed model into an index that is an input for the GR3J and IHAC rainfall-runoff models. For possible operational use of GR3J and IHAC rainfall-runoff models, one should establish a set of consistent values of parameters A&B for each catchment and establish WIs for wet, average and dry years for the same catchment. Subsequently, for a particular catchment and particular year (dry, wet, average) one could use the specific values of A&B parameters of that catchment to estimate the expected runoff volume for given event(s). It is also possible that some other rainfall-runoff models will not need A&B parameters and can

make direct use of the WI obtained from distributed models for wet, dry and average years. In such case, one could select a WI curve near enough to the hydrological situation (i.e. dry, wet, etc.) and estimate the runoff for a given number of events. In that context, an attempt has been made to use the WI of DiCaSM directly into a rainfall-runoff HYDROMED model (Ragab *et. al.* in press) using the Serein catchment. The results were very encouraging. That would open up the possibility of using the WI directly in other rainfall-runoff models.

In using point measurements of soil moisture (TDR), the variational method results indicated that the updating process is efficient in most cases; the gain in persistence is improved by 1.5 to 17% for the studied catchments. Incorporating observed soil moisture did provide significant improvement of the initial parameter updating methodology. In using the EO, the variational method did not perform better than the point measurements case due to the reduced size of soil moisture data. Moreover, there was an additional problem related to the fact that in some cases the days with EO data did not correspond to the days of flood events. The assimilation methodology, therefore, could not be carried out on such occasions. Generally with EO data, the gain in persistence is improved by 2 to 5% for the studied catchments. This could have been improved further if the number of flood days corresponding with the EO measurements days was high enough to make a good size data set. These findings showed that the variable parameter updating methodology could be used in a flood-forecasting context to assimilate catchment *in situ* soil moisture index into hydrological models.

The sequential updating method can also be used for flood-forecasting. The procedure has been tested on flood periods. The efficiency over the different basins could reach up to 32% when assimilating streamflow and soil moisture but the gain against the assimilation of streamflow alone was less significant than over the entire test period. Compared with the variational method, the total gain against the results without assimilation is higher but compared with the assimilation of streamflow alone it is less significant. However, the tests showed that the updating of the model was efficient for more than one week. This may indicate that the sequential procedure is better adapted to scarce data, like those derived from Earth Observation data in future satellite missions, than the variational one. Currently using ERS/SAR data, results showed little improvements in the updating methodologies due to the scarcity of available EO data but should be more efficient for higher data acquisition frequency.
Due to the scarcity of available EO data and the low frequency of the soil moisture estimation, better results are expected if more frequent data become available. However, the future SMOS mission with data acquisition of 3-day frequency should offer a new potential for the use of this information in hydrology.

The results obtained so far are very encouraging and worth further investigation. It the first or one of the very few attempts by hydrologists to try to use remote sensing to predict the stream flows. As such, it is a first step towards a future operational system that can directly employ more frequent Remote sensing data to predict stream flows. The latter can help reservoir managers to better allocate water resources among users depending on the supply, demand and the list of priority.

References

Aubert D., Loumagne C., Weisse A., Le Hégarat-Mascle, (2001), Assimilation of Earth Observation data into hydrological models : the sequential method. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, accepted.

Baudez J.C., Loumagne C, Michel C, Palagos B, Gomendy V et Bartoli F, (1999). Modélisation hydrologique et hétérogénéité spatiale des bassins : vers une comparaison de l'approche globale et de l'approche distribuée.. *Etude et Gestion des Sols* , **6**, 165-184.

Beven, K.J. (1997). TOPMODEL : a critique. Hydrological Processes, **11**(9), 1069-1085.

Edijatno and Michel, C. (1989). Un modèle pluie-débit journalier à trois paramètres (A three parameter daily rainfall-runoff model). *La Houille Blanche,* 2, 113-121.

Edijatno, Nascimento, N.O., Yang, X., Makhlouf, Z. and Michel, C. (1999). GR3J: a daily watershed model with three free parameters. Hydrological Sciences Journal, **44**(2), 263-278.

Engman E.T. (1990). Use of microwave remotely sensed soil moisture in hydrologic modeling. In: *Application of Remote Sensing in Hydrology*, ed. G.W. Kite & A. Wankiewicz, 279-292. Proc. Symp. No.5, NHRI, Saskatoon, Canada.

Jakeman, A.J., Littlewood, I.G. and Whitehead, P.G. (1990). Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. Journal of Hydrology, **117**, 275-300.

Kalman R.E. (1960). A new approach to linear filtering and prediction problems. *J.Basic Eng.,* **vol 82 D**, 35-45

Klemes, V. (1986). Operational testing of hydrologic simulation models. Hydrological Sciences Journal, **31**(1), 13-24.

Le Hégarat-Mascle, S., Alem, F., Quesney, A., Normand, M., Loumagne, C., (2000). Estimation of watershed soil moisture index from ERS/SAR data, *Proceedings of EUSAR2000, in Munich, Germany, on May 23-25, 2000,* pp.679-682.

Lettenmaier, D. P., Wood, E. F. (1992) Hydrologic forecasting. Chapter 26 in Handbook of hydrology, ed. by D.R. Maidment , McGrawHill.

Loumagne C., Michel C., Normand M. (1991). Etat hydrique du sol et prévision des débits. *Journal of Hydrology*, **vol. 123, n°1-2**, 1-17.

Loumagne, C., Chkir, N., Normand, M., Ottlé, C., Vidal-Madjar, D. (1996). Introduction of soil/vegetation/atmosphere continuum in a conceptual rainfall/runoff model. Hydrological Sciences Journal, **41**(6), 889-902.

Littlewood, I.G. and Jakeman, A.J. (1994). A new method of rainfall-runoff modelling and its applications in catchment hydrology. *Environmental Modelling, Vol. II, P. Zannetti (Ed.), Computational Mechanics Publications, Southampton, UK,* 143-171.

Littlewood, I.G., Down, K., Parker, J.R. and Post, D.A. (1997). The PC version of IHACRES for catchment-scale rainfall-streamflow modelling. Version 1.0. User Guide. Institute of Hydrology, 89 p.

Makhlouf, Z. (1994). Compléments sur le modèle Pluie-débit GR4J et essai d'estimation de ses paramètres. Thèse de doctorat, Université paris XI Orsay, 426 p.

Michel (1989). Hydrologie appliquée aux petits bassins ruraux (Applied hydrology for small rural catchments). Cemagref (Antony), 528 p.

Morel-Seytoux, H. (1998). Présentation du modèle HMS. Note to the IIBRBS.

Nascimento, N.O. (1995). Appréciation à l'aide d'un modèle empirique des effets d'actions anthropiques sur la relation pluie-debit à l'échelle du bassin versant (Assessment with an empirical model of the impacts of human activities on catchment-scale rainfall-runoff processes). *Unpublished PhD dissertation. CERGRENE / ENPC,* 550 p.

Nash, J.E. and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models. Part I: A discussion of Principles. Journal of hydrology, **27**(3), 282-290.

O'Connell P.E., Clarke RT. (1981). Adaptive hydrological forecasting a review. *Hydrol.Sci. Bull,* **26(2)**, 179-205.

Oudin L., Weisse A., Loumagne C., Le Hégarat-Mascle S., (2001). Assimilation of soil moisture into hydrological models for flood forecasting: a variational approach. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, accepted.

Perrin, C. (2000). Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. *PhD Thesis, INPG (Grenoble) / Cemagref (Antony)*, 530 p.

Perrin C., Michel C., Andréassian V., (2001) : Does a large number of parameters enhance model performance ? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology,* 242, 275-301

Quesney A., Le Hégarat-Mascle S., Taconet O., Vidal-Madjar D., Wigneron J.P., Normand M., Loumagne C. (2000). Estimation of watershed soil moisture Index from ERS/SAR data. *Remote Sensing of Environment,* **vol 72**, 290-303.

Quesney, A., François, C., Ottlé, C., LeHégarat-Mascle, S., C. Loumagne and Normand, M., (2001). Sequential assimilation of SAR/ERS data in a surface hydric model coupled to a global hydrological model with an Extended Kalman Filter,, Actes du Huitième *Symposium International "Mesures Physiques et Signatures en Télédétection", Aussois, France*, pp. 689-693.

Refsgaard J.C. (1997). Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrol.,* **28**, 65-84.

Ragab, R., Cain, J. D. (1997). Aggregation of input parameters for catchment scale model and its implication to model validation. Monitoring and modelling of soil moisture: Integration over time and space. Workshop program W3. 5th Scientific Assembly of the International association of Hydrological Sciences (IAHS), Rabat, Morocco, 23- April - 3 May 1997. (Abstract of a paper presented at the assembly).

Ragab R., Perrin C., Littlewood I, Bromley J., France M., (2001). Prediction of runoff to surface reservoirs using the remotely sensed catchment wetness index. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, accepted.

Ragab, R., B. Austin and D. Moidinis. 2002. the Hydromed model and its application to semi-arid Mediterranean catchments with hill reservoirs. 1- the rainfall-runoff model using a genetic algorithm for optimisation. Journal of Hydrology and Earth System Sciences (in Press).

