THEME ENV.2012.6.1-1

EUPORIAS

(Grant Agreement 308291)

European Provision Of Regional Impact Assessment on a Seasonal-to-decadal timescale

Deliverable D23.3

Report on initialisation of impacts models for seasonal predictions

Deliverable Title	**Report on initialisation of impacts models for seasonal predictions**	
Brief Description	Correct initialisation of impacts models on the seasonal timescale may improve the predictive skill of the impacts model (e.g. crop sowing dates). The standard approach will be to initialise from spin-up using reanalysis products (e.g. ERA-Interim), or from climatology. This report describes progress on sensitivity experiments designed to advise on how impact models should be initialised in the EUPORIAS project.	
WP number	23	
Lead Beneficiary	Pete Falloon, Met Office	
Contributors	Met Office: Karina Williams, Pete Falloon, Carlo Buontempo, Paula Newton WUR: Ronald Hutjes, Wouter Greuell, Iwan Supit University of Leeds: Kathryn Nicklin SMHI: Ilias Pechlivanidis University of Cantabria: Rodrigo García-Manzanas University of Lund*:* Anna Maria Jönsson IC3*:* Rachel Lowe CETaqua: Laurent Pouget MeteoFrance: Jean-Pierre Ceron EDF : Laurent Dubus Predictia : Max Tuni San Martin	
Creation Date Version Number Version Date	16/07/2014 1.3 30/07/2014	
Deliverable Due Date Actual Delivery Date	31/07/2014	
Nature of the Deliverable	X	*R - Report*
		P - Prototype
		D - Demonstrator
		O - Other
Dissemination Level/ Audience	X	*PU - Public*
		PP - Restricted to other programme participants, including the Commission services
		RE - Restricted to a group specified by the consortium, including the Commission services
		CO - Confidential, only for members of the consortium, including the Commission services

Version	Date	Modified by	Comments
0.5	16/07/2014	P Falloon	Initial draft
0.6	21/07/2014	P Falloon	Added SMHI contribution
0.7	21/07/2014	P Falloon	Updated Met Office contribution
0.8	21/07/2014	P Falloon	Updated MeteoFrance contribution

0.9	21/07/2014	P Falloon	Initial edits
1.0	21/07/2014	P Falloon	First full draft version
1.1	24/07/2014	P Falloon	Edited MO/Leeds section
1.2	28/07/2014	P Falloon	Edited CETAqua section
1.3	29/07/2014	P Falloon	Edits from P Newton

Table of Contents

1. Executive Summary

An important aspect of most seasonal prediction systems is that the models are initialised using recent observations of various aspects of the earth system, such as sea surface temperatures, land surface properties (notably soil moisture and snow) and other factors. Initialisation may therefore also be an important aspect for using impact models driven by the outputs of seasonal prediction systems. This report summarizes findings on the importance of initialisation for Work Package 23 of the EUPORIAS project. Key findings are listed below:

1. Literature studies show that the impact of initialisation model outputs (and its relative contribution to overall skill) differs between types of impact models, differs between seasons and regions, with lead time and with variable;
2. The literature studies discussed here are mainly for hydrology and there are generally fewer studies for other impact sectors. In hydrology, initial hydrological conditions can play a crucial role in overall impact skill;
3. It may be important to consider socio-economic factors in initialisation and model setups for the work-package (for instance, dam operation rules and water demand), especially at smaller spatial scales; in addition it may be important to consider climate extremes which may not be represented in seasonal means;
4. Studies performed by Work Package 23 partners generally confirm the literature findings in 1) above; hence the need for model spin-up and initialisation varies with sector/impact and model. In general the importance of initialisation and spin-up is greatest for hydrological models, and least for crop models.

2. Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOW, Section B1.1):

No.	Objective	Yes	No
1	Develop and deliver reliable and trusted impact prediction systems for a number of carefully selected case studies. These will provide working examples of end to end climate-to-impacts-decision making services operation on S2D timescales.	X	
2	Assess and document key knowledge gaps and vulnerabilities of important sectors (e.g., water, energy, health, transport, agriculture, tourism), along with the needs of specific users within these sectors, through close collaboration with project stakeholders.		
3	Develop a set of standard tools tailored to the needs of stakeholders for calibrating, downscaling, and modelling sector-specific impacts on S2D timescales.	X	
4	Develop techniques to map the meteorological variables from the prediction systems provided by the WMO GPCs (two of which (Met Office and MeteoFrance) are partners in the project) into variables which are directly relevant to the needs of specific stakeholders.		
5	Develop a knowledge-sharing protocol necessary to promote the use of these technologies. This will include making uncertain information fit into the decision support systems used by stakeholders to take decisions on the S2D horizon. This objective will place Europe at the forefront of the implementation of the GFCS, through the GFCS's ambitions to develop climate services research, a climate services information system and a user interface platform.		
6	Assess and document the current marketability of climate services in Europe and demonstrate how climate services on S2D time horizons can be made useful to end users.		

3. Introduction

Despite the recent effort to develop underpinning climate prediction science for seasonal to decadal (S2D) predictions, there has been relatively little uptake and use of S2D forecasts by users for decision making in Europe (Dessai and Soares, 2013). On the other hand, there is a much longer tradition in applying seasonal forecast information for user applications in other parts of the World, notably in Africa, the USA and Australia (Hansen et al. 2011; Dessai and Soares, 2013); one notable exception is the use of precipitation forecasts for hydropower generation management by EDF (Dubus, 2012,2013). In part, this is related to the relatively limited skill of S2D forecasts in Europe; in contrast predictability in decadal hindcasts (forecasts of the past) is greatest in the Tropics (MacLeod et al. 2012). This illustrates the importance of understanding skill in user uptake of such products (Meinke et al. 2006; Davey & Brookshaw, 2011; Demeritt et al. 2013). However, it should be noted that accuracy, lead time, and appropriate spatial and temporal scale of S2D forecast information may not be the main (or only) factors influencing user uptake; potential economic and environmental benefits may be of greater importance (Marshall et al., 2011). In addition, probabilistic (ensemble) prediction systems are more commonly used in medium-range applications, which bring additional challenges in communicating forecast information to end-users.

The use of basic S2D forecast outputs (e.g. temperature, precipitation etc) has significant potential to support both shorter-term decision making (thus helping avoid potential risks and losses, and optimize profits), and longer-term climate adaptation plans in numerous sectors (e.g., agriculture, water, health and energy – Van der Linden and Mitchell, 2009). Further benefit could also be realized by providing information more directly relevant to potential users, such as changes in crop yields, river flows, and forest productivity which we refer to here as “impacts” for the purpose of this report. In addition, further processing of direct S2D forecast outputs and the use of impact models may improve the usability of S2D forecasts with weak skill (Dubus 2012, 2013). However, as noted above, the skill of S2D forecasts for impacts (as opposed to generic assessments of weather and climate skill) may limit the usability of S2D impacts products.

S2D predictions of weather and climate can be derived both from statistical (or empirical) and dynamic models (Davey & Brookshaw, 2011). The former approach is usually based on regional historic relationships between climate variables; most recent dynamic approaches use fully coupled ocean-atmosphere general circulation models (CGCMs). For instance, the Met Office Hadley Centre (MOHC) currently has two operational CGCM-based systems for S2D forecasting, the latest versions of which are both based on the HadGEM3 model: GloSea5 (Global Seasonal Prediction system version 5) and DePreSys (Decadal Prediction System).

Some S2D forecasting systems, particularly the CGCM approaches, may include impact-relevant outputs directly, for example via river flow models, soil moisture calculations, or estimates of vegetation productivity. Validation and skill assessment in these systems may also provide valuable information on the overall performance of the seasonal prediction system itself. For example, rivers integrate land hydrology over large geographic areas and are important sources of freshwater input to the oceans (Falloon et al. 2011). However, drivers for, and focus of skill assessments for CGCM development versus impact (and user) application may differ.

3.1 Approaches for estimating impacts from S2D forecasts

As noted above, there is considerable diversity in the methods used to produce estimates of climate impact from S2D forecasts, including:

1. Outputs directly from the CGCM models themselves (e.g. extreme temperature and precipitation, storm tracks and cyclones, soil moisture and runoff; as noted above, some CGCMs may include river flow or vegetation models, and so may produce estimates of river flow, vegetation productivity and other impacts). These are termed coupled or online approaches;
2. Simple (offline) metrics derived from S2D forecast outputs (e.g. statistical relationships between temperature and crop yield, or rainfall and river flow; heating and cooling day requirements); a brief review for agriculture is provided by Hansen et al. (2006);
3. Outputs from more complex (offline) impact models, using S2D outputs as their inputs (e.g. crop, hydrological, ecosystem or other dynamic models). Hansen et al. (2006) provide a review relevant to agriculture; and
4. Classification and analogue approaches, as reviewed by Hansen et al 2006. These may include classifying the climate into certain typical phases or types (e.g. for ENSO – El Nino and La Nina) and then producing weather data from past years related to these phases (analogs) as input to impact models (or statistical approaches). Weather classification approaches are similar, but instead cluster historic data into particular circulation patterns or weather types.

The primary focus of this report is on methods 1 and 3, particularly on complex (offline) impact models.

3.2 General methodological issues for estimating skill in S2D impacts

This diversity in potential methods implies that there may also be a range of potential ways of assessing skill. An important consideration for evaluating skill of S2D impacts is therefore that the approach taken (including the variables studied and methods used) need to be relevant to the application in question. Examples of the potential range of phenomena to be evaluated could include (Falloon et al. 2013):

- The occurrence of events (e.g. crop failure – yield below a threshold; the existence of a heat-wave, drought, or flood)
- The magnitude and timing of events (e.g. river flow patterns, anticipated crop yield amount and harvest date)
- The probability of particular events (in both categories above) occurring

Furthermore, impacts themselves will be affected by different outputs of S2D forecast models (themselves with different levels of skill) in different ways, depending on their relative importance for the different sectors. In addition, even within one sector (e.g. energy), different processes will be affected to different extents by weather and climate events (Dubus, 2010).

The methods for evaluating S2D impact skill require a broad consideration of various aspects of the impact estimation process (Challinor et al. 2005), including:

1. Experimental design, especially how models are initialised and spun-up (e.g. Cosgrove et al. 2003); and experiments used for comparison with S2D hindcast impact estimates (e.g. impact models driven with observed climatology);
2. Post-processing applied, including bias-correction and downscaling methods (see the section below for further details);

3. Statistics/metrics used for validation; different techniques may be needed for deterministic or probabilistic approaches (Falloon et al. 2013); and
4. Observed data to be used for validation. For S2D hindcasts, the specific periods of data availability (e.g. yield time-series) may be important, since hindcasts are often only run for certain periods. In addition, it may be necessary to correct observed data – for instance, technology trends lead to increases in yield with time regardless of climate (Challinor et al. 2005), and anthropogenic influences on river flow are often not included in impact models (Falloon et al 2011; Haddeland et al 2011).

This report focuses particularly on 1. above, especially regarding model initialisation, as part of the "production chain" for seasonal impact assessments in the EUPORIAS project.

3.3 Recent studies on the impact of initial conditions to forecast skill

An important aspect of CGCM-based S2D prediction systems is that the models are initialised using recent observations of various aspects of the earth system, such as sea surface temperatures, land surface properties (notably soil moisture and snow) and other factors. The effect of initialisation uncertainties differs between (type of) impact models, differs between seasons and regions and is probably a significant fraction of the overall impact forecast uncertainty, or inversely its skill.

There are generally more studies for hydrology compared to other impacts. Several studies have investigated the relative importance of initial hydrological conditions (IHCs) and climate forecast (model) skill (CFS) in the overall skill of seasonal impact forecasts. Several studies have applied an Ensemble Streamflow Prediction (ESP) approach to investigate this issue, both for the USA and globally (Shukla & Lettenmaier, 2011; Shukla et al. 2013).