Thomas, H.A. (1981). Improved methods for national water assessment, Report to U.S. Water Resources Council, Contr. WR15249270, U.S. Water Resources Council, Washington, D.C.

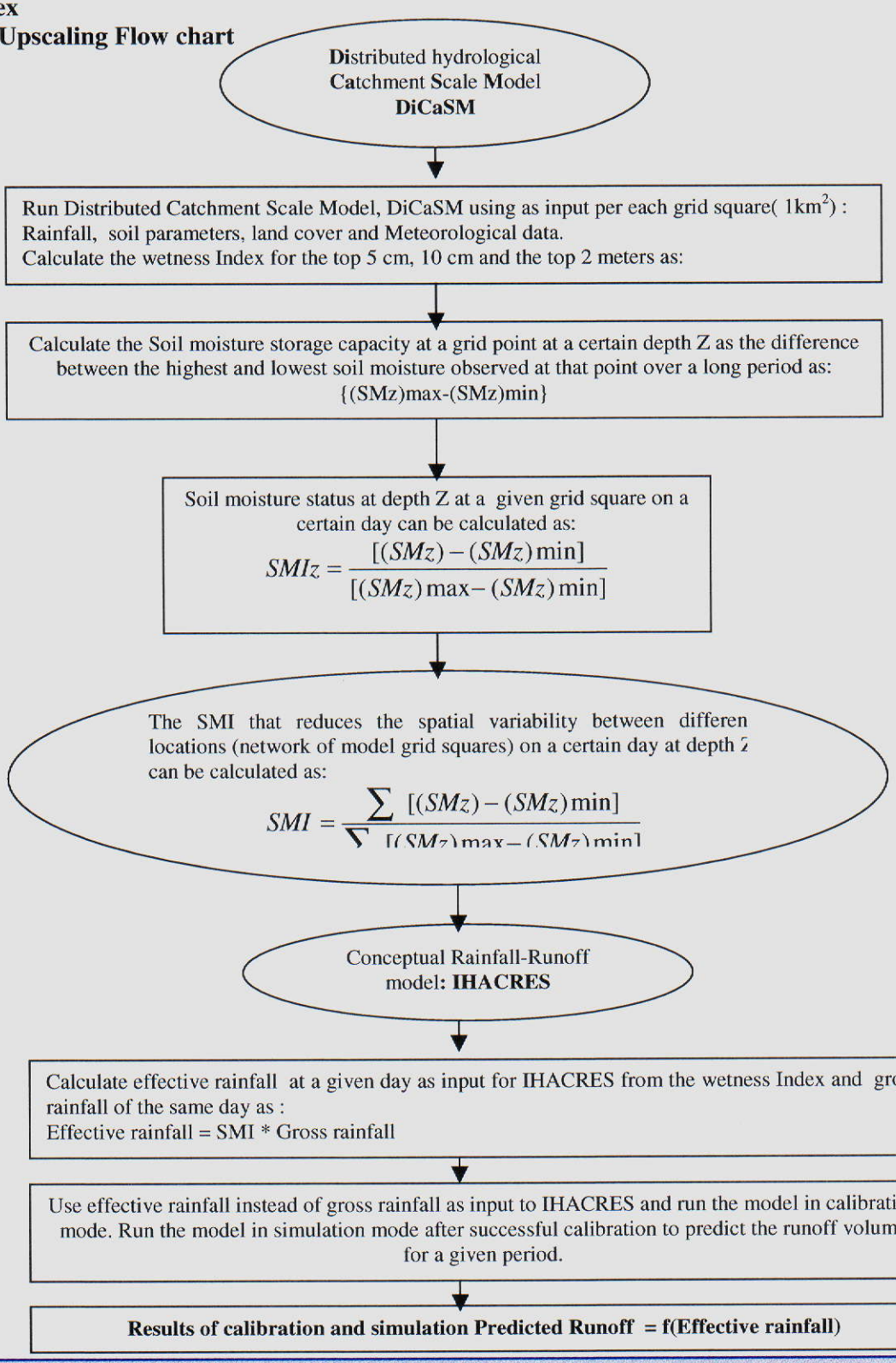
Wallis, J. R. and Todini, E. (1975) Comment upon the residual mass curve coefficient. *J. Hydrol.*, 24, 201-205.

Weisse A., Michel C., Aubert D. and Loumagne C. (2001). *Assimilation of soil moisture in a hydrological model for flood forecasting.* Soil-Vegetation-Atmosphere Transfer Schemes and Large-Scale Hydrological Models. Proc. of the 6th IAHS Scientific Assembly at Maastricht, The Netherlands, July 2001, *IAHS Publ.* no.°270, 249-256.

Yang, X. and Michel, C. (2000) Flood forecasting with a watershed model: a new method of parameter updating. Hydrol. Sci. J. 45(4), 1-10.