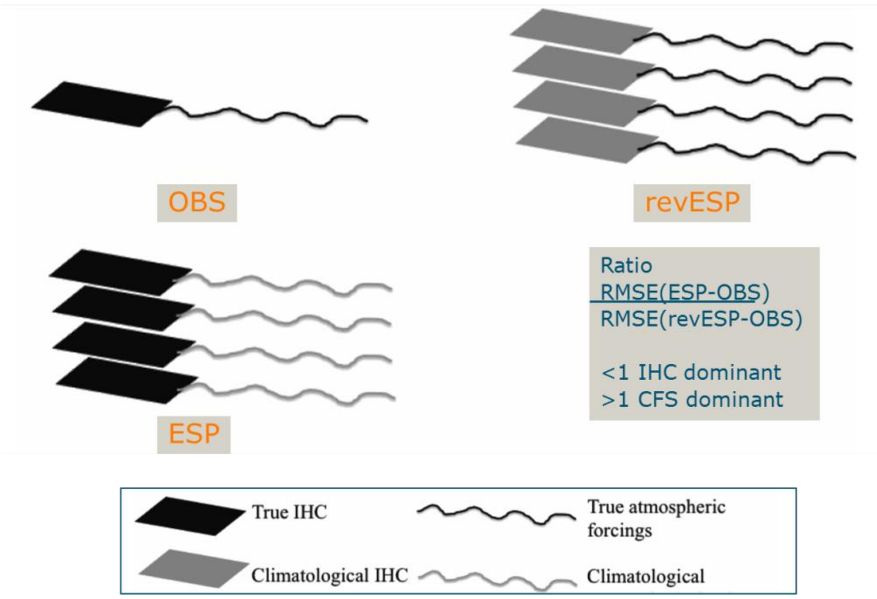


Figure 1: from Shukla & Lettenmaier (2011): schematics of analyzing influence of Initial Hydrological Conditions versus Climate forecast Skill on runoff forecasts

Paired ESP and reverse-ESP experiments were used to assess the importance of IHCs and CFS in determining forecast skill. The ESP experiment gains all its forecast skill entirely from knowledge of the IHCs whereas in the reverse-ESP (rESP) forecast skill is derived only from knowledge of the atmospheric forcings. Using the ratio of Root Mean Square Error (RMSE) in predicting cumulative run-off and mean monthly soil moisture of each experiment allowed identification of the variability of the relative contributions of the IHCs and CFS spatially throughout the year.

Shukla & Lettenmaier (2011) studied the significance of initial conditions versus forecast quality in determining the skill of predicting six month cumulative runoff, using an approach schematised in *Figure 1*. The ESP experiments of Shukla & Lettenmaier (2011) consisted of a land surface model driven with observed forcings until the forecast initialisation date thus generating the IHCs. During the forecast period, an ensemble of forcings was created from the time series of observations starting at the forecast initialisation date and continuing until the end of the forecast period. For rESP experiments, the IHCs on the forecast date were taken from each of the historical years of simulation and during the forecast the land surface model was run with gridded observations for that year.

Shukla & Lettenmaier (2011) found that IHCs generally had the strongest influence over cumulative run-off and soil moisture over the first month of a forecast beyond which their influence decreases at rates depending on location, lead time and initialisation date (*Figure 2*). Beyond one month, IHCs influence the cumulative runoff and soil moisture during spring and summer months, mostly over the western USA. CFS dominated both runoff and soil moisture forecast skill beyond one month mainly over the Northeast USA throughout the year. For the rest of the region, CFS dominated forecasts during autumn and winter. The relative contributions of IHCs and CFS had a first order relationship with the ratio of initial total moisture variability to the variability of precipitation during the forecast period.

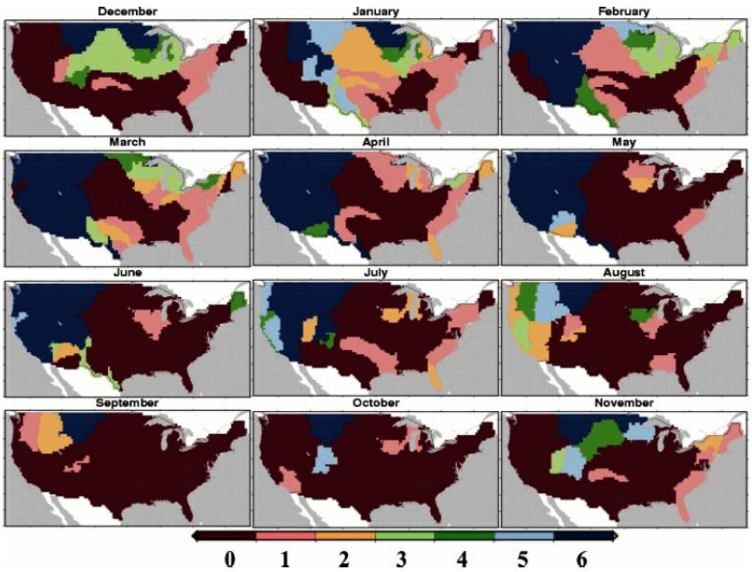


Figure 2: from Shukla & Lettenmaier (2011): Plot of the maximum lead (in months) at which Initial Hydrological Conditions dominate over Climate forecast Skill, for 6-month cumulative runoff forecasts, initialised on the beginning of each month

Shukla et al. (2013) created ESP experiments with the VIC hydrological model initialised with true IHC and forced with an ensemble of atmospheric forcings randomly sampled from the period 1961-2007. In the rESP experiments, the model was initialised with ensembles of the

IHCs randomly sampled from the same climatological period. Each ensemble was then forced from the forecast date onwards with observed atmospheric forcings for the target year. IHCs were found to play a crucial role in determining seasonal hydrological skill globally.

Generally, the contributions of the IHCs were greater than the contribution of the CFS over Northern Hemisphere during the forecast period starting in October and January (Southern Hemisphere – April and July), mainly over shorter lead-times. Over snow dominated regions the Northern Hemisphere IHCs dominate runoff forecast for up to six months during the forecast period starting in April. Overall the contribution of the CFS was higher than IHCs over the tropics throughout the year. The contribution of IHCs especially over the first month was greater for soil moisture than run-off.

Cosgrove et al. (2003) initialised models with re-analysis, 100% wet and 100% dry soil moisture conditions and found that the ‘memory’ of the initial conditions varied regionally and was shorter for 100% wet initial conditions.

Demirel et al. (2013) investigated the effect of uncertainty from model inputs, parameters and initial conditions on 10 day ensemble low flow forecasts. They used two hydrological models (GR4J and HBV) applied to the Moselle river. Forecasts were generated using an ECMWF model ensemble, consisting of 51 members, thus providing the uncertainty range of model inputs. The Generalised Likelihood Uncertainty Estimation (GLUE) approach was used to estimate parameter uncertainty. Part of the GLUE parameter set was directly linked to river water storage estimates, and the observed discharge on the forecast issue day were used to update model stores and reflect uncertainty in initial conditions.

The GR4J model was found to overestimate low flows whereas HBV was more prone to underestimating low flows, especially if parameter uncertainty was included in forecasts. The forecasts of the HBV model incorporating input uncertainty resulted in the most reliable forecast distribution. Parameter uncertainty reduced the number of event ‘hits’, and the false alarm rate of GR4J was approximately twice that of HBV. In general, Demirel et al. (2013) found that parameter uncertainty had the largest effect on the medium range low flow forecasts, whereas the input uncertainty had the smallest effect.

Koster et al. (2010) assess the relative contributions of early-season snow and soil moisture information to the skill of riverflow forecasts in a suite of land-modelling systems (four different Large Scale Hydrological Models, or LSHMs: VIC, Noah, Catchment and Sac) for the USA, using the snow and soil moisture information both together and separately to derive seasonal forecasts. They used river flow gauge observations which were “naturalised” to remove the effect of anthropogenic influences such as dams, withdrawals and reservoir evaporation. In their study, skill was assessed using the square of the correlation coefficient (r^2) between the observed yearly time series of seasonal (March-July) river flows and the corresponding multimodel average time series during periods of overlap. In other words, skill was assessed in terms of reproducing observed interannual variability in streamflow and bias was not considered.

Their skill analysis showed that early-season snow-water storage generally made the greatest contribution to skill for predicting variability in more northerly and mountainous catchments, but the contribution of early-season soil moisture was also significant in the more southerly regions (*Figure 3).* Koster et al. (2010) suggest that the current land-surface models driven with large-scale meteorological data can produce estimates of soil and snow water storage that are useful for basin-scale prediction.

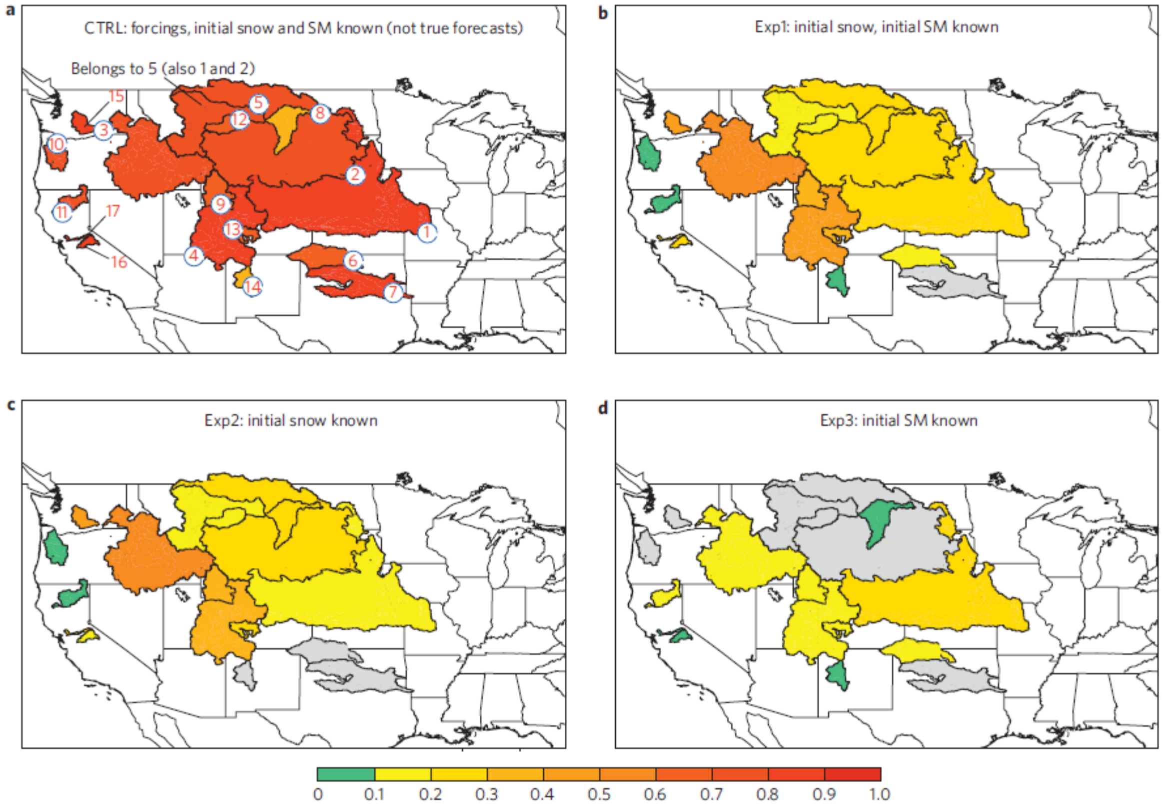


Figure 3: from Koster et al. (2010): Streamflow skill levels in the Western United States achieved in the simulation experiments, plotted by basin. Skill is measured as the square of the correlation coefficient (r^2) between MAMJJ total streamflows from simulations and corresponding (naturalized) measurements.

The studies discussed above illustrate that the duration of the influence of initial conditions varies regionally, with lead time, season, forecast starting date and with variable (e.g. soil moisture, snow, river flow).

4.Summary of progress under work package 23

4.1 Implications of previous studies and deliverables D23.1/D23.2 for work package 23

The work package has:

- developed a prototype operational workflow to use the impact models in S2D forecast mode, both through defining ways to initialise the models via a workshop and series of reports (D23.1, D23.3), and through agreeing a modelling protocol for using seasonal hindcasts with the models. Most partners have set up their models with observed forcing datasets as a baseline, and have begun to download the currently available seasonal hindcasts from the ECOMS portal; and

- Begun to assess and improve the predictive skill of impact models by analysing simulations driven by seasonal hindcasts. For example, CETAqua have analysed sensitivity of water impacts to seasonal climate conditions and the Met Office have run crop simulations using GloSea5 seasonal hindcasts. The work package has produced milestone reports on low- and high-end impact events/case studies to focus on (e.g. discharge, crop yields, etc).

As noted in section 3.3 above, recent literature studies have demonstrated that the effect of initial conditions varies regionally, with lead time and with forecast starting date, and is an important aspect to consider in seasonal impact prediction systems. *Table 1* below summarises the impact models being used in Work Package 23 (WP23). A wide range of models are being applied, both across and within sectors, each with differing needs and potential benefits from initialisation.

Following on from the workshop on model initialisation (and in D23.1 and D23.2) it was noted that impact models targeting systems that exhibit distinct memory effects (known or presumed) may need proper initialisation of their state variables at the start of a forecast/hindcast simulation. This may be expected to apply especially to models of hydrological systems where significant stores of soil moisture, snow and surface water in lakes/reservoirs/wetlands may reflect accumulated effects of past fluxes. Similarly this would apply to models of vegetation dynamics though probably more so for perennial vegetation (e.g. forests) than for annual vegetation and crops. As a result impact models for sectors that build on these, e.g. hydropower or forestry, likewise may be sensitive to initial states. Initialisation of impact models for systems that are sensitive to instantaneous weather impacts only, e.g. solar and wind power or tourism, on the contrary is likely to be relatively unimportant.

The relative importance of initialisation, and its dependency on the nature of the impact model being used was confirmed by the workshop participants representing the various impact modelling groups, based on their expert judgement and existing literature. For the particular models used in the consortium the effect of various possible approaches towards initialising relevant state variables based on model spin-up, on climatology or on observations (e.g. remote sensing) needs to be assessed. It was agreed during the workshop (and noted in D23.1/D23.2) that sensitivity experiments would only be performed for those models where initialisation is considered critical.

In D23.2, we proposed adopting an approach similar to that used in Shukla & Lettenmaier (2011) for exploring the sensitivity of the impacts models to initial conditions. This should provide a means of quantifying the uncertainty in model predictions that related to the value of the initial conditions as a function of lead time, start time in the seasonal cycle and region. Such sensitivity experiments can be done on full climatological skill statistics, i.e. for the full period for which the GloSea5 or System4 hindcasts are available, but also on (common) studies of particular events in specific regions. The latter may provide more insights as to why our impact models do or do not show skill in high/low anomalies in certain regions/lead times through detailed analysis of propagation of errors in initial conditions, forcing data or parameters. Selection criteria for case studies include that the impact anomaly (e.g.

anomalous crop yield) must be caused by a climatological anomaly. Regional stakeholder expertise may be needed to determine which events were driven by climatological rather than socio-economic effects (e.g. CAP reforms in the EU, or civil unrest in E-Africa).

Table 1: Impact models being used in EUPORIAS WP23

Sector	*Model*	*Forcing*	*Scale*	*Resolution*	*Forecast Variables*
Agriculture	JULES/ JIM *MO*	WFD, CRU-NCEP	Global	0.5 degree and 1.25*1.874 and 2 degree versions	Crop Yield Crop NPP River flow
	GLAM crop model *Leeds*	Daily Min and max temp, precipitation and solar radiation	Regional (e.g. all of India, semi-arid West Africa, China)	Typically 0.5 degree to 2.5 degree grid cells.	Crop yield Crop biomass.
	LPJmL *WU*	WFD	Global	0.5 degrees	Crop Yield River discharge Reservoir volume
	CGMS *WU*	WFD	Regional	25km	Crop yield
Hydrology	VIC *WU*	WFD	Regional	0.25 degrees	River discharge Water Temperature
	MORDOR *EDF*	ECMWF	North Atlantic/ Europe	2.5 degrees	River flow
	E-HYPE *SMHI*	ERA-INTERIM with monthly bias correction against GPCC	Europe	215 km2	Discharge Water quality
	Coupled models for decision making at the river basin agency level *CETaqua*	Seasonal forecast data	River basin	Various	River flow System reliability
Forestry	GUESS Storm-Ips *Lund*	Daily Temp, Precip, Radiation, Wind	Europe	0.5 degrees or lower	Risk of damage to forest
Health	Temperature related mortality statistical model *IC3*	ERA-Interim temperature	Europe	NUTS2 administrative regions	Mortality

Case study selection is also dependent on e.g. response options stakeholders may have had in any particular event depending on the outcome and skill of a forecasts would that have been available at the time. Stakeholder engagement is vital to the case study selection process, but may also influence the selection of appropriate skill metrics or their visualisation.

D23.2 noted that possible sources for data to initialise hydrology and vegetation related state variables in the impact forecast models include the following (availability may differ as for historical data for hindcast initialisation or (near-) real-time data for real forecast initialisation):

1. Real observations reflecting actual status at simulation start time or representing a climatological average for that moment in the seasonal cycle. For example, observed snow cover and snow depth from weather stations and/or satellite products, observed water levels in lakes reservoirs, soil moisture status from satellite products, vegetation status (biomass, LAI) from satellite products. Translation of observations to model variables is not always trivial.
2. Assimilated products from other operationally run models. For example, soil moisture/snow status from (re)analysis products from the operational weather centres, or from off line assimilation systems (e.g. LDAS, GLEAM, etc.). Translation of variables between models may lead to (arguably relatively small) inconsistencies.
3. Using appropriate spin-up times for the impact models themselves, forcing them with observed or (re-) analysed weather data. Translation issues mentioned above are naturally prevented. However, drift in the impact model may cause biased initialisation.
4. 'Guestimates' of initial states. For example, at the end of the dry season in semi arid climates the soil moisture can simply be set to very low values. Crop models generally start from zero biomass.

The second method was briefly discussed during the workshop on model initialisation. The impact models may be initialised with seasonal climate model forecast output. However, problems may arise as the latter models are tuned; i.e., soil moisture from the seasonal climate forecast model may not be appropriate to input to the impacts model as it has been adjusted to reduce biases in the 2m air temperature. In addition, literature suggests that initialisation from a different model may cause problems.

For example, Cosgrove et al. (2003) initialised an hydrology model (MOSAIC LSM) with either soil moisture status taken from NCEP re-analysis, or starting with 100% wet and 100% dry soil moisture conditions respectively and found that the 'memory' of the initial conditions varied regionally (across the USA) and was shortest for re-analysis initialisation (0-18 months, average nine months), about two years longer for 100% wet initial conditions and another two years longer for the 100% dry initialisation. These time scales roughly apply equally to total column, root zone soil moisture and evaporation. Spin-up to equilibrium was much sorter for soil temperature. Soil moisture memory varied strongly between climate zones and between different LSMs. Obviously, careful assessment of such effects needs to be done in case initialisation states are taken from independent models. In WP23, none of the partners is presently planning to use this method.

The third method from the list above is the preferred method to be used by the consortium members in EUPORIAS WP23 and WP31. Having discussed various observational datasets, it was decided that the Watch Forcing Data ERA-Interim (WFDEI) which combined the ERA-interim and GCPC products would be the most appropriate for spin-up and initial conditions and to produce a climatology of impacts. This dataset covers the 1979-2011 time period at 50 km resolution and daily (and if needed three hourly) resolution. With this dataset, all models can perform a single continuous run for the whole period from which initial states can be taken for the seasonal hindcast runs, for example being forced by the GloSea5/System4 data (*Figure 4*). There is additional consistency in this in the sense that also the seasonal climate forecasts from both GloSea5 and System4 are themselves initialised from ERA-Interim. This is the preferred approach ideally to be followed for all the impact models to be run for the European domain. For the agricultural models to be run for the East African domain the fourth method from the above list may optionally be used, as any carry-over of soil moisture and or crop status from the previous year is likely to be negligible.

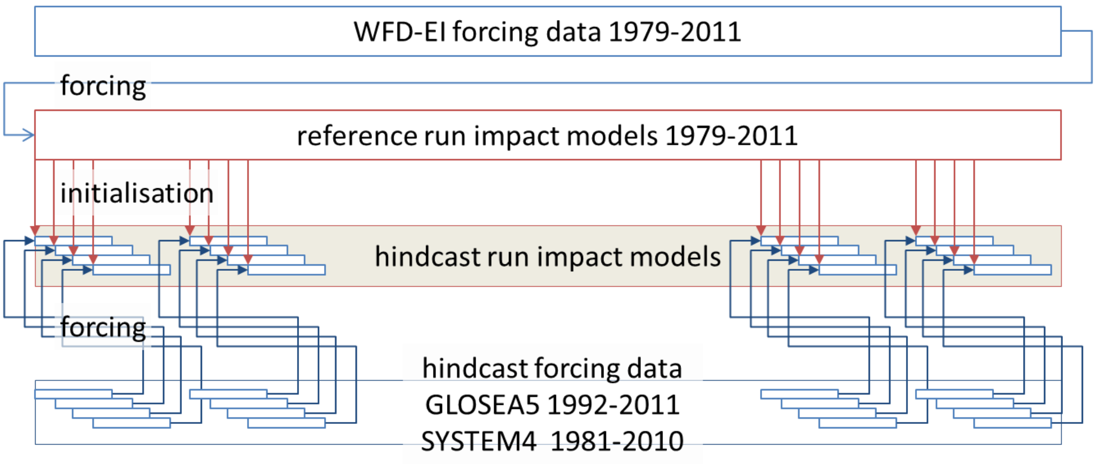


Figure 4: Schematic representation of modelling approach in EUPORIAS WP23

In summary, the EUPORIAS WP23 and WP31 partners have agreed that:

- Our overall aim is to provide the best model performance possible to meet stakeholder needs, rather than to perform a strict model inter-comparison experiment;
- There is a need to perform sensitivity experiments test using to assess the effects of different climate model forcing data (with/without bias-correction), the effect of impact model initialisation uncertainties (using various sources, or arbitrary changes e.g. +/-20% soil moisture/snow values), and compare against our "best" forcing and initialisation estimates;
- A common climatology (WFDEI) would be used for reference forcing and general initialisation. We would favour:
    - Spin-up using WFDEI, the period depending on the model being used.
    - Run using both raw and bias corrected seasonal forecast/hindcast model data;
- Sensitivity experiments can be performed using full climatological skill statistics, but also on (common) studies of particular events (exhibiting both weather and impact anomalies, and the latter caused by the former, not e.g. socioeconomic conditions).

The latter may provide more insights as to why our impact models do or do not show skill through detailed analysis of propagation of errors in initial conditions, forcing data or parameters; and

- Stakeholder engagement is vital to the case study selection process, but may also influence the selection of appropriate skill metrics or their visualisation.

4.2 Progress in model initialisation for the water sector

CETaqua have tested the use of seasonal predictions as input of the impacts models to predict the potential impacts. The first results have been commented to the project stakeholders. In particular, they have performed experiments to check the sensitivity of different type of impacts (reservoir filling, urban water demand) to seasonal climate conditions. The response to the climate drivers was analyzed over a range of different time scales for a small river basin in Spain (part of the Ebro river basin, Najerilla sub-basin with the Mansilla dam; *Figure 5*) and, regarding water demand, for an urban area close to Barcelona (*Figure 6*).

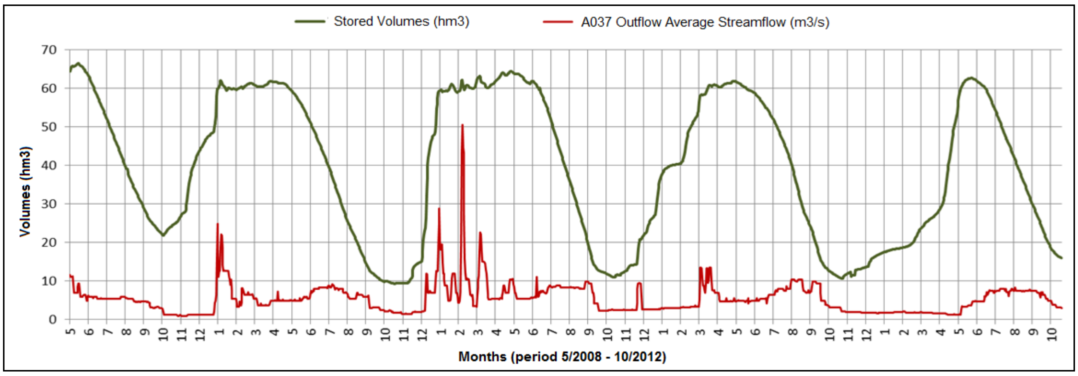


Figure 5: Volumes in the Mansilla dam and releases for the period 2008-2012

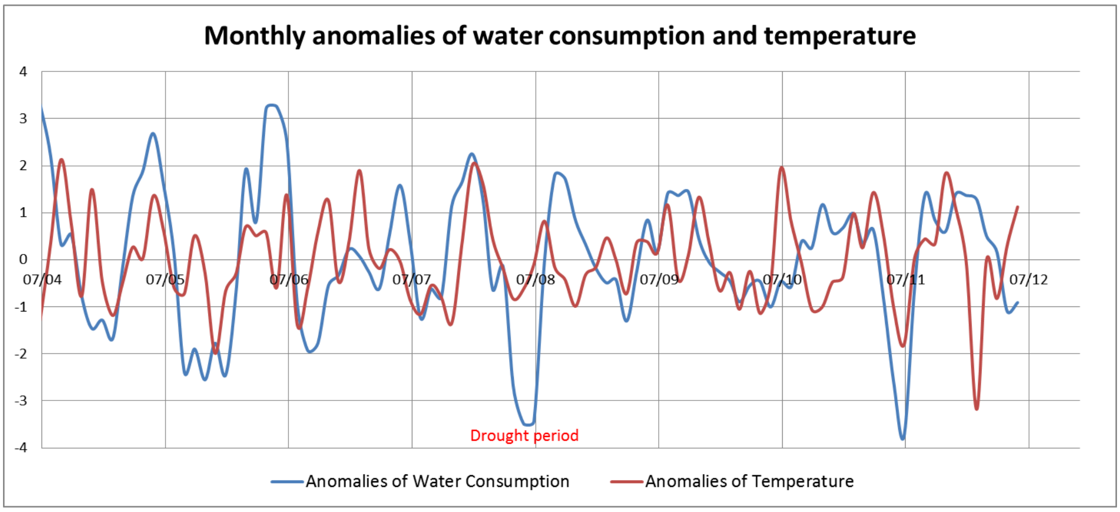
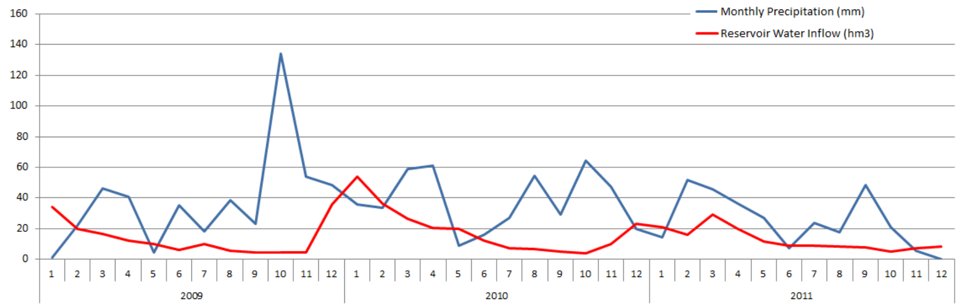