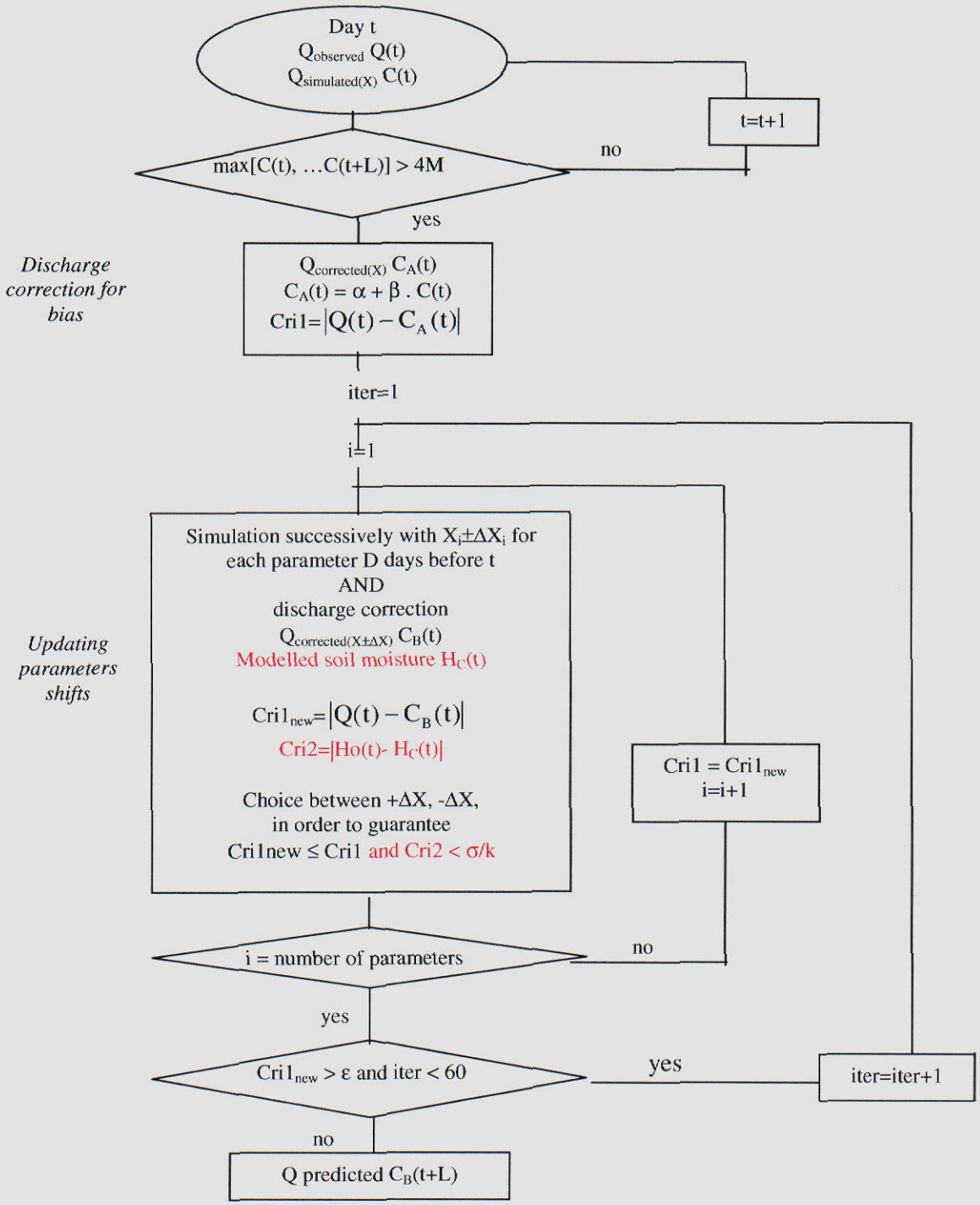
Annex

A.1. Upscaling Flow chart

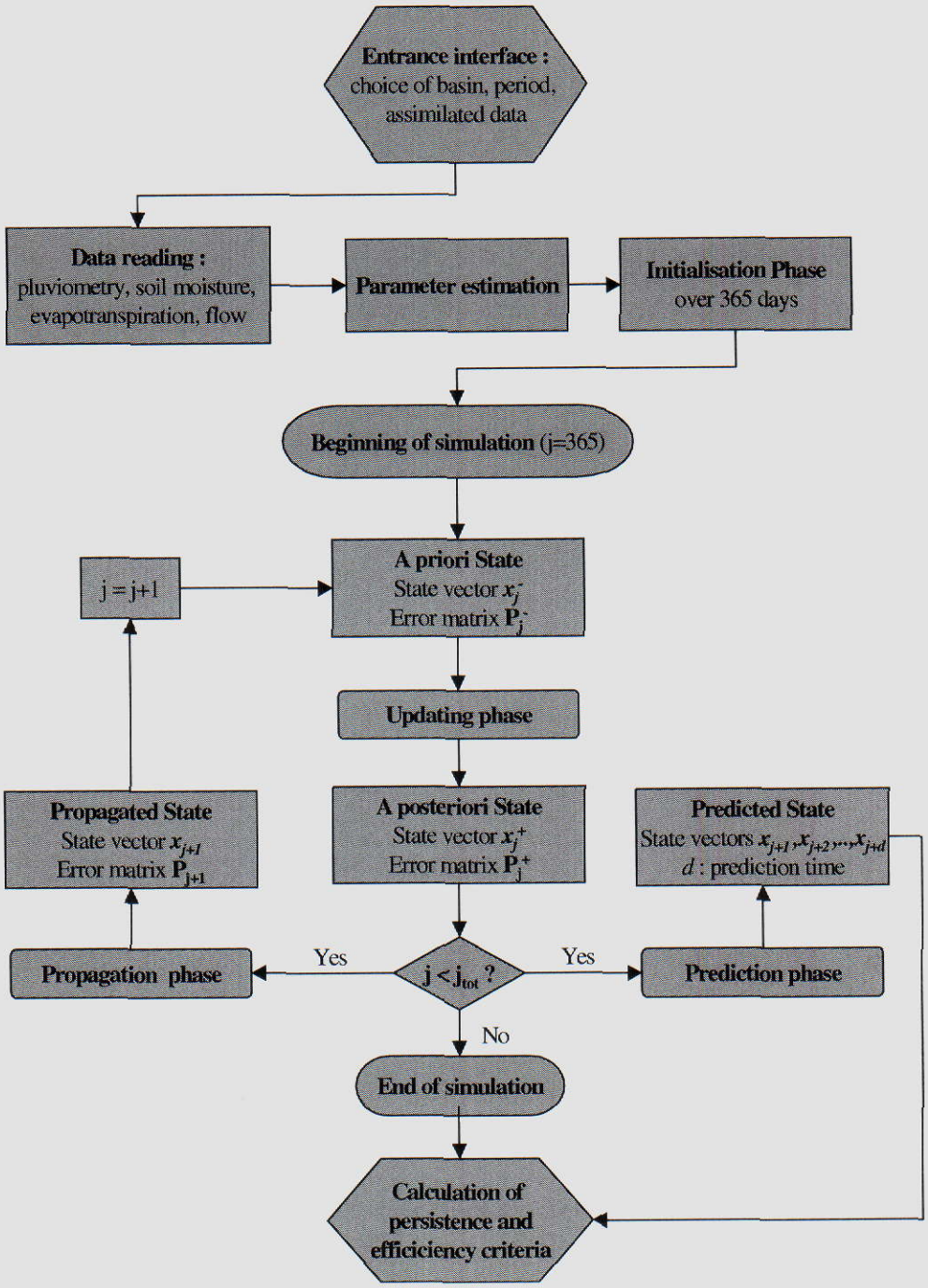


A.2. Variation Flow chart

In the following flow chart, the sequences in red are underlining the actions specific to the assimilation of soil moisture.



A.3. Sequential Flow chart



5. IMPLEMENTATION IN AN OPERATIONAL CONTEXT

Regarding user's needs and the adequacy of the assimilation procedures set up during the project some suggestions have been provided on how these methodologies can be implemented in reservoir operation. The adequacy of assimilation procedures have been analysed from a user perspective considering not only the improvements but also the gaps or difficulties of possible implementation in an operational context. In this chapter a presentation of the most appropriate forecast procedures and reservoir operations are provided along with the analysis of the current reservoir management leading to different rules of reservoir operation.

5.1 Use of EO data for the improvement of Reservoir Operations

For a better management of their system of reservoirs, users would like to improve on all aspects of the process of reservoir releases. Thus, they want to improve their models and to use all readily available information to gain a better knowledge of future rains, make a better simulation of river flows, and determine an optimal set of reservoir releases.

Part of these needs can be addressed based on the assumption that the amount of runoff that will result from a given rainfall will depend greatly on the initial degree of saturation of the basin, or in other words, on the amount of water that is held in storage in the basin, predominantly in the soil. A model that calculates runoff given rainfall can use information about the state of the system to improve its results in forecasting mode. Thus, a better knowledge of the soil moisture state of the system will improve forecast of runoff and consequently should, in principle, improve operations of the reservoirs.

This is the main assumption underlying the AIMWATER project. So, in order to meet user needs the project has to address two different issues: i) the possibility to measure soil moisture over the entire watershed in an operational way, ii) the selection of appropriate models that can incorporate this new information in the rainfall runoff process to improve the prediction of runoff by deterministic simulation using observed rainfall. The first issue is presented in chapter 3 and in two papers (S. Le Hégarat-Mascle et al, 2001 and J. Moreno et al, 2001). The second issue is the subject of this document.

There are two tasks involved in selecting the most appropriate forecast models and releases from a reservoir or several reservoirs. The first one is: provide the current knowledge of the system and predict its evolution into the near future, say for the next week. The second is: given the knowledge of the state of the system, currently and into the future, determine the optimal pattern of releases for, say, the next week to meet specific objectives at various points of interest in the system. Our purpose in this report is to situate the role of the project AIMWATER in that scheme of things and to see where and how well its results fit into the broader goal.

- **What is the AIMWATER project about?**

The AIMWATER project is a scientific project geared to determine whether the additional knowledge of a basin moisture index obtained by remote sensing can improve the prediction of runoff. So far we have used the term "prediction" in a loose sense. One could have used the term "forecast" as well. In this report these words will be used with a specific meaning in mind. We shall refer to "prediction" when an estimate at a future time of a quantity, such as discharge in a river at a given point, is made based on a model of the hydrologic system while the actual value of the discharge is already known. Thus given the current knowledge of the system and its future evolution, using a river basin model, one can predict the flows at a given point for the next week. The values of discharges at this given point have been observed and they are known. The purpose of the prediction is to assess the accuracy of the model and it is quantified by comparison with historical records.

In that sense, the project AIMWATER used different models to predict runoff from a variety of basins to compare the accuracy and robustness of these models (Perrin *et al*, 2001). The project then ranked the models and was able to select a few that, overall, performed better and can integrate soil moisture information into their algorithmic structure (Loumagne *et al*, 2001). Models were tested on a sample of 429 catchments worldwide.

Considering the results, several model structures were recommended because of their consistent performance and reliability: GR models: GR4 model (Edijatno *et al.*, 1999) and GRHUM model (Loumagne *et al.*, 1996), IHACRES model (Littlewood *et al.*, 1997) and TOPMODEL (Beven, 1997). These models were then used to carry the other steps planned for the study.

Then the project looked at modifications of the selected models to incorporate the additional information about the state of the system provided by soil moisture. In the context of AIMWATER this step has been referred to as "assimilation". The original models did not necessarily predict soil moisture as part of their internal structure. Consequently the internal structure had to be modified and, following that modification, the next task is to define how the observed new information (soil moisture index) is to be used within the model that does calculate soil moisture.

Is the EO (Earth Observation) measurement to be fully trusted or only partially, being itself subject to errors? This question is discussed in chapter 3 and in other papers (S. Le Hégarat-Mascle et al, 2001 and J. Moreno et al, 2001). Usually, when additional information about a system is to be used to improve performance of a model, but is known to be somewhat in error, a "filtering" technique is used (Refsgaard, 1997). Fundamentally the model uses a weighted mean between the value calculated by the model and the observation, to make predictions into the future. Then the improvement in performance is assessed by comparing predictions by models with or without assimilation, with or without filtering. Not surprisingly the models with assimilation and with filtering performed better. The improvement in performance has been quantified and it is significant (Oudin et al, 2001, Aubert et al, 2001). AIMWATER has fulfilled its scientific goal.

• Optimal Rule of Operations: Theory

The next step for implementation is to develop a framework to integrate the EO data information and the newly developed simulation tools into a set of operational rules for the timing and magnitude of the releases to be made of all reservoirs. For that purpose we need to review available theory for the optimization of reservoir releases. This will be discussed first in the context of a deterministic future (Massé, 1946). Though, the applicability and use of this methodology is relevant to general reservoir operations, the discussion will focus here on the Seine river basin system reservoir upstream of Paris, managed by one of the project customer.

An optimal rule of operations for the releases from a single reservoir or from several reservoirs was derived rigorously previously (Morel-Seytoux, 1999b) within, of course, a certain set of assumptions. The derived optimal rule of operation, which for simplicity we shall discuss here only in the context of a single reservoir, states that the "memory-integrated future marginal value" of the release must remain constant in time. The meaning of this statement will be clearer by looking at its mathematical expression and defining the symbols that appear in it.

$$\int_{t}^{t+M} \frac{\partial F_B[\tau, q_B(\tau)]}{\partial q_B} k_B(\tau - t)\, d\tau = \text{constant for all times } t \qquad (1)$$

In this expression F_B is the objective function to be optimized at target point B, say Noisiel or Paris, located at a certain distance from the release point A, say the reservoir Marne (Figure 1).