Figure 6: Monthly comparison between anomaly in temperature (red) and water demand (blue) for and urban area close to Barcelona, for the period 2004-2012

Using the data available, CETAqua assessed the response of a sub-basin in the Ebro River basin to monthly and seasonal climate drivers. The objective was to estimate the inflow to a downstream reservoir (Mansilla dam). The preliminary tests performed by CETaqua do not show a very clear relationship between regional and local precipitation (at one weather station in the basin) with the Mansilla dam filling (estimated from a downstream gauging station and the water level in the dam; *Figures 6 and 7*). Potential reasons are the non-representativeness of the local weather data, the imprecision in the estimation of the inflow, and the great importance of various processes not considered and occurring at smaller scale (intense local precipitation, evapotranspiration, small aquifers, etc.). In this case, a well calibrated and initialised hydrological model would be necessary to estimate the future inflows for the dam.

The usefulness of getting such inflow prediction for the Mansilla dam has also been assessed. It would be limited to some optimization of the energy production during the dam filling period. Indeed, the current management rules do not allow much flexibility since the main purpose of the dam is for irrigation (the dam should be filled before the irrigation period, which has been the case in the last 10 years) and since the size of dam do not allow pluriannual management (the dam fill and empty every year). In conjunction with AEMET and CETaqua it was decided to extend the geographical scope of the case study to cover a larger basin with pluriannual dam(s) and more data available. In terms of model initialisation, these findings imply that it may be important to consider socio-economic factors in initialisation, particularly for smaller scale basins, given the weak relationships found between seasonal climate drivers, reservoir filling and urban water demand.



.

Figure 7: Monthly comparison between estimated inflow to the dam and local rainfall data

SMHI are using the E-HYPE hydrological model in WP23. The Hydrological Predictions for the Environment, HYPE, model is a semi-distributed rainfall-runoff model capable of describing the hydrological processes at the basin scale (Lindström et al., 2010). The model represents processes for snow accumulation and melting, evapotranspiration, soil moisture, discharge generation, groundwater recharge, and routing through rivers and lakes. HYPE simulates the water flow paths in soil which is divided into three layers with a fluctuating groundwater table. In addition, parameters are more linked to physiographical characteristics in the landscape, such as Hydrological Response Units (HRUs) linked to soil type and depths and vegetation. Elevation is be used to get temperature variations within a sub-basin to influence the snow conditions. The model requires information on terrain, soil and land use, lakes and reservoirs and irrigation as input, which, in this application, has been obtained from the global sources.

Irrigation in HYPE is simulated based on crop water demands calculated either with the FAO-56 crop coefficient method or relative to a reference flooding level for submerged crops (e.g. rice). The demands are withdrawn from rivers, lakes, reservoirs, and/or groundwater within and/or external to the sub-basin where the demands originated. The demands are constrained by the water availability at these sources. After subtraction of conveyance losses, the withdrawn water is applied as additional infiltration to the irrigated soils from which the demands originated.

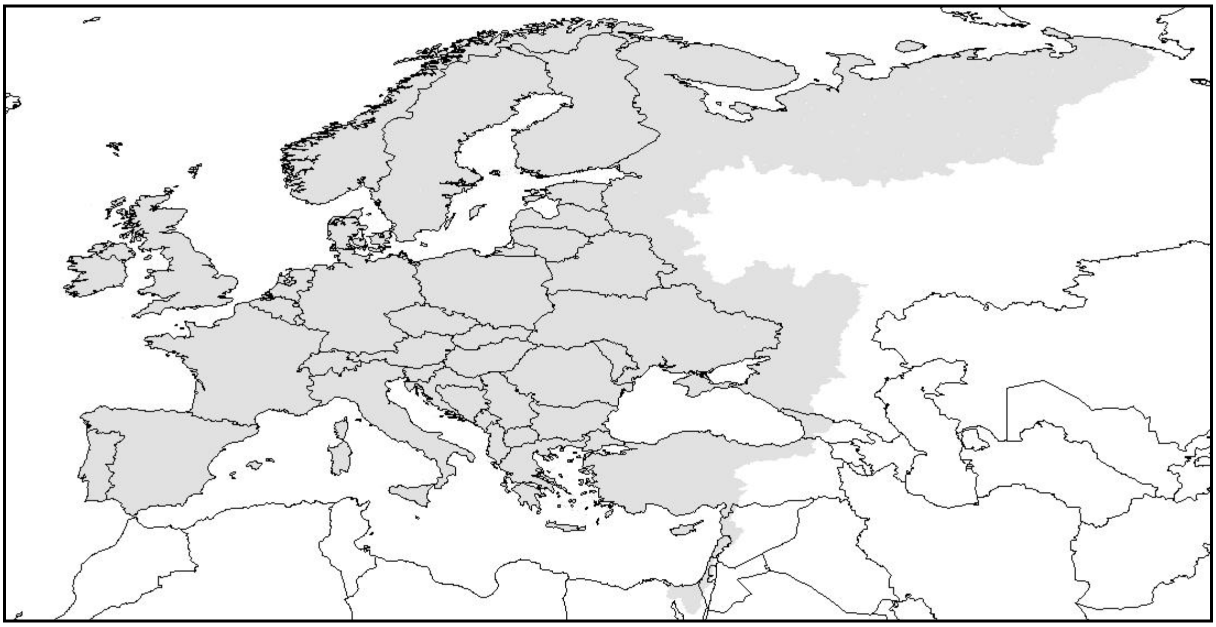


Figure 8: Domain of the EHYPE model

Table 2: Data sources and characteristics of the model setup

Characteristic/Data type	Info/Name	Provider
Total area (km²)	8.8 million	-
No. of sub-basins	35408 (mean size 215 km²)	-
Topography (routing and delineation)	hydroSHEDS (15 arcsec)	Lehner et al. (2008)
Soil characteristics	Harmonised World Soil Database (HWSD)	Nachtergaele et al. (2012)
Land use characteristics	CORINE	Bartholomé et al. (2002)
Reservoir and dam	Global Reservoir and Dam database (GRanD)	Bernhard et al. (2011)
Lake and wetland	Global Lake and Wetland Database (GLWD)	Lehner & Döll (2004)
Irrigation	Global Map of Irrigation Areas (GMIA)	Siebert et al. (2005)
Discharge	GRDC, EWA and others (2690 stations)	http://www.bafg.de/GRDC
Precipitation	WFDEI (0.5^o x 0.5^o)	Weedon et al. (2011)
Temperature (mean, min, max)	WFDEI (0.5^o x 0.5^o)	Weedon et al. (2011)
Snow cover area	GlobSnow	Weedon et al. (2011)

The HYPE model setup for the pan European region (8.8 million km^2; see *Figure 8*) is currently improved. The model has a spatial resolution of 35,408 sub-basins, i.e. in average 215 km^2 and is referred to as EHYPE. The model runs at a daily time step and is currently evaluated at 2690 discharge stations. The model is setup based on global available datasets which are listed in *Table 2*. Forcing data are based on the WATCH-ERA INTERIM (WFDEI) product.

The EHYPE hydrological model needs initial conditions (level in surface water, i.e. reservoirs, lakes and wetlands, soil moisture, snow depth) that will be obtained by driving the model using observations for a spin-up period. The model can further run using pre-calibrated model parameters. The model will be evaluated using the following performance criteria:

- Root mean square error
- Tercile probability
- Outer quantile
- Reliability diagrams
- ROC scores

In their investigation, SMHI aim to develop an impact model directly been useful to the end-users; hence an adequate model performance in terms of discharge and other hydrological variables is important. EHYPE model is currently been setup and calibrated using observed discharge stations. A preliminary analysis of model adequacy focus on the Kling-Gupta Efficiency (KGE; Gupta et al., 2009). KGE values can vary between -∞ (poor agreement between modelled and observed data) and 1 (perfect agreement between modelled and observed data). Current results show an overall good performance over the entire European domain. *Figure 9* illustrates the performance as a function of runoff coefficient and wetness index. This analysis points towards limitations of the current model setup and requirements for improvements. Similarly, *Figure 10* illustrates the spatial variability of the KGE performance.

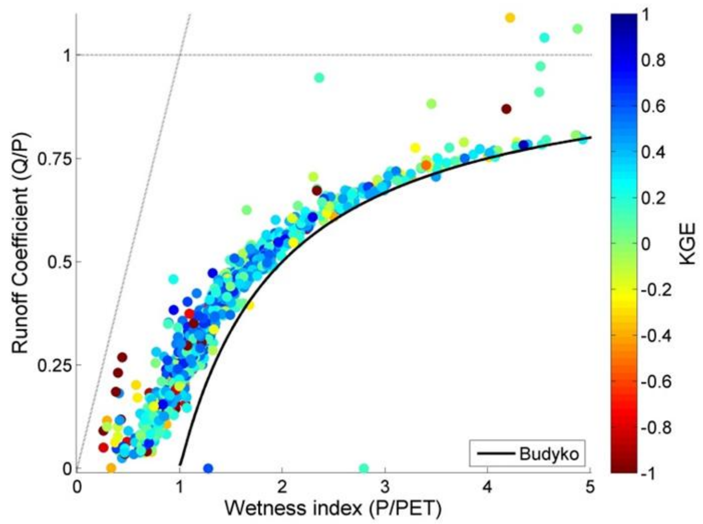


Figure 9: Performance analysis based on the Budyko framework

SMHI have also worked on improving the representation of WFDEI forcing data (i.e. precipitation and temperature) in the model in order to adequately estimate the sub-basin mean fluxes. The difference between WFDEI and E-HYPE spatial resolution could result into under- / over-estimation of extremes; hence different spatial interpolation methods are investigated and their corresponding runoff is assessed. This process has already been completed. Substantial effort is currently also being input to improving the parameterization of the model and further improving the models performance over the hindcast period. The new improved E-HYPE version will become operational (hence replace the current existing) and will be tested within EUPORIAS.

In particular, SMHI have refined the E-HYPE hydrological model which will have an impact on both the models initialisation and performance. The current operational version of E-HYPE is based on HYPE model v. 4.3.1 (following version 4.1.0), which overcomes some technical problems and allows a better initialisation of state model variables.