This objective function could be the square of deviations between actual flow at B, q_B, at time t, from a desired target value, which changes with the seasons, and that objective should be minimized.

$\frac{\partial F_B[\tau, q_B(\tau)]}{\partial q_B} k_B(\tau - t)$ is the marginal value of the flow at B at time t, whereas on the other hand, $\frac{\partial F_B[\tau, q_B(\tau)]}{\partial q_B}$ is the marginal value of the release, and $k_B(t)$ is the instantaneous unit hydrograph for flow routing from the point of release to the target point B.

M is the memory of the system of propagation from point of release A to B, which in the case of the Marne reservoir to Paris is around 7 days. The discharge at B depends on the natural flow at point A, $q_A(t)$ augmented

by the release x (t) and also on the many natural contributions of the tributaries, $q_{NT}(t)$, occurring between the point of release A and the target point B.
It can be expressed as:

$$q_B(t) = \int_{t-M}^{t} [q_A(\tau) + x(\tau)] k_B(t-\tau) d\tau + q_{NT}(t) \qquad (2)$$

It is clear from Eq.(2) that the flow at point B at time t depends on releases and natural flows for the M previous days. Thus looking back at Eq.(1) one can see that the optimal criterion at a given time involves not only the flows and releases for that day but also their values M days later and M days earlier. The integral in Eq.(1) has a clear economic significance. It is the marginal value of a release at time t, and that value is the integral with respect to future times of the marginal values weighted by the kernel of the convolution for propagation. Very naturally a unit impulse of release at time t creates a variation for $q_B(t)$ at time τ equal to $k_B(\tau - t)$.

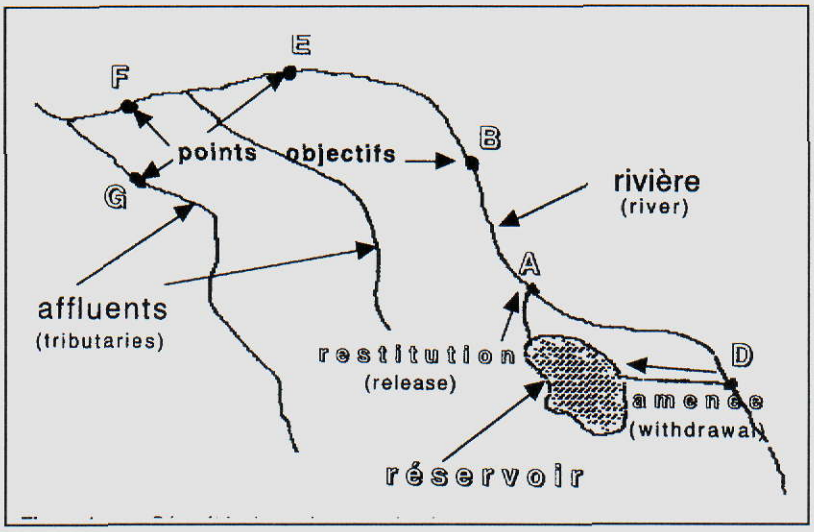


Figure 1: System geometry and important points

Put it plainly, today's release cannot be calculated optimally without calculating at the same time the releases for the next M days and without having forecasts of the natural flows for the next M days. A system of M equations must be solved for M releases. The coefficients depend on the characteristic of the system (geometry, e.g. where the tributaries feed into the main river, and river propagation characteristics such as Manning 's roughness, etc.). The right hand sides depend on forecasts of the natural flows for the next M days.

The unanswered question is: will the substitution of a forecast in the equations in place of the true value, as yet unknown, invalidate the optimality of the decision? In other words will the substitution of an estimate of the forecast in the decision equations still lead to optimal decisions "on the average"?

This is almost a moot point because the forecast being the only estimate available there is practically no alternative to its use. However one will not trust the decisions for the releases for days far from today. The saving grace for the optimal pattern is that even though releases for (M+1) days must be calculated today, only one is implemented and "tomorrow is another day", when calculations are repeated based on what happened today. Naturally all sorts of variations are possible and ultimately only simulations under "real time" conditions will confirm the validity of systematically using the forecasts in the decision equations to determine the releases.

Figure 2 illustrates the shape of a flow hydrograph at Austerlitz using a traditional operational procedure and the optimal one, however determined under a deterministic future. Obviously if the input forecasts are good a similar result would be obtained.

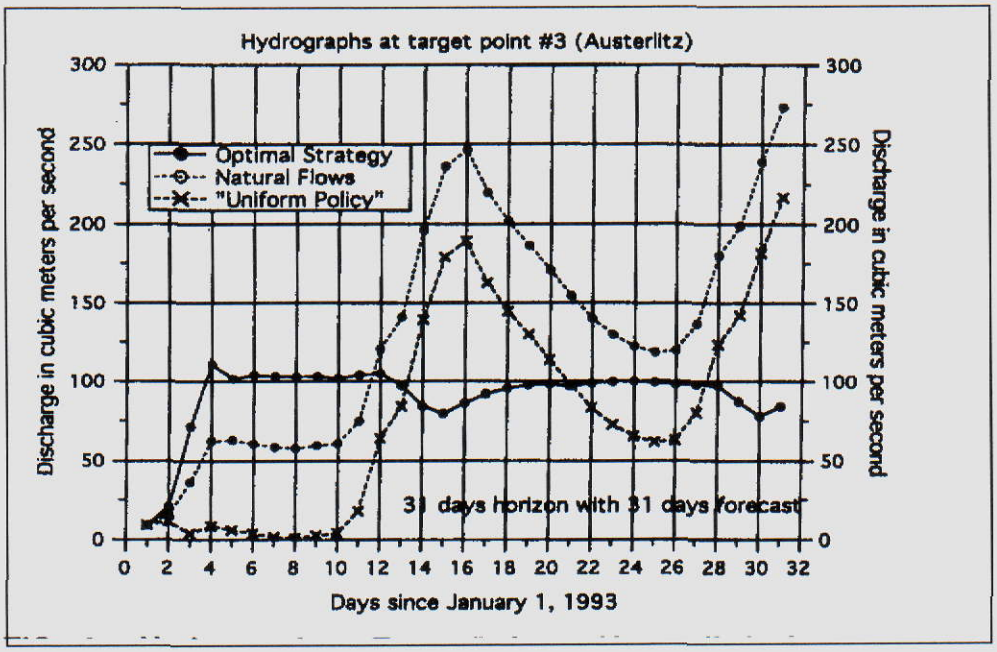


Figure 2: Hydrographs at Target point 3 (Austerlitz) with 31 day Forecast

- **Precipitation forecasting**

In real time one can use meteorological forecasts for an estimation of rainfall amounts. These forecasts are not very accurate on the long term and naturally are not available for historical records dating back to 1900. Short of meteorological forecasts, or in conjunction with them, one can define and use the statistical characteristics of rainfall records at all stations in the river basin. A previous statistical analysis for most stations in the Seine river basin upstream of Paris: Morel-Seytoux , 1998, produced the information necessary to forecast rain into the future, for tomorrow and on, as an expected value conditioned on the known situation of yesterday.

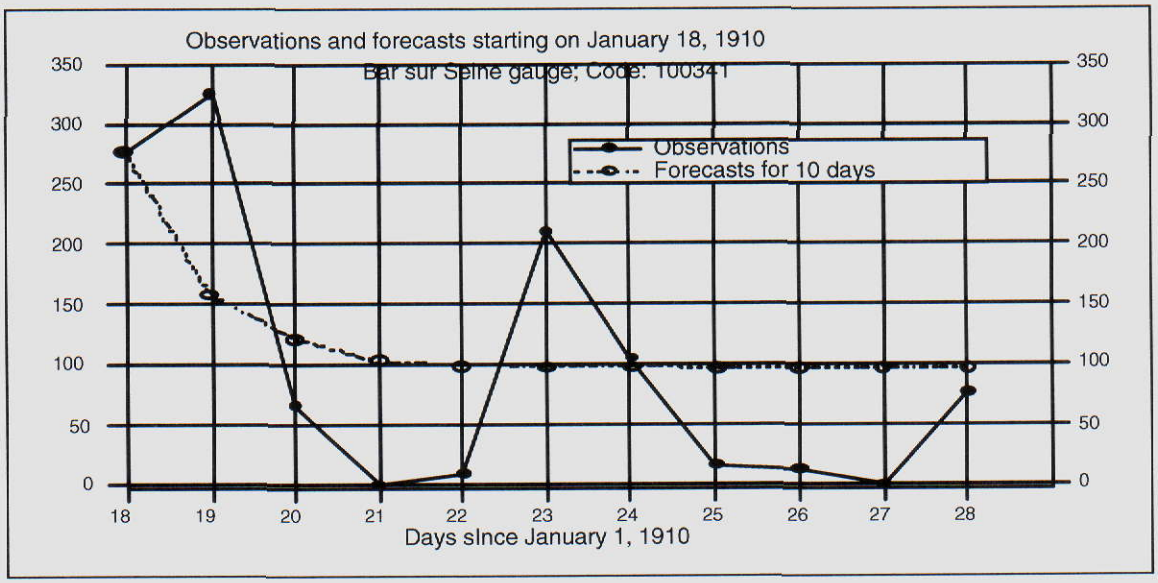


Figure 3: Rain Forecast at Paris for the 1910 Flood

Origin: H. Morel-Seytoux/Cemagref/ U. Independente
Distribution: Cemagref/CETP/CEH/U.Valencia/U. Independente/ARBSLP/IIBRBS/CEE

Figure 3 illustrates a forecast starting on January 18 of the year that saw the famous 1910 record flood in Paris. Clearly for greater lead times the influence of the initial condition on January 18 diminishes so that 10 days later the forecast tend toward the expected value of rainfall on that date.