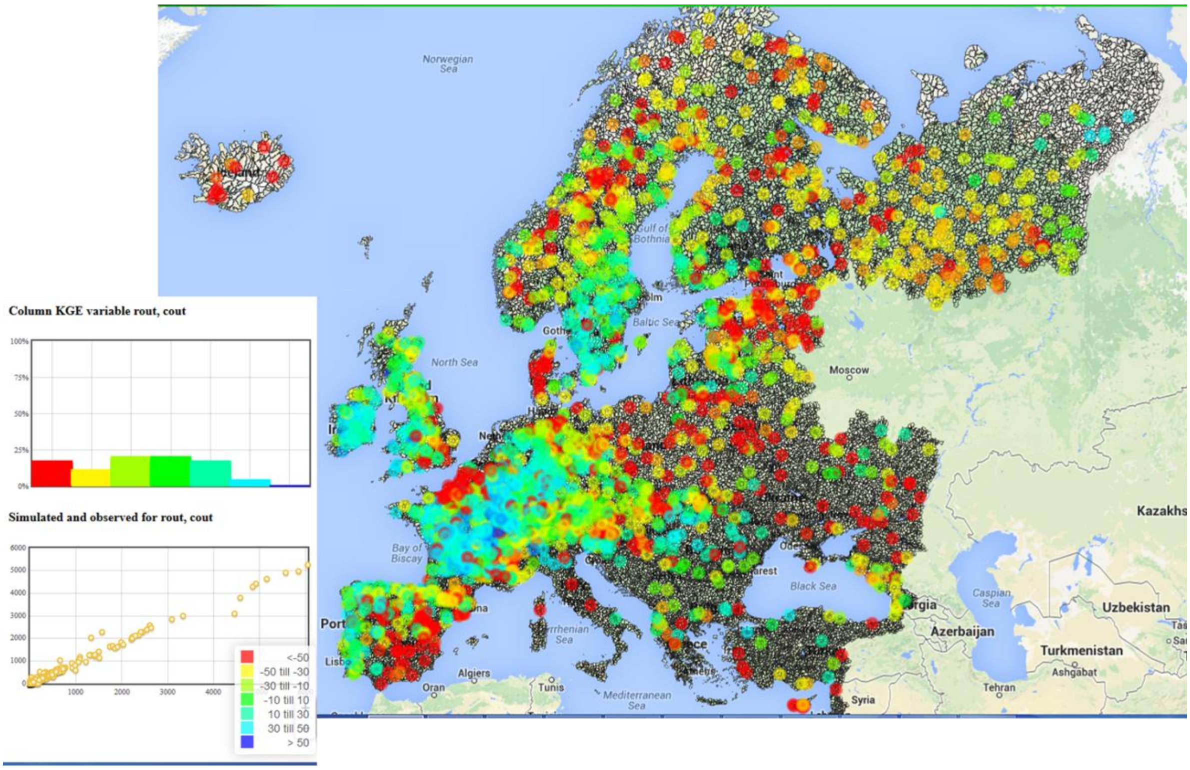


Figure 10: KGE EHYPE model performance

In particular:

- Spin up can be used to estimate the initial water level in the dams;
- Alternative/additional processes are now present allowing a better representation of the hydro-climatic processes; hence extraction of additional variables;
- Introduction of new performance criteria which are useful in the evaluation of hydrological forecasts; and
- In parallel to model developments, SMHI have been working on the coupling of forcing data (mainly precipitation) and model.

EDF's overall goal is to run their hydrological model (MORDOR) with seasonal forecasts, in order to test seasonal forecasts for hydropower production management applications. The model will be run on a large set of 35 watersheds in France (*Figure 11, Table 3*), for which temperature, precipitation and river flow data are available (from both EDF's own and confidential data sources). The idea is to run and/or compare different forecasting methods to provide river flow forecasts from 1 to 3-6-12 months ahead.

The study will be based on previous achievements using ECMWF VarEPS's monthly forecasts (Dubus, 2012), and results from Météo-France (Singla et al. 2012), and should consist in a comparison of different forecasting strategies. EDF will use different daily mean temperature and precipitation forecasts on these 35 locations obtained from:

1. bias-corrected ECMWF S4 forecasts of Z700 & Z1000 + an analog method ➔ dataset analog_ECMWF
2. Météo-France's SIM model's (8 km grid) ➔ dataset SIM_MF
3. bias-corrected Météo-France S3 forecasts of Z700 & Z1000 + an analog method ➔ dataset analog_MF

These three datasets should then feed the MORDOR hydrological model to get daily river flow forecasts up to 3-6-13 months (depending on the available data from the different SF systems). The basic objective is to run option 1 above, and possibly options 2 and 3 if there is enough time. In any case, the river flow forecasts (from 1 and possibly 2/3) will be compared to observations and to Météo-France river flow forecasts from their S3/SIM model.

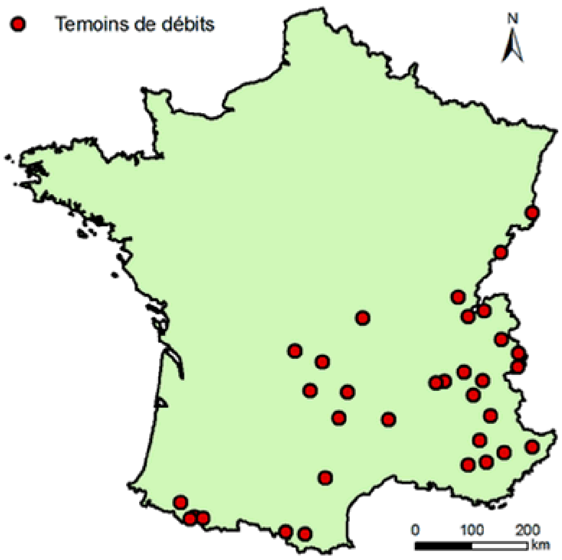


Figure 11: Map showing the set of 35 watersheds to be simulated using MORDOR

ECMWF System 4 seasonal hindcasts have been downloaded (1981-2010) for Z700, Z1000 over the North Atlantic / Europe sector. The reason for applying an analog method (1 above) is that direct temperature and precipitation forecasts at the local scale are not skilful enough.

Table 3: Watersheds to be simulated using MORDOR

FRHI	Rhin	Rhin@Kembs
FALN	Alpes Nord	Q1 : Q0001 : Fond de France Q2 : Q3660 : Doubs aux Brenets Q3 : W1410010 : Isère à Grenoble Q4 : Arve@Arthaz
FRHOI	Rhône-Isère	Q1 : BviRhone@Bugey Q2 : Ain@Vouglans Q3 : Q0024 (IsereàPizancon) Q4 : Q0037 (BourneàPontManne) Q5 : Rhone@Scex
FDT	Durance-Tinée	Q1:Q0314 (TinéePontLune) Q2:BviDurance@Cadarache Q3:BviDurance@Escale Q4:BviVerdon@SteCroix
LDD	Drac-Durance	Q1 : Drac@Sautet Q2 : Durance@SerrePoncon Q3 : Verdon@Castillon
LALP	Alpes	Q1 : Isère ValdIsère (Q3302) Q2 : Romanche@Chambon Q3 : Cenise@MontCenis Q4 : Doron@Roselend Q5 : Ain@Vouglans
		MASSIF-CENTRAL
DOR	Dordogne	Q1 : Dordogne@Bort Q2 : Q2454 : Cere à Cantales Q3 : Q2154 : Vezere à Montceaux
TRU	Truyère	Q1 : Truyere@Grandval Q2 : Q2704 : Lot à Castelnau
LOI	Loire	Q1 : Allier@StYorre Q2 : ChassezacPiedBorne (Q0384) Q3 : Arn à StPeyre (Q1643)
		PYRENEES
FPYR	Pyrénées	Q1 : Q1043 (Aspe à Pont d'Escot) Q2 : Arriougrand@Migouelou Q3 : Aston@Laparan Q4 : Aude@Puyvalador Q5 : Brousset@Pontdecamps Q6 : Gave@LaRaillere
LPYR	Pyrénées	Q1 : Q1043 (Aspe à Pont d'Escot) Q2 : Arriougrand@Migouelou Q3 : Aston@Laparan Q4 : Aude@Puyvalador Q5 : Gave@LaRaillere

The analog method for producing precipitation and temperature forecasts is that briefly described in the proceedings of the ECMWF 2012 seminar (also used in Workpackage 21): http://www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2012/Dubus.pdf. This stage includes an evaluation of the Z700/Z1000 forecasts with respect to reanalysis (ERAinterim and/or NCEP), and an evaluation of the temperature/precipitation forecasts on the watersheds.

The impact (hydrological) model will be run by the EDF operational hydrometeorological forecasting division once they are provided with the analog dates and temperature/precipitation forecasts. The joint analysis of results will include comparison with Meteo-France's hydrological forecasting chain.

The MORDOR hydrological model needs initial conditions (water/snow stocks, water level in reservoirs, etc) that will be obtained from observations. It then needs temperature and precipitation forecasts, which will be obtained from the different options described above. Therefore, there is no specific need for a spin-up of the model.

However, in order to get the best possible results, the input variables of the analog method and/or the MORDOR model need to be as good as possible. The main requirement for this study is then to have geo-potential height, and/or temperature and precipitations forecasts which are bias corrected. A further calibration of the temperature and precipitation forecasts would provide added value.

EDF’s approach to initialisation is therefore:

1. to use bias-corrected / calibrated Z700 and Z1000 forecasts over the North-Atlantic / European sector, from ECMWF S4 and possibly Météo-France S3;
2. to get help on calibration of temperature & precipitation forecasts (either scientific/technical advice or ready to use R routines)

Figure 12, from Dubus (2012) illustrate the effect of initialising the MORDOR model with numerical weather prediction (NWP) values of temperature and precipitation, obtained from the ECMWF monthly forecast system, combined with EDFs analog method. It shows, for each week from October 2004 to April 2010, different forecasts of the monthly cumulated inflow for the Drac river at Sautet hydropower station. The grey curves correspond to the climatological distribution of river flow; the red curves are the forecasts obtained with the MORDOR hydrological model, when it is forced with historical time series of temperature and precipitation (~50 years); the blue curves are the MORDOR forecasts, forced with the temperature and precipitation forecasts obtained from ECMWF monthly forecasts + EDF’s analogs method; green dots are the observations. In both cases, the MORDOR model is initialised with observed conditions for snow cover, and reservoir levels etc. The only difference between red and blue curves is that in the latter case, the actual dynamical forecasts are taken into account, rather than the climatological view based on historical time series. The plot shows that, particularly during in autumn, when the snow stock has melted, initial conditions weakly constrain the hydrological model (contrary to late winter/early spring, when the flow is mostly determined by the snow stock in mountains), and that the current NWP forecast brings valuable information, and strongly reduces the uncertainty in the forecast.

On average over the 43 french watersheds where this model was tested, there is an improvement up to week four, in particular for the most extreme quantiles (*Figure 13*).

A preliminary study has shown that for seasonal lead times, the extension of this method based on Z1000/Z700 large scale forecasts and the analogs methods may have a positive impact as well, but the bias in the Z1000/Z700 forecasts need to be removed.

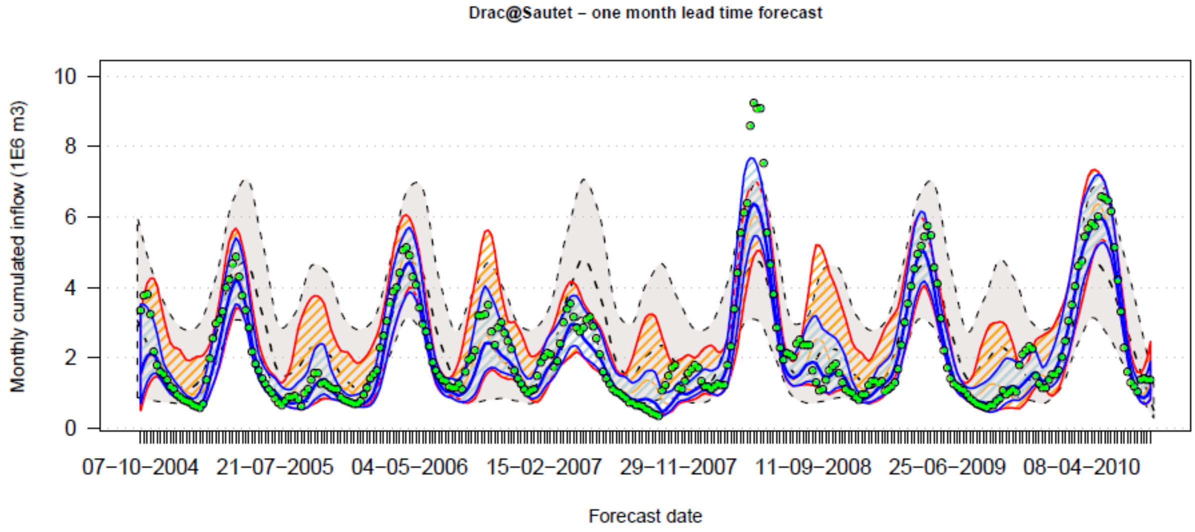


Figure 12: weekly forecasts of the monthly cumulated inflow on Drac river at Sautet hydropower station. Grey: climatological distribution of the inflows; red: MORDOR model forced by historical time series of temperature and precipitation; blue: MORDOR model forced by ECMWF monthly forecasts (Z700/Z1000) + analogs method; green dots: observations