• Conclusions

The AIMWATER project demonstrated that the determination of a soil moisture index via use of EO data was an operational feasibility. It also showed, on selected sub-basins in the Seine river basin and on the Arade basin, that an improvement in prediction of runoff could be expected. It remains to demonstrate that what appears to be a logical inference, namely that such prediction improvement would lead also to a forecast improvement in the system as a whole and to an improvement in reservoir operations, can be confirmed through a well controlled simulation experiment.

We have prepared the ground for the design of such an experiment by indicating the essential components of such an experiment. At present the two principal ingredients discussed here, namely: (1) a procedure to forecast rain and integrate it in a forecast of runoff and river discharge for several days, and (2) an optimal algorithm to select the releases to be made from reservoirs, are typically lacking in operational practice.

These will have to be integrated in the current reservoir release decision procedures, which are reviewed in the section below along with suggestions made for an improvement in the procedures.

5.2 Suggested steps toward the improvement of Reservoir Operations in actual practice

This section focus on a discussion of the steps that remain to be completed so that use of EO data can actually be integrated in an improved procedure for optimal reservoir operations.

• Background

The more information one has of a river basin system, the better one is able to understand its behaviour and, probably, manage it more properly. Thus the basic premise for this research project was that information about soil moisture basin conditions would provide additional information about the system that could be put to good use.

At the scale of interest of several thousand square kilometers in large basin operations, one is interested in an overall catchment wetness index rather than a sample soil moisture value collected over a few square meters small plot that is not necessarily (and usually) representative at a scale several orders of magnitude larger. Thus EO data provide information at the appropriate scale and would seem to be ideally suited for the purpose of hydrologic simulation of a river basin.

One of the first tasks of the project was to tackle the several problems encountered in converting a backscattered radar signal into a meaningful catchment wetness index. The specific procedure used for that purpose is described in chapter 3. In particular a recent article (Le Hégarat-Mascle et al, 2001) outline the 6 practical steps required to carry on this passage from acquisition of EO signal to the quantification of a catchment wetness index during a calibration phase and the 5 steps needed during the operational phase.

The next task was to find an "equivalence" relation between the EO data derived wetness index with the conceptual soil moisture storage representation existing in the models used by the operators to make forecasts of runoff from the watersheds that feed the rivers. Several studies were undertaken to (1) decide which rainfall-runoff models would be simple and robust enough to make adequate predictions of runoff given rainfall and (2) could accept readily without any, or with minimal, change the wetness index in the performance of these very predictions. This subject is discussed in chapter 4.

In the assimilation of data providing information about the internal state of a model, one runs into two major sources of errors. The soil moisture data are not free from errors and the rainfall-runoff model is only a rough approximation of the complex reality. Thus techniques had to be developed, and that was a third task, to "filter" these errors (or noises) out of the distorted signal in order to receive it clear. Chapter 4 displays results of various assimilation procedures that were tested on several basins for various lead times from assimilation of TDR estimates of soil moisture, and from assimilation of EO estimates of soil moisture. These results showed the improvement brought about by the assimilation of soil moisture for flood forecasting purposes.

The last task was to take all this information and new tools and to integrate them in a system of reservoir operating rules. This task it is now addressed in the specific context of how operations are currently carried out by IIBRBS one of the project customer. However this last implementation will require a great deal of effort on the part of IIBRBS. One can expect, from potential users, that they would need to be convinced of the merit of the proposed changes before they would be willing to incur the associated risk and expenses.

• The need for a convincing demonstration

Will an improvement in prediction of runoff in upstream or lateral basins result in improvement of operations of the reservoir in the basin as demonstrated, for example, by meeting the "target" discharge at a given point? This raises the next question: do actual operations, today, use the forecasted runoff effectively? Before we answer that question, we need to define "forecast" as opposed to prediction. In this report we refer to "forecast" when an estimate at a future time of a quantity, such as discharge in a river at a given point, is made based on a model of the hydrologic system while the actual value of the discharge (and of the rainfall) is still unknown.

For users involved in the AIMWATER project (IIBRBS (France) and ARBSLP (Portugal) the answer to the original question is No. Why is it so? Quite simply because the operators do not trust the forecasted runoff

because it depends on forecasted rainfalls and these are notoriously inaccurate even for a lead time of just one day, and forecasts are needed for up to seven days, the propagation time between the most upstream and downstream points in the Seine basin for example. Typically the rules of operations of reservoirs have been derived from simulations that assume a perfect knowledge of the future rainfalls or runoffs (deterministic future).

When the runoff simulation estimates are calculated by the model, the forecast rainfalls are the historical rainfalls. In other words could a rule of optimal operations, derived under the assumption that forecasts for the future are perfect, lead to better operations than a rule that would ignore these forecasts or use them in a less than supposedly optimal way? All AIMWATER results were done under the assumption of a deterministic future and so were all simulations that led to the determination of the "Target filling curve" or "Courbe Objective de Remplissage" (COR) for the Seine river basin. It is true that in the determination of the COR, some traces were generated stochastically but once generated they are treated as a deterministic future.

- **What demonstration?**

A convincing demonstration would combine: (1) a rule of operations that account for all the past and future flows that optimization theory says should appear in the determination of a release value under a deterministic future, (2) a way in which the future flows forecasts would appear as surrogate estimates for the true but unknown values, and (3) a forecast of inputs to calculate the estimates of runoff and river discharges by the river basin model.

A measure of performance is then obtained by comparison with historical records. Models of runoff and discharge propagation with or without assimilation of EO soil moisture data, with a standard rule of operations or with an optimal rule, can then be compared as to their merit in terms of improving reservoir operations. The optimal rules of operations and a statistical forecasting technique for rainfall have been discussed in more technical details in the above section and in the literature (Morel-Seytoux et al, 2001; Morel-Seytoux, 1998, 1999b). Now we need to confront these theoretical procedures with release rules currently in use.

- **Real time rules and legal constraints**

It is quite difficult to determine exactly how reservoirs operate on a day-to-day basis because it is almost always done on the basis of the long practical experience of the operators, naturally within a set of regulations, which do leave some room for flexibility. Nevertheless models are used more and more for decision support, and for simulations of various scenarios. From a study of the impact of these various scenarios on the behavior of the system, the operator then decides on a particular release to be made.

Currently, the simulated management procedure includes neither past releases nor future ones; experience in its use has shown that this procedure leads to oscillations in the pattern of the releases. The procedure then can be modified by an updating methodology to avoid these oscillations and by placing limitations on the range of changes in the values of the releases from one day to the next. Since theory suggests that past and future releases do influence the release to be made today, it can be sound to modify the current simulation procedure for operations with the most realistic theoretical approach developed in the context of a deterministic future and then extended to the case of an uncertain one. Unfortunately there is not a single unique objective way to include the uncertainty in the optimization formulation for the problem. The formulation is subjective depending on the attitude of the manager with respect to risk.

Many years ago a "filling and draining curve", COR, (Courbe Objectif de remplissage) was developed for each reservoir operated by IIBRBS. Specific reservoir volume contents defined in a legally binding document the "arrêté préfectoral" are to be met at monthly or bimonthly dates, such as January 1, 15 and February 1, etc. There is some flexibility between these dates as to the rate at which one will reattach to the COR if one had deviated from it. There are obvious reasons why one does not follow the COR exactly all the time. Sometimes there is simply not enough water in the river to divert to the reservoir under a drought condition. And even if there is water in the river there may be barely enough to meet minimum stream requirements downstream. Vice versa it is not advisable to hurry to reattach to the COR when a flood is coming and storage will be required to attenuate its impact. Within the binding legal rules there is some flexibility for operations but the decision depends much on the conjuncture, with all the uncertainties involved. How precisely the decisions on how much

to release at a given reservoir are made is not something that one can read in an IIBRBS report. One can find general guidelines or operating principles and one can read the "arrêté préfectoral". Eventually the best that can be offered from a reservoir manager tool like PEGASE, modified to assimilate the soil moisture information, is to serve in making the decision; it will not provide a decision per se.

This being said one can try to codify the rules of operations in a model to perform simulations that mimic real time operations. In that case it is a necessity to formulate the decision process in a rigorous algorithmic format. It does not represent how the system is really operated but it is a plausible rational and mathematical expression of how it is most likely to be operated.

• Simulation of reservoir operations in model PEGASE

The particular rule to be described here is called: the "extended rule of operations" (règle de gestion étendue). It is, as the name implies, a variation on a previous, somewhat less flexible, rule.

- *General principle for the filling and drainage of the Seine reservoirs*

First let us define a few terms. "Reference flood (downstream) discharge" (débit de référence en crue) refers to a maximum allowable discharge at a specified point downstream from the reservoir. If at all possible that value should not be exceeded. "Target filling curve" (Courbe-objectif de remplissage or COR for short) is a curve that defines the ideal scheduling of filling and drainage of the reservoir.

A current situation is labeled "flood" if both: (1) the uncontrolled downstream discharge would (or does in fact) exceed the reference value and (2) the volume of water in the reservoir exceeds that prescribed by the COR. During a "flood" period water is diverted into the reservoir while limiting the downstream discharge to the reference value in flood until the reservoir is full. Augmentation of summer low flows is only allowed if the three following conditions are obtained:

a. It is the summer period (summer refers to a condition of low flow, not the calendar season)

b. The summer reference discharge ("débit de référence secondaire d'étiage) is not exceeded downstream from the reservoir

c. The flood reference discharge is not exceeded.

- *Computation of the releases*

If the conditions for low flow augmentation are not satisfied then the situation is one for filling the reservoir. If in that situation, water is diverted to fill the reservoir but a minimum flow (débit réservé or reserved flow) must be allowed past the point of diversion and this procedure is followed until the reservoir content matches the value of the COR and beyond, while maintaining a downstream discharge at the reference level.

If the natural downstream discharge is less than the summer reference discharge a release is made of value calculated by a formula with variables are defined as follows:

x(j) is the release for day j; Qnatu(j) is the natural discharge just upstream of the release point; Q2(j-1) is the discharge observed at a (downstream) target point the day before day j ; Qtar is the target (desired) discharge at the target point; amp, epsi and Ti are parameters in the process of regulation and gam is a "filter" parameter.

The formula itself is:

$$x(j) + Qnatu(j) = x(j-1) + Qnatu(j-1) - amp.[Q2(j-1) - Q2(j-2)] + amp.epsi.[Q2(j-1) - Qtar)]/Ti - gam.[x(j-1) - x(j-2)] \quad (3)$$

The justification for this empirical formula is that the controlled total discharge (release + natural flow) at the point of release should be kept constant if the discharge at the target point does not change and has the desired target discharge value. On the other hand if the target value was not met on the previous day then the total discharge needs to be changed and somewhat in proportion to the deficit in meeting that target. The quickness to reestablish the desired target discharge is modulated by the parameter Ti. The term with the parameter gam is

meant to prevent too rapid changes in the releases. Naturally if the calculated value for the release is negative the release is zero.

Further discussion for this release rule, called Extended Operation Rule ("Règle de Gestion Étendue") is shown in the Appendix. We now concentrate on the limitations of the formula used.

- **Theoretical deterministic Optimal Rule of Operations**

We review briefly what was discussed already in the section above, applying here the procedure to the Seine river basin. The derived optimal rule of operation, for a single reservoir, states that the "memory-integrated future marginal value" of the release must remain constant in time. In algebraic terms this statement can be given the form:

$$x(j) + Q_{natu}(j) = \sum_{v=j-M}^{j+M} \{a(v)[x(v) + Q_{natu}(v)] + b(v)[Q_{lat}(v)\} \quad (4)$$

M is the memory of the system of propagation from point of release A to B, which in the case of the Marne reservoir to Paris is around 7 days (see Figure 4).

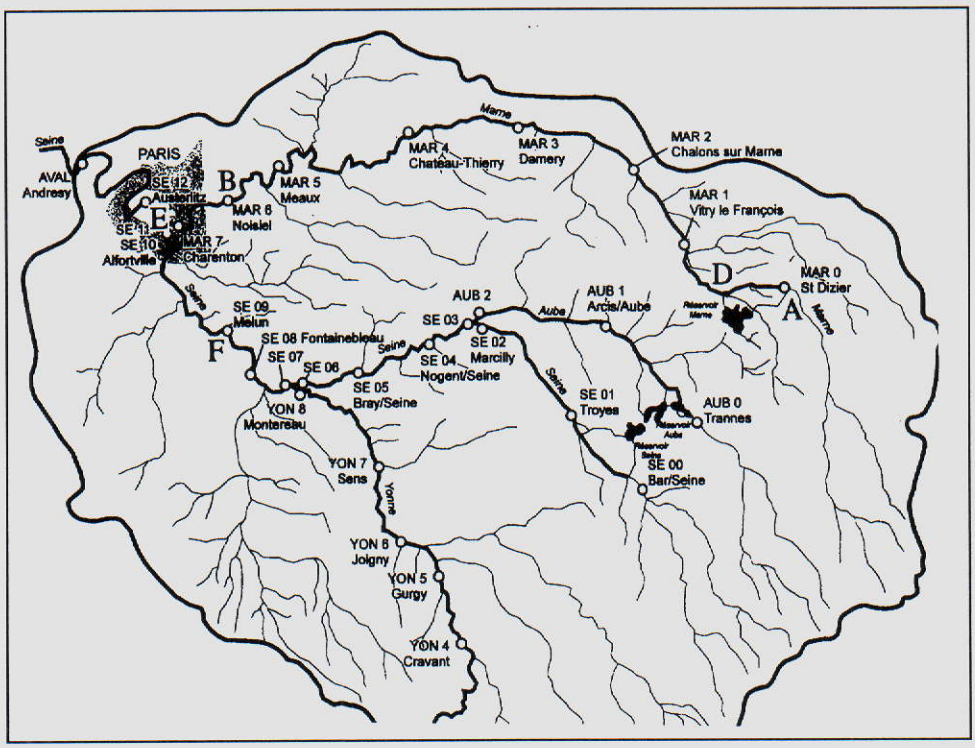


Figure 4: Geometry and Important Points and Features of System, for the Seine River Basin

Looking at Eq.(4) one can see that the optimal criterion at a given time involves not only the flows and releases for that day but also their values M days later and M days earlier, since the discharge at target point B depends upon the M antecedent releases and natural flows. It follows that today's release cannot be calculated optimally without calculating at the same time the releases for the next M days and without having forecasts of the natural flows for the next M days. A system of M equations must be solved for M releases. The coefficients

depend on the characteristic of the system (geometry, e.g. where the tributaries feed into the main river, and river propagation characteristics such as Manning's roughness, length, cross-section, etc.). The right hand sides depend on values of the natural flows for the next M days.

- **Comparison of the two rules**

As stated previously the coefficients a(n) and b(n) appearing in Eq.(4) depends on the geometric and hydraulic characteristics of the river system and in the case of the Seine river are fairly well known. They can be readily calculated as demonstrated before (Morel-Seytoux, 1999b). If one compares with Eq.(3) one should notice that Eq.(3) only involves the current release and no future ones (in the case of the Marne there should be about 6 future ones) and no future natural flows, upstream or lateral ones.

With that policy everything is dependent on what happened in the past. It is a policy that is based on the presumption that what happens tomorrow and the following days can be deduced from what happened in the previous couple of days by linear interpolation. One can be sure that this policy will fail whenever there is a change in magnitude of the trend and especially under the most important case of a reversal of trend. Note that the original formulation for the extended rule of operations, did not include the filtering terms and wide oscillations in release values occurred.

That did not increase the confidence of IIBRBS in the model and in modeling. Including these filtering terms in the rules of operation, oscillations are dampened through the last term in Eq.(3). Better rules of operations including the prediction coming from the model should be incorporated in PEGASE in order to help with the actual real time operations.

- **Suggested steps for the implementation in reservoir operations**

The implementation will require many steps. The IIBRBS would have to secure EO radar data for all sub basins in the Seine basin upstream of Paris. It would have to develop the regression relation between signal and observed field soil moisture, which implies the development of an extensive network of field measurements. Naturally there are shortcuts and these should be taken first, before engaging into expensive alternatives that may not be necessary. For example one could assume that the slope of the regression obtained on the Grand Morin, the Petit Morin and the Serein is representative and use an average value for all other catchments, thus bypassing the need for a full instrumentation of the basin. Similarly the ordinate at the origin, which is variable from catchment to catchment, could be estimated roughly on the basis of the studied three basins and eventually modified by checking on the sensitivity of that parameter on the runoff prediction. Next, one needs to develop a relationship between equivalent bare soil moisture deduced from the EO signal and the rainfall-runoff model internal storage variable. Again possibly an average relation could be developed based on the three catchments studied, and tested by simulation. Another problem for implementation is that the current GR model in PEGASE is an old version compared to the new versions needed for assimilation and recalibration will be necessary.

Having completed all these steps and presuming then that calibrated rainfall-runoff models are available for all catchments in the Seine river basin, it remains to integrate these models into a reservoir release scheme. Now of all of this could be used to predict runoff and use the existing so-called "extended rule". However as discussed in earlier sections this is fundamentally a very sub optimal rule. It should be replaced by a more optimal procedure, which incidentally could have been programmed in old PEGASE, independently of using EO signals. All in all the steps required involve a well defined procedure for acquisition of data and their treatment, and a great deal of programming work to process the information to the point that a set of releases for each reservoir is recommended every day on a real time basis.