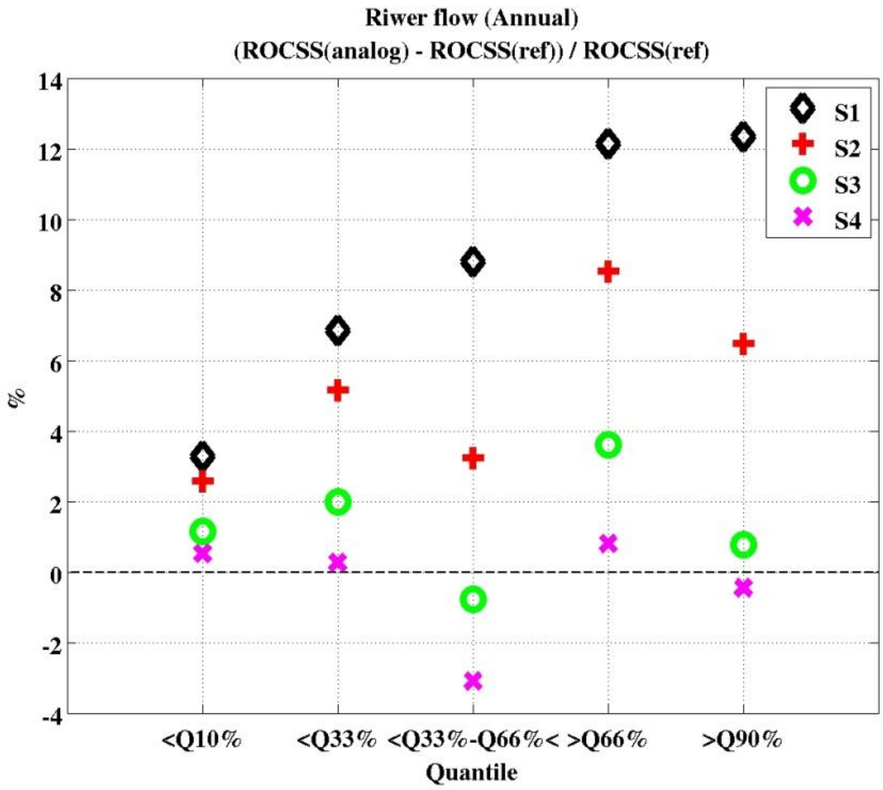


Figure 13: Improvement in Relative Operating Characteristic Skill Score (ROCSS) for river flow forecasts over 43 watersheds in France, with the ECMWF/analogs method (analog) compared to the use of historical temperature/precipitation time series (ref) for different quantiles of the forecast distribution, and for the 4 weeks of the monthly forecast

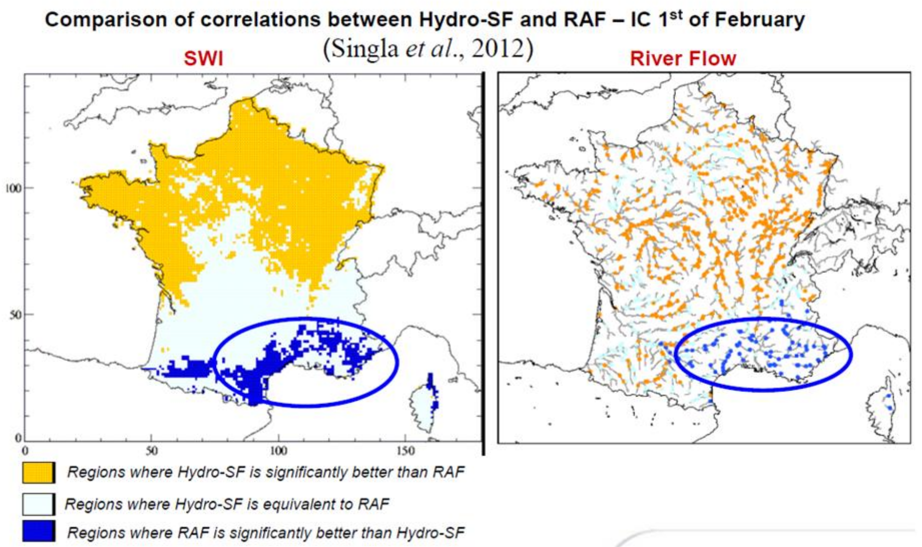


Figure 14: MeteoFrance hydrological forecast results for spring (MAM), comparing the HydroSF and RAF methods (Singla et al. 2012), with initial conditions for 1st February. Results are compared using spatial representation of time correlations over 1960-2005. left panel: soil water index; right panel: river flow

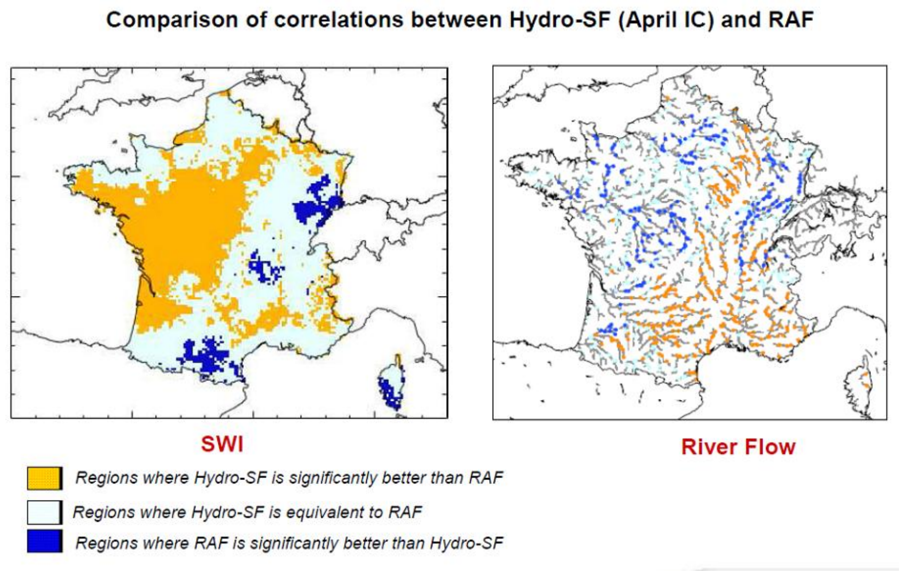


Figure 15: MeteoFrance hydrological forecast results for summer (JJA), comparing the HydroSF and RAF methods (Singla et al. 2012), with initial conditions for 1st April. Results are compared using spatial representation of time correlations over 1960-2005. left panel: soil water index; right panel: river flow

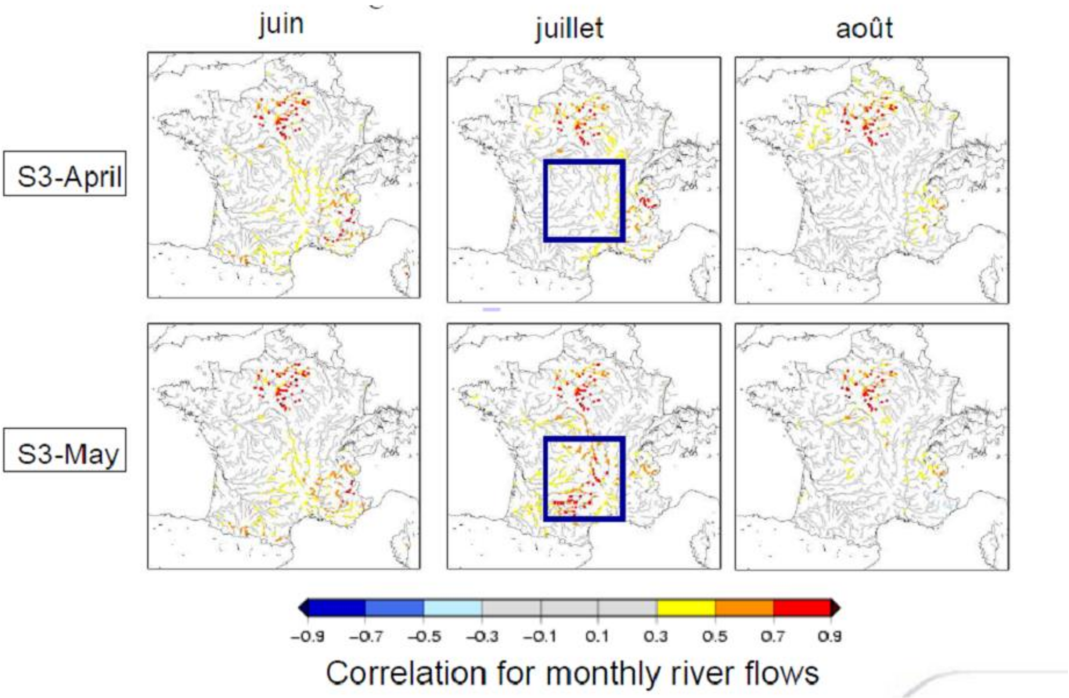


Figure 16: Summer results (JJA) for the Hydro-SF method, showing the spatial correlation for monthly river flows as a function of initial conditions (top row – April initial conditions; bottom row – May initial conditions) over the months June, July and August

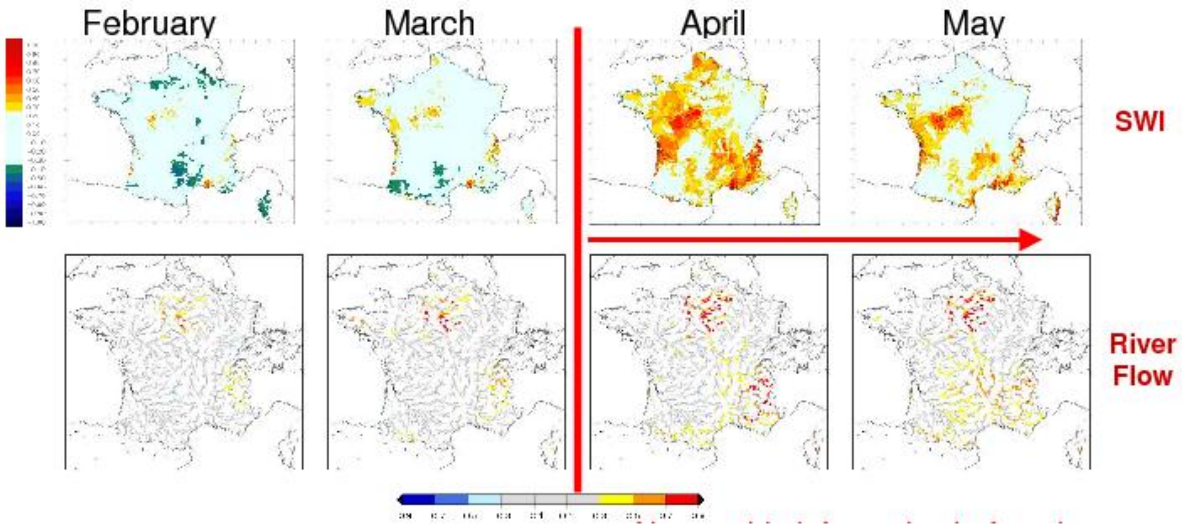


Figure 17: Illustrating the effect of initial conditions on MeteoFrance hydrological simulations. Correlation for soil water index (SWI) and river flows over the 1979-2007 period (HYDRO-SF/ARPEGE-S3) using different initial conditions for the summer forecast (JJA). correlations > 0.3 are significant.

Meteo-France started the production of the impact variable hindcast using the downscaled data provided in the WP21. Two seasons were investigated first; the Spring and Summer seasons as they are crucial in term of water resource management. These results will be extended in the next for other periods of interest. A hindcast of 30 years was issued for both the Soil Wetness Index (on an 8 km grid over France) and the River Flow (for more than 900 stations along the rivers). These hindcasts (termed Hydro-SF) were evaluated against a relevant hydrological reanalysis (so called SIM reanalysis) which was validated against observations (also making a contribution to D23.4). Both probabilistic and deterministic scores were computed. In addition, to demonstrate the added value brought by such a forecasting suite, a specific experiment using random atmospheric forcing (and so call RAF experiment) was prepared. The same scores for these experiments were also performed.

The SWI and River Flow hindcasts for MAM and JJA are available to Project partners for case studies. *Figures 14* and *15* illustrate that the relative value of the full hydrological forecast method (Hydro-SF) compared to using random forcings (RAF) varies with season, region and variable.

Figure 16 illustrates the effect of initial conditions, showing that better simulation of JJA river flows results when using May initial conditions compared to April initial conditions, and in general, better river flow prediction is found in July for the Massif Central region. Similarly, *Figure 17* shows a clear improvement in hydrological forecast correlations when driving the model with April, rather than March initial conditions; in general, little usable information appears to be provided from the impact forecast before the beginning of April.