To recognize more specifically the nature and extent of the information transfer required to proceed from AIMWATER scientific results to an implementation in the IIBRBS decision process for reservoir releases, let us imagine two scenarios. In one scenario a single scientist well familiar with the AIMWATER approach and results is in charge of the project. We refer to him as the contractor. Then that contractor would have to become familiar with the PEGASE system. Where would he find the necessary information? There is one basic report available in nine volumes prepared by the consulting firm BCEOM (1994) that developed the system for operational use. However, fortunately and somewhat in anticipation of incoming changes to PEGASE as a result of the AIMWATER study, IIBRBS has contracted with the BCEOM to prepare a full report describing the system as it

exists today and in the form also of a complete user's manual so that it could be used in house without a great deal of expertise about the internal intricacies of the system. Just as an example one would have to look at where the old model (routine) GR comes to play and replace it with the new version. Can one use for calibration of the new version the same code that served to calibrate the watersheds using the old version? Of course none of these steps are difficult, in principle, but they will require a great deal of reanalysis, redesign and software redevelopment.

Vice versa consider the scenario that an IIBRBS technician knowledgeable with PEGASE is the contractor. What would he need to know? Certainly the cursory description of steps described in chapter 3 to derive soil moisture indices and the description of the assimilation procedures in chapter 4. Scientists rarely prepare user's manuals. It is not their vocation. Unfortunately that's what's needed to put new research developments into general practice by technicians. Essentially one really needed step for eventual use by IIBRBS is the preparation of a comprehensive manual including only what, in the end, is germane to application by IIBRBS in complete details, proceeding from step to step, from basic acquisition of the data to the final calibrated rainfall-runoff real-time mode software.

Without trying to go into great details let us broadly define how that manual should be constructed. For every step the following (types of) questions should be addressed.

(1) What specific data acquisition is necessary? For example radar signal.
(2) Where and how does one get it? For example EO images.
(3) In what form does it come? For example as an EXCEL file.
(4) What transformation is required of the data so that it can be processed and be in a useful form for the next steps? For example removal of the vegetation effect.
(5) What procedures and algorithms need to be followed to accomplish this item? For example determining the Kalman gain in the filter.
(6) What software is needed and where does one get it? For example the GR model.
(7) In what form will it be available? For example as a FORTRAN source code
(8) What level of expertise on the part of the user is required to fulfill this step? For example an engineer degree, with a specialty in hydrology, to reanalyze the calibration of watershed parameters.
(9) Where is that personnel likely to be found? For example in a particular CEMAGREF or BCEOM group.

This is of course not an exhaustive list nor necessarily an appropriate one in all the steps needed for implementation in practical application of AIMWATER results into reservoir management. The establishment of a list of questions of that type should be a prerequisite for the preparation of each section of the manual.

- **Conclusion**

In practical terms the implementation of the methodology developed by AIMWATER involves many complex logistical issues, such as how to get the data in a timely fashion, how to get the satellite pictures, how to relate the EO data to actual basin moisture index for each of the sub basin of a large basin, etc. We have limited ourselves here to look at the type of model that would have to be implemented if one were to use the improvement in runoff forecast brought about by the knowledge of soil moisture to bear on the decision process to choose a release.

The proposed approach is one that would not require additional information from the one already available as parameters in the PEGASE model (geometric and hydraulic characteristics of the river system). However many alternative uncertainty formulations can be thought so that the decision made is not the best just in an expectation sense, relative to the objective, but one that has also the smallest variance. Since it is not possible to obtain that result (the "zero law" of optimization stating that it is only possible to optimize a single objective function) it is necessary to compromise by constraining the optimization for the expected objective to stay within a given range of variance or vice versa. In other words not just do the best on the average but eliminate the risk of making sometimes a very bad decision.

Given the fact in the face of uncertainty there is not a single objective that can be set unambiguously, the user will need to implement the new methodology of assimilation of EO data within a new optimization framework that is capable of displaying the impact of the magnitudes of various uncertainties and of the particular objective selected on the calculated values of the releases. The proposed demonstration would then not only show the merit of assimilation of a wetness index but also provide the decision maker with an array of options as to which objective to choose to implement in a given situation.

References

Aubert D., Loumagne C., Weisse A., Le Hégarat-Mascle, (2001), Assimilation of Earth Observation data into hydrological models : the sequential method. *International Symposium of Remote Sensing in Hydrology, IASH,* Montpellier, France, accepted.

BCEOM (1994). Étude d'optimisation de la gestion coordonnée des barrages réservoirs du bassin de la Seine. Study report. 9 volumes, June 1994.

Beven K.J. (1997). TOPMODEL: a critique. *Hydrological Processes*, **11(9)**, 1069-1085.

Edijatno, Nascimento N.O., Yang X., Makhlouf Z. and Michel C. (1999). GR3J: a daily watershed model with three free parameters. *Hydrol. Sci. J.*, **44(2)**, 263-278.

Le Hégarat-Mascle S., Poirier- Quinot M, Alem F, Weisse A., Loumagne C., (2001). Validation of a methodology to Monitor Soil Moisture from C-band SAR Spaceborne in an Operational Way. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, accepted.

Littlewood I.G., Down K., Parker J.R. and Post D.A. (1997). The PC version of IHACRES for catchment-scale rainfall-streamflow modeling. Version 1.0. User Guide. *Institute of Hydrology,* 89p.

Loumagne C., Chkir N., Normand M., Ottlé C. and Vidal-Madjar D. (1996). Introduction of soil/vegetation/atmosphere continuum in a conceptual rainfall-runoff model. *Hydrological Sciences Journal*, **41(6)**, 889-902.

Loumagne C., Normand M., Riffard M., Weisse A., Quesney A., Le Hegarat-mascle S., Alem F. (2001). Methodology for integration of remote sensing data into hydrological models for reservoir management purposes. *Hydrological Sciences Journal*, 46(1), 89-102.

Massé P., (1946). Les réserves et la régulation de l' avenir dans la vie économique: 1. Avenir Déterminé. Hermann & Cie, Publishers, Paris, France, 138 pages.

Morel-Seytoux H.J. (1998). Analyse Statistique des Pluies pour l' Utilisation de Séries Passées dans des Simulations en Temps Réel. December 1998, HYDROLOGY DAYS Publications, 57 Selby Lane, Atherton, CA 94027-3926, Report HWR 98. 5, 130 pages.

Morel-Seytoux H.J., (1999a). Needs of the IIBRBS and ARBSLP for additional data acquisition and/or modeling capability for better management of their reservoirs. Annual progress report for the first year. September 15, 1999.

Morel-Seytoux H. J., (1999b), Optimal Deterministic Reservoir Operations in Continuous Time. *J. of Water Resources Planning and Management.,* 25 (3), 126-134

Moreno J., Cuñat C.,Alonso L., Gonzalez MC, Garcia JC., 2001. Operational methodologies to retrieve surface parameters from EO data: Synergistic use of optical and SAR data. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, accepted.

Perrin C., Michel C., Andréassian V., 2001 : Does a large number of parameters enhance model performance ? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology,* 242, 275-301

Refsgaard J.C. (1997). Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrol.,* **28**, 65-84.

Oudin L., Weisse A., Loumagne C., Michel C., Le Hégarat-Mascle S., (2001). Assimilation of soil moisture into hydrological models for flood forecasting: a variational approach. *International Symposium of Remote Sensing in Hydrology, IASH,* Montppellier, France, accepted.

Appendix

Specific algorithm for the Extended Operation Rule

There are three main steps in the determination of the release or withdrawal. First one needs to establish the current situation in the reservoir. Second one needs to calculate the required volume in the reservoir for the current date depending on the current situation and the needs. Third one must account for the various additional constraints and then select the compatible desired reservoir content

First step: Determination of reservoir situation

One calculates the known actual reservoir content known, which is the content at the end of the previous day. One calculates the volume prescribed by the COR for the current day. One assesses for the current date the state of several logical operators (variables taking values only TRUE or FALSE or YES and NO). These indicators are: "drainage_period (VIDANG)", "local_operations (LOCALE)", "flood (CRUE)", "late_summer (ETGTAR)", "summer_operations (VIDETG)".

a. "drainage_period (VIDANG)" : If current date is within the drainage period, then VIDANG = TRUE, otherwise VIDANG = FALSE.

b. "local_operations (LOCALE)" : If there is a model point where a secondary flood reference discharge has been established, then LOCALE = FALSE, otherwise LOCALE = TRUE.

c. "flood (CRUE)" :
First case: If (LOCALE = TRUE) and if the either the upstream or downstream value of discharge exceeds the flood reference discharge for the current date, then it is a flood situation and CRUE = TRUE.
Second case: If (LOCALE = FALSE) then CRUE = TRUE. if either the upstream or the downstream discharge exceeds the flood reference discharge for the current date OR if the discharge at a model point (where a secondary flood reference discharge has been defined) exceeds that secondary flood reference discharge. In other words there is a state of flood at one or both two reference points (either just downstream from the reservoir or at a further downstream model point).

d. "late_summer (ETGTAR)" : ETGTAR = TRUE if the end of the drainage period is past and the discharge at the reference point is (still) less than the summer reference value.

e. "summer_operations (VIDETG) : VIDETG = TRUE if the three following conditions obtain simultaneously: (1) VIDANG = TRUE, (2) there is a model point where demand for water is estimated (through a set summer target discharge) and (3) the reservoir content is greater than the value on the bottom drainage curve ("courbe plancher de vidange). Otherwise VIDETG = FALSE.

Second step: Calculation of required volume in reservoir

a) If (VIDANG = TRUE) and if (reservoir content is less than the one defined by the COR) and if (the secondary summer reference discharge is exceeded) then the required volume in the reservoir is the one set by the COR. Proceed to third step.
If (CRUE = TRUE) then the reservoir is filled to the maximum allowed. Proceed to third step.
Else if (CRUE = FALSE) and if (the reservoir content exceeds the maximum volume) then the reservoir is to be drained as rapidly as possible to the COR value. Proceed to third step.
In all other conditions proceed to b).

b) If (VIDETG = TRUE) the release is calculated according to Eq.(1). Based on this calculated value the reservoir content is determined but is bounded upward by the value of the COR and downward by the bottom (ground floor , " plancher") curve. Proceed to third step.

Otherwise if (ETGTAR = TRUE) one drains the available volume in the remaining summer reserve block with a depletion rate based on the time difference between the end of the year and the end of the drainage period. Proceed to third step.
Otherwise if (VIDANG = TRUE; Note that in this case the variables previously tested VIDETG and ETGTAR, both have the value FALSE) one defines a new COR with a linear variation between the (low) content actually attained at the beginning of the drainage period and the normal volume to be attained at the end of the drainage period. Proceed to third step.
If in none of the four previously described situations, the required volume is that of the COR. The rate at which the reservoir content is to catch up with the desired COR volume is set as a parameter in the operations rules. Proceed to third step.

Third step: Accounting for the regulatory constraints

a) "Dead Block" ("tranche morte"). It is not permissible for reservoir content to drop below that volume and consequently previously estimated volume must be adjusted to meet that constraint.

b) Withdrawal or release. Given the now established required reservoir content to be attained at end of current day, one calculates the required mean daily discharge to attain the required volume.

First case: the required discharge is positive (diversion from river into reservoir). The value is constrained by the maximum discharge that can be diverted (canal capacity), the reserved discharge that must flow in the river past the point of diversion and the flood reference discharge. This latter constraint applies only if (CRUE = TRUE) and if there is no secondary flood reference discharge. If that situation obtains, the withdrawal amount must be such that the discharge downstream from the point of release be at least equal to the reference discharge. If on the other hand there is a secondary flood reference discharge (i.e. LOCALE = FALSE) then the reservoir is filled to the maximum.

Second case: the required discharge is negative (case of release). The release amount is limited as a function of the natural downstream discharge and the minimum discharge between the downstream reference value and the maximum discharge to be generated downstream by the release.

6. CONCLUSION AND PERSPECTIVES

The methodologies presented in this report have been validated over the selected areas of the project under two contrasting European climates, humid temperate and semi-arid Mediterranean. A critical assessment has been carried out to compare results provided by classical rainfall-runoff models and models including Earth Observation data with constant attention to the practical needs of the users for better management of their reservoirs.

Based on the results obtained so far, customers cannot use the methodologies set up for the project without applying the different steps suggested in the report to accommodate their current reservoir operations. Further work needs to be undertaken before these methodologies can be transferred in an operational context. Nevertheless the most important results of the project concern the derivation of EO soil moisture indices at a basin scale and the assimilation of this information into rainfall-runoff models. An overview of these results is provided in the Technological Implementation Plan along with their possible dissemination and potential use. Some perspectives on the use of EO data in the reservoir community are presented below.

With the implementation of these methodologies in reservoir operation it is expected that the project should yield real improvement in water management with economic and scientific benefits. First, there is the mitigation of detrimental effects of floods or long periods of low flows and the improvement of the cost-effectiveness of the existing models in the operational reservoir context. Second, there is the improvement in model performance as a result of assimilating Earth Observation data and the development of generic algorithms to derive soil moisture indicators from SAR data.

Perspectives on the use of EO data in the reservoir community

In AIMWATER the main parameter considered is soil moisture. For its retrieval active microwaves were selected for the better spatial resolution (10-20 m) of available platforms at the time of project development despite of their low temporal resolution (35 days).

ERS-SAR was the selected platform which provides data limited to C-band and only one polarization (VV), but different studies shows that the optimal retrieval of soil moisture needs several incidence angles and that diversity of frequency and/or polarization allows the discrimination of the vegetation and soil signal contributions, and the different effects of roughness and moisture in the soil signal component. At near future the ENVISAT platform will provide such characteristics. It is equipped with an Advanced Synthetic Aperture Radar (ASAR), operating at C-band; ASAR ensures continuity with the image mode (SAR) and the wave mode of the ERS-1/2 AMI. This sensor is technologically much more advanced than ERS and will acquire images in VV (as ERS), HH, and cross polarizations allowing for the first time the use of polarimetric information from satellite.

In this project we implemented some methodologies to the synergistic use of optical and radar images. The main limitation of such techniques is the temporal difference between the acquisition of the microwaves and optical data. Until ENVISAT there was not a platform capable to acquire both type of data over the same area. ENVISAT includes a MERIS sensor. MERIS optical sensor will provide high spectral resolution images (bandwidth of nm), which will allow for new applications not possible with broadband satellites (LANDSAT, SPOT). Although its spatial resolution (300×300 m) is far from the spatial resolution of SPOT (20×20 m), the radar/optical synergy will be possible.

About optical data a better spatial resolution and a high number of bands will increase the discrimination of different soil covers when land cover classifications are elaborated. One of the key responsibilities of NASA's Earth Science Office is to ensure the continuity of future Landsat data. The New Millennium Program's (NMP) first Earth Observing flight (EO-1), managed by NASA's Goddard Space Flight Center (GSFC), will validate revolutionary technologies contributing to the reduction in cost and increased capabilities for future land imaging missions.

Three revolutionary land imaging instruments on EO-1 will collect multispectral and hyperspectral scenes over the course of its mission in coordination with the Enhanced Thematic Mapper (ETM+) on Landsat 7. Breakthrough technologies in lightweight materials, high performance integrated detector arrays and precision spectrometers will be demonstrated in these instruments. Detailed comparisons of the EO-1 and ETM+ images will be carried out to validate these instruments for follow-on missions.
This NASA spacecraft is an order of magnitude smaller and lighter than current versions. The EO-1 mission will also provide the on-orbit demonstration and validation of several spacecraft technologies to enable this transition. Key technology advances in communications, power, propulsion, thermal and data storage are also included on the EO-1 mission.

EO-1 was launched on November 19, 2000. EO-1 flies in a 705 km circular, sun-synchronous orbit at a 98.7-degree inclination. This orbit allows EO-1 to match within one minute, the Landsat 7 orbit and collect identical images for later comparison on the ground.

Once or twice a day, sometimes more, both Landsat 7 and EO-1 will image the same ground areas (scenes). All three of the EO-1 land imaging instruments will view all or subsegments of the Landsat 7 swath. Reflected light from the ground will be imaged onto the focal plane of each instrument. Each of the imaging instruments has unique filtering methods for passing light in only specific spectral bands. Bands are selected to best look for specific surface features or land characteristics based on scientific or commercial applications.

For each scene, over 20 Gbits of scene data from the Advanced Land Imager, Hyperion, and Atmospheric Corrector will be collected and stored on the on-board solid-state data recorder at high rates. When the EO-1 spacecraft is in range of a ground station, the spacecraft will automatically transmit its recorded image to the ground station for temporary storage. The ground station will store the raw data on digital tapes, which will be periodically sent via overnight mail delivery to the Goddard Space Flight Center for processing and sent to the EO-1 science and technology teams for validation and research purposes.
The Hyperion instrument provides a new class of Earth observation data for improved Earth surface characterization. The Hyperion provides a science grade instrument with quality calibration based on heritage from the LEWIS Hyperspectral Imaging Instrument (HSI). The Hyperion capabilities provide resolution of surface properties into hundreds of spectral bands versus the ten multispectral bands flown on traditional Landsat imaging missions. Through this large number of spectral bands, complex land eco-systems shall be imaged and accurately classified.

The Hyperion provides a high-resolution hyperspectral imager capable of resolving 220 spectral bands (from 0.4 to 2.5 μm) with a 30-meter resolution. The instrument can image a 7.5 km by 100 km land area per image and provide detailed spectral mapping across all 220 channels with high radiometric accuracy. The instrument originally conceived a drop in to the ALI instrument and is now baselined to be a standalone instrument on EO-1. The major components of the instrument include the following:

- System fore-optics design based on the KOMPSAT EOC mission. The telescope provides for two separate grating image spectrometers to improve signal-to-noise ratio (SNR).
- A focal plane array, which provides separate short wave (SWIR) and visible spectral (VNIR) detectors based on spare hardware from the LEWIS HSI program.
- A cryocooler identical to that fabricated for the LEWIS HSI mission for cooling of the SWIR focal plane.

Hyperspectral imaging has wide ranging applications in mining, geology, forestry, agriculture, and environmental management. Detailed classification of land assets through the Hyperion will enable more accurate remote mineral exploration, better predictions of crop yield, and assessments, and better containment mapping. The perspectives of such EO data use in the reservoir community should be increased along with developed methodologies for their assimilation in operational tools.