Wageningen University has worked on setting up the VIC and LPJml models for hydrological forecasts in Europe (initially; later perhaps also for East Africa). Both have been run with WFDEI forcing to provide a base run for initialisation of the hindcasts and as reference. The base run has been validated against an observed discharge dataset based on GRDC augmented for some basins. Presently, WU are developing scripts to perform the hindcast runs starting with System4 data.

The System4 hindcast has been downloaded from the ECOMS-UDP, re-gridded to 0.5^{o}, reformatted to NETCDF conform protocol, a land mask applied and re-organised to one annotated file (with all variables) for each forecast (i.e. a total of 5400 files = 360 7-month forecasts x 15 members). These can be made available to other partners upon request. A bias and skill assessment has been performed for selected grid boxes across Europe (see WP31/32 reports).

Figure 18 shows examples of the performance/validation of both VIC and LPJml models for a selected basin. The highly variable temporal coverage of observed data in GRDC may be somewhat problematic for validating the model outputs.

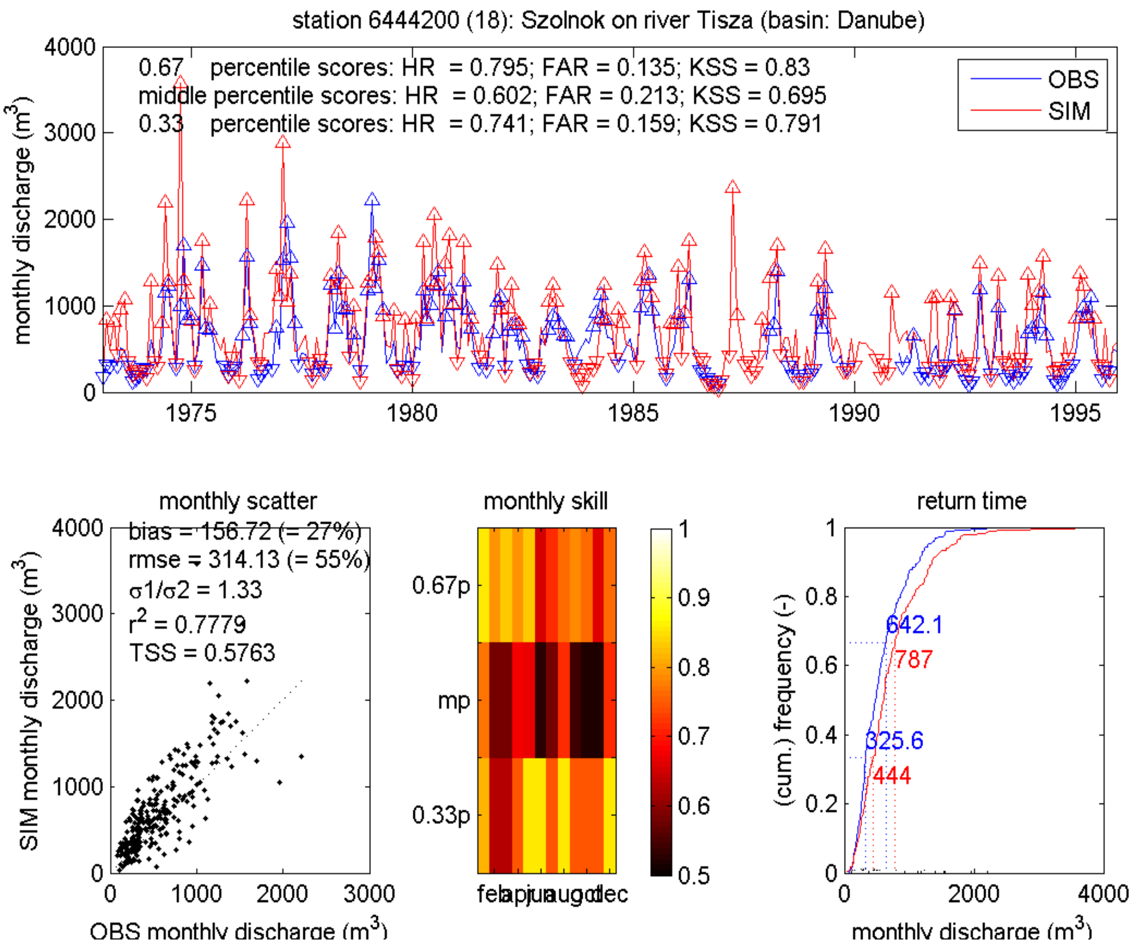


Figure 18: Example discharge validation LPJml forced by WFD for the Tisza. The bottom-centre figure presents the skill (HansenKuiper score) for above/below normal discharges using the reference forcing, setting the target for the hindcast ensemble

4.3 Progress in model initialisation for the agriculture sector

The University of Leeds have been getting GLAM-maize ready to simulate yields in East Africa. This has involved checking through the code and contacting the different people working on GLAM-maize and making sure that the model version has all the relevant updates and recent bug fixes. Appropriate parameter values for maize in East Africa are also being assessed, and a strategy for dealing with the many different maize varieties grown across the region is being developed. In collaboration with Wageningen, a likely approach will be to categorise the huge number of maize varieties into a few distinct groups and then find suitable parameter values for each group of maize varieties. As discussed during the workshop and noted in D23.1 and D23.2, model initialisation is not likely to have a significant impact in the GLAM model for crop production in East Africa. The approach proposed will be to assume that the soil is completely dry at the end of the dry season/start of the rainy season, and observed weather from the start of the rainy season until the forecast date. This is a valid approach since generally-speaking; in tropical regions the soils are dry at the beginning of the crop season. However, as noted in D23.1, for crop modelling studies in Europe the impact of snow cover on crop (winter wheat) productivity may be significant and therefore crop models could be sensitive to initial snow conditions.

The Met Office has begun to use GloSea5 seasonal hindcast data with the JULES impact model, and to assess the impact of data disaggregation on model results. The latter work is particularly important since sub-daily driving data, needed for running the JULES model, may only be available from a more limited set of seasonal hindcasts. The Met Office has also been working on developments to the JULES model, including on the JULES_crop crop model, on including an inline forcing data disaggregator, on adding a crop product pool, and enabling spatially varying heights of forcing data in the model. Further work is also ongoing to include an irrigation scheme which will eventually include both demand (crop requirement) and supply (removal of water from rivers and groundwater). The global-scale performance of the JULES-crop model has been assessed, driving the model with both CRU-NCEP and WFDEI forcing data.

In particular, the Met Office performed a simulation using JULES-crop with CRU-NCEP-forcing data for 1960-2009 with a 50-year spin up. CRU-NCEP has been used for initial tests rather than WFDEI because it takes less time to run, due to lower resolution, and also covers a longer time period (which will also lower the error on the correlations). The results from this simulation were then used to initialise a large number of smaller runs, also forced by CRU-NCEP. Each of the smaller runs finishes at the end of the calendar year, and the smaller runs start (for example):

- 15th Jan 1960, 15th Jan 1961, ..., 15th Jan 2009
- 15th Feb 1960, 15th Feb 1961, ..., 15th Feb 2009

The initialisation used for each run was the climatology of the 21 days around the start date from the original run. In other words, all runs starting on the 15th January were initialised from the mean of 5th-25th January (inclusive) for 1960-2009. No spin up was used for these runs (although the first two weeks of each of these is technically spin up, since the 'main run' starts on the 1st of the following month, so that monthly means can be outputted).

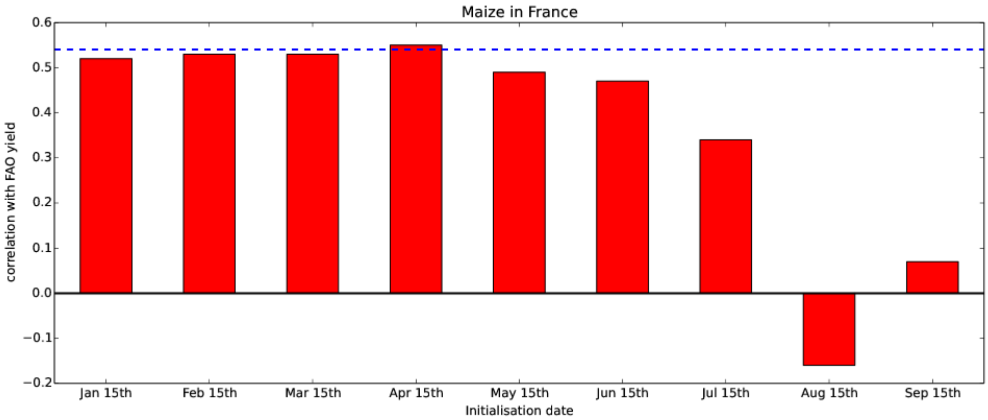


Figure 19: Comparison between observed (FAO) and modelled maize yields for France, using JULES-crop simulations with CRU-NCEP forcing for 1960-2009, as a function of initialisation. Blue dotted line: Correlation (Pearson correlation coefficient) of simulated yield versus FAO yield for CRU-NCEP-forced JULES-crop run for 1960-2009. Red bars: Correlation (Pearson correlation coefficient) of simulated yield versus FAO yield for the CRU-NCEP-forced JULES-crop runs initialised by climatology each year. Each bar is labelled by the start date of the runs

We show an example of the results from this study for Maize in France (*Figure 19*). The sowing date used for maize in France is beginning of May, with a harvesting date of the middle of October (Sacks et al. The results show generally little impact of initialising the model with climatological (rather than “actual”) weather data in the early part of the season, until at least June/July when the correlation with observed yields begins to degrade.

Initial assessment of results for maize in Ethiopia showed no significant correlation for the full CRU-NCEP simulation (correlation = 0.2, not significantly different from zero at 95% confidence level). This could be because the annual variability in Ethiopian yield depends on a wider range of factors than simply meteorological drivers, inaccuracies in the FAO observations for Ethiopia (as discussed in Greatrex (2012)), the parameters for maize used not being representative of Ethiopian maize, or a dependence on effects not modelled (e.g. water stress, high temperature stress).

This study seems to suggest that JULES-crop maize runs can be initialised with climatology, without suffering a large reduction in skill, assuming that the results for France are more generally applicable. Assuming the model is driven by seasonal forecasts which have been bias corrected to WFDEI, this also implies that JULES-crop maize simulations runs for EUPORIAS can be reliably initialised using WFDEI climatology. The Met Office have already performed a global WFDEI-forced JULES-crop run with the right outputs to allow calculation of initialisation files.

The Met Office and the University of Leeds have experimented with forcing GLAM with GloSea5 hindcast data for groundnut over West Africa, using the procedure formulated in Nicklin (2013) for use with the ECMWF System 3 hindcast. The model was initialised by assuming no soil moisture and no crop at the start of the rainy season and then running with observed weather from the start of the rainy season up to each hindcast start date, using the GPCP 1DD dataset for rainfall and ERA-Interim for temperature and solar radiation. The GloSea5 hindcast set had three start dates (25th April, 1st May and 9th May), 10 ensemble members for each start date, each with a run length of seven months. No bias correction was applied to the seasonal forecast and each ensemble member was used to drive a separate GLAM run.

Model runs were carried out both with and without yield calibration, using a yield gap parameter (YGP) that acts to reduce the leaf area of the crop in order to reduce crop yields. It was not possible to fully calibrate yield in the north-west of the domain, where the observed yields were greater than the modelled yield even with no reduction to the modelled yield from the YGP. This could be due to a spatial bias in the seasonal forecast precipitation field, irrigation not being included in the model or a lack of spatial resolution in the yield observations.

We carried out a preliminary evaluation using correlation maps, ROC curves, ROC scores, reliability diagrams and sharpness diagrams. These results indicate some skill in the prediction of crop failures although the forecasts were generally overconfident. Correlations between forecasted and observed groundnut yields were found to be typically higher in the western and northern regions and small or negative in the eastern and southern parts of the domain.

Wageningen University has worked on setting up the CGMS/WOFOST and LPJml models for agricultural forecasts in East Africa (with potential to perform later simulations for Europe). Both have been run with WFD-EI to provide a base run for initialisation the hindcasts and as reference. The base run is presently being validated against an observed crop production dataset based on FAOSTAT, augmented with sub-national statistics for Kenya and Tanzania. Initially we will focus on Maize production, later other crops may be added (e.g. wheat, millet). Presently, we are developing scripts to perform the hindcast runs starting with System4 data.

The System4 hindcast has been downloaded from the ECOMS-UDP, re-gridded to 0.5^{o}, reformatted to NETCDF conform protocol, a land mask applied and re-organised to one annotated file (with all variables) for each forecast (i.e. a total of 5400 files = 360 7-month forecasts x 15 members). These can be made available to other partners upon request. A bias and skill assessment has been performed for selected grid boxes across East Africa (see WP31/32 reports).

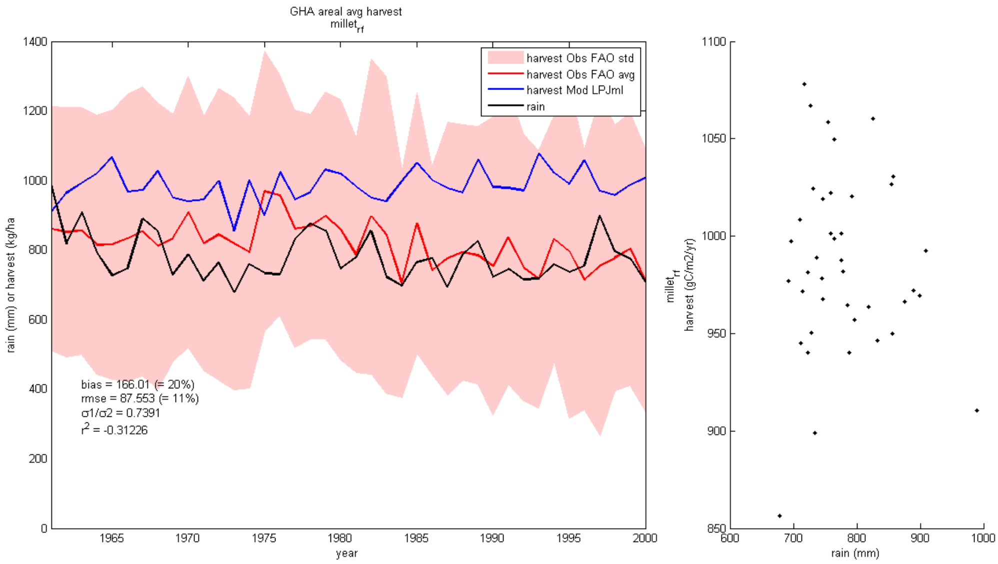


Figure 20: Observed (FAOSTAT) Millet production for GHA compared to modelled by LPJml (forced by WFD). Major driver rain is also plotted

We show examples of the performance/validation of both WOFOST and LPJml models for the whole of the Greater Horn of Africa (GHA) using national statistics from the FAO only (*Figure 20*). Several problems arise with the observed statistics that are currently being worked on: calendar issues (different databases use different calendars, causing shifts in the series), aggregation issues (reported production at different administrative levels do not always add up to total of next higher level; administrative reorganizations; see *Figure 21*), only part of interannual variability is climate related (check for major alternative causes, e.g. political upheaval), records at all levels exhibit considerable gaps. Work is underway to overcome these issues as far as is possible. We may also have to consider assessing the skill of hindcasts not only with respect to observed crop yields, but also with respect to those simulated in the reference run forced with WFDEI data. *Figure 20* illustrates the strong link between both observed and modelled yields and precipitation changes, which supports the argument to not need long initialisation models for crop models in East Africa proposed by the University of Leeds.

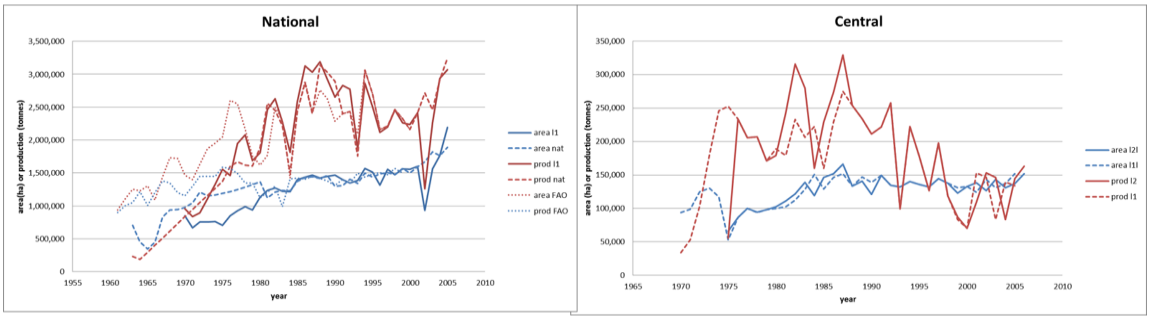


Figure 21: Inconsistencies between Maize production in Kenya as reported at 4 different levels. By resp FAO, national total, sum of l1 (8 provinces) and sum of l2 (46 districts) subdivision

4.4 Progress in model initialisation for the forestry sector

The University of Lund have started to test the impact models (forestry sector) and asked SMHI to provide relevant S2D model data. They are proposing, and working on a forestry case study focused on planning of harvesting activities in the winter season. The reason for this is twofold. Firstly, the best prospects for achieving some skill at the seasonal time-scale are likely to be in predicting winter NAO conditions, which in turn have a high correlation to Scandinavian winter temperature (and precipitation). Secondly, winter-time harvesting poses major logistic and environmental challenges in which climate and weather is a major component. Thinning and clear cutting are commonly practiced in winter when the soil is frozen. However, rainy autumns and mild winters make it difficult to harvest forest standing on wet soils, as the heavy machines cause driving damage (soil compaction, deep wheel tracks that permanently changes water flow). Seasonal forecasts of enough skill will provide a tool to substantially improve the planning of harvesting and associated logistic of forest companies.

Lund University and SMHI have been collaborating in developing the local surface climate downscaling tool/model LDCLIM, which integrates existing modeling components into a tool for describing local climate conditions and ecological impacts as the combined effect of regional atmospheric forcing and local physiographical and biological conditions. For the case study, the following variables are required as input:

- daily 2-metre temperature and precipitation (after bias correction)
- 3-hourly temperature, specific humidity, wind speed (u and v) at the lowest (preferably) model level of the seasonal forecasting system, and additionally downward radiation shortwave (preferable divided into direct and diffuse component) and longwave, as well as rain and snow intensity and surface pressure.

The University of Lund approach has been to work on and apply different model components in parallel, both the fully integrated LDCLIM model and the model component (LPJ-Guess) for forestry and ecological impacts.

4.5 Progress in model initialisation for the transport sector

Predictia will use the numerical model METRo (Model of the Environment and Temperature of Roads) to forecast road conditions. This model has an initialisation phase in which, for each forecast, an initial road temperature profile is needed. To produce such a profile METRo uses road temperature observations at the surface and subsurface from the last two days to force the heat-conduction model. In order to use METRo to perform forecasts on the seasonal timescale the seasonal model will take the role of those observations. Moreover, according to experimentation comparing forecasts with full/small initialisation periods (Crevier & Delage 2001), only small differences in road temperature should appear after a 24-h forecast. Hence initialisation is not considered to be critical for this application.

References

Bartholomé, E., Belward, A.S., Achard, F., Bartalev, S., Carmona Moreno, C., Eva, H., Fritz, S., Grégoire, J.-M., Mayaux, P., Stibig, H.-J., 2002. GLC 2000 Global Land Cover mapping for the year 2000. European Commission, DG Joint Research Centre, EUR 20524 EN.

Bernhard, L., Liermann, C., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rödel, R., Sindorf, N., Wisser, D., 2011. Global Reservoir and Dam (GRanD) database - Technical documentation version 1.1.

Challinor AJ; Slingo JM; Wheeler TR; Doblas-Reyes FJ (2005) Probabilistic simulations of crop yield over western India using the DEMETER seasonal hindcast ensembles, TELLUS A, 57, pp.498-512.

Cosgrove, B. A., and Coauthors, 2003: Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS). J. Geophys. Res. 108, 8845, doi:10.1029/ 2002JD003316.

Crevier, LP and Delage, Y (2001) METRo: A New Model for Road-Condition Forecasting in Canada, Journal of Applied Meteorology 40, 2026-2037.

Davey M, Brookshaw A (2011) Long-range meteorological forecasting and links to agricultural applications, Food Policy 36, S88–S93.

Demeritt, D, Sebastien Nobert, Hannah L. Cloke, Florian Pappenberger (2013) The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management, Hydrological Processes, 27, 147–157, DOI: 10.1002/hyp.9419

Demirel M.C., Booij M.J., Hoekstra A.Y. 2013. Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models. Water Resources Research doi: 10.1002/wrcr.20294

Dessai, S, Soares, MB (2013) EUPORIAS (Grant agreement 308291) Deliverable 12.1, Systematic literature review on the use of seasonal to decadal climate and climate impacts predictions across European sectors, University of Leeds, Leeds, UK.

Dubus, L. 2010. Practices, needs and impediments in the use of weather/climate information in the electricity sector In: TROCCOLI, A. (ed.) Management of Weather and Climate Risk in the Energy Industry. Springer

Dubus, L. 2012. Monthly and seasonal forecasts in the French power system. ECMWF Seminar, Personal presentation.

Dubus, L. 2013. Use of monthly and seasonal to decadal forecasts in the Energy sector: EDF's experience modelling in Europe. Tellus Series a-Dynamic Meteorology and Oceanography, 57, 476-487.

Falloon, P, Betts R, Wiltshire A, Dankers R, Mathison C, McNeall D, Bates P, Trigg M (2011). Validation of river flows in HadGEM1 and HadCM3 with the TRIP river flow model. Journal of Hydrometeorology,12,1157-1180. doi: 10.1175/2011JHM1388.1

Falloon P, Fereday D, Stringer N, Williams K, Gornall J, et al. (2013) Assessing Skill for Impacts in Seasonal to Decadal Climate Forecasts. J Geol Geosci 2:e111. doi: 10.4172/2329-6755.1000e111

Greatrex, H. (2012). "The application of seasonal rainfall forecasts and satellite rainfall estimates to seasonal crop yield forecasting for Africa", PhD Thesis, University of Reading, UK

Gupta, H. V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P., and Yeh, P.: Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results, J. Hydrometeorol., 12, 869–884, doi:10.1175/2011JHM1324.1 , 2011.

Hansen, J. W., Mason, S. J., Sun, L. & Tall, A. 2011. Review of seasonal climate forecasting for agriculture in Sub-Saharan Africa. Experimental Agriculture,47, 205-240.

Hansen, J.W., Challinor, A., Ines, A.V.M, Wheeler, T., Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges. Climate Research 33:27-41.

Koster, R. D., et al. (2010), Contribution of land surface initialisation to subseasonal forecast skill: First results from a multi-model experiment, Geophys. Res. Lett., 37, L02402, doi:10.1029/2009GL041677.

Lehner, B., Döll, P., 2004. Development and validation of a global database of lakes, reservoirs and wetlands. J. Hydrol. 296, 1–22.

Lehner, B., Verdin, K., Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos, Trans. AGU 89, 93–94.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., Arheimer, B., 2010. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. Hydrol. Res. 41, 295–319.

Macleod, D., Caminade, C. & Morse, A. 2012. Useful decadal climate prediction at regional scales? A look at the ENSEMBLES stream 2 decadal hindcasts. Environmental Research Letters, 7, 044012

Marshall, N. A., Gordon, I. J. & Ash, A. J. 2011. The reluctance of resource-users to adopt seasonal climate forecasts to enhance resilience to climate variability on the rangelands. Climatic Change, 107,511-529.

Meinke, H., Nelson, R., Kokic, P., Stone, R., Selvaraju, R., Baethgen, W. (2006): Actionable climate knowledge – from analysis to synthesis. Climate Research 33, 101-110.

Nachtergaele, F., van Velthuizen, H., Verelst, L., Wiberg, D., 2012. Harmonized world soil database version 1.2. FAO, Rome and IIASA, Laxenburg, Austria.

Nicklin, K (2013) Seasonal crop yield forecasting in semi-arid West Africa Ph.D thesis University of Leeds.

Siebert, S., Döll, P., Hoogeveen, J., Faures, J.-M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrol. Earth Syst. Sci. 9, 535–547.

Shukla, S. and Lettenmaier, D. P.: Seasonal hydrologic prediction in the United States: understanding the role of initial hydrologic conditions and seasonal climate forecast skill, Hydrol. Earth Syst. Sci., 15, 3529-3538, doi:10.5194/hess-15-3529-2011, 2011.

Shukla, S., Sheffield, J., Wood, E. F., and Lettenmaier, D. P.: On the sources of global land surface hydrologic predictability, Hydrol. Earth Syst. Sci. Discuss., 10, 1987-2013, doi:10.5194/hessd-10-1987-2013, 2013.

Singla, S. et al (2012) Predictability of soil moisture and river flows over France for the spring season, *Hydrol. Earth Syst. Sci.*, 16, 201-216, 2012.

van der Linden P, Mitchell JFB (2009) ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, Exeter EX1 3PB, UK.

Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J.C., Bellouin, N., Boucher, O., Best, M., 2011. Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. J. Hydrometeorol. 12, 823–848